Electronic Colloquium on Computational Complexity, Report No. 105 (2015)

Towards a Characterization of Approximation Resistance for Symmetric CSPs

Venkatesan Guruswami* Euiwoong Lee†

Computer Science Department
Carnegie Mellon University
Pittsburgh, PA 15213.

Abstract

A Boolean constraint satisfaction problem (CSP) is called approximation resistant if independently setting variables to 1 with some probability α achieves the best possible approximation ratio for the fraction of constraints satisfied. We study approximation resistance of a natural subclass of CSPs that we call *Symmetric Constraint Satisfaction Problems (SCSPs)*, where satisfaction of each constraint only depends on the number of true literals in its scope. Thus a SCSP of arity k can be described by a subset $S \subseteq \{0, 1, \ldots, k\}$ of allowed number of true literals.

For SCSPs without negation, we conjecture that a simple sufficient condition to be approximation resistant by Austrin and Håstad is indeed necessary. We show that this condition has a compact analytic representation in the case of symmetric CSPs (depending only on the gap between the largest and smallest numbers in S), and provide the rationale behind our conjecture. We prove two interesting special cases of the conjecture, (i) when S is an interval (i.e., $S = \{i \mid l \leqslant i \leqslant r\}$ for some $l \leqslant r$) and (ii) when S is even (i.e., $s \in S \Leftrightarrow k - s \in S$). For SCSPs with negation, we prove that the analogous sufficient condition by Austrin and Mossel is necessary for the same two cases, though we do not pose an analogous conjecture in general.

*Supported in part by NSF grant CCF-1115525. guruswami@cmu.edu
†Supported by a Samsung Fellowship and NSF CCF-1115525. euiwoonl@cs.cmu.edu

ISSN 1433-8092

1 Introduction

Constraint Satisfaction Problems (CSPs) are among the most fundamental and well-studied class of optimization problems. Given a fixed integer k and a predicate $Q \subseteq \{0,1\}^k$, an instance of CSP(Q) *without negation* is specified by a set of variables $X = \{x_1, \ldots, x_n\}$ on the domain $\{0,1\}$ and a set of constraints $\mathscr{C} = \{C_1, \ldots, C_m\}$, where each constraint $C_j = (x_{j,1}, \ldots, x_{j,k})$ is a k-tuple of variables. An assignment $X \mapsto \{0,1\}$ satisfies C_j if $(x_{j,1}, \ldots, x_{j,k}) \in Q$. For an instance of CSP(Q) *with negation*, each constraint C_j is additionally given offsets $(b_{j,1}, \ldots, b_{j,k}) \in \{0,1\}^k$ and is satisfied if $(x_{j,1} \oplus b_{j,1}, \ldots, x_{j,k} \oplus b_{j,k}) \in Q$ where $\oplus$ denotes the addition in $\mathbb{F}_2$. The goal is to find an assignment that satisfies as many constraints as possible.

CSPs contain a large number of famous problems such as Max-SAT (with negation), and Max-Cut / Max-Set-Splitting (without negation) by definition. They have always played a crucial role in the theory of computational complexity, as many breakthrough results such as the NP-completeness of 3SAT, the Probabilistically Checkable Proofs (PCP) theorem, and the Unique Games Conjecture (UGC) study hardness of a certain CSP.

Based on these works, recent works on approximability of CSPs focus on characterizing every CSP according to its approximation resistance. We define *random assignments* to be the class of algorithms that assign $x_i \leftarrow 1$ with probability α independently. A CSP is called *approximation resistant*, if for any $\varepsilon > 0$, it is NP-hard to have a $(\rho^* + \varepsilon)$-approximation algorithm, where ρ^* is the approximation ratio achieved by the best random assignment. Even assuming the UGC, the complete characterzation of approximation resistance has not been found, and previous works either change the notion of approximation resistance or study a subclass of CSPs to find a characterization, and more general results tend to suggest more complex characterizations.

This work considers a natural subclass of CSPs where a predicate Q is *symmetric* — for any permutation $\pi : [k] \mapsto [k]$, $(x_1, \ldots, x_k) \in Q$ if and only if $(x_{\pi(1)}, \ldots, x_{\pi(k)}) \in Q$. Equivalently, for every such Q, there exists $S \subseteq [k] \cup \{0\}$ such that $(x_1, \ldots, x_k) \in Q$ if and only if $(x_1 + \cdots + x_k) \in S$. Let SCSP($S$) denote such a symmetric CSP. While this is a significant restriction, it is a natural one that still captures the following fundamental problems, such as Max-SAT, Max-Not-All-Equal-SAT, t-out-of-k-SAT (with negation), and Max-Cut, Max-Set-Splitting, Discrepancy minimization (without negation). Except the work of Austrin and Håstad [2], many works on this line focused CSPs with negation, while we feel that the aforementioned problems without negation have a very natural interpretation as (hyper)graph coloring and are worth studying.

There is a simple sufficient condition to be approximation resistant due to Austrin and Mossel [4] with negation, and due to Austrin and Håstad [2] without negation. For SCSPs, we show that these simple sufficient conditions can be further simplfied and understood more intuitively, and suggest that they might also be necessary for and thus precisely characterize approximation resistance. We prove it for two natural special cases (which capture all problems mentioned in the last paragraph) for both SCSPs with / without negation, and provide reasons that we believe this is true at least for SCSPs without negation.

1.1 Related Work

Given the importance of CSPs and the variety of problems that can be formulated as a CSP, it is a natural task to classify all CSPs according to their computational complexity for some well-defined task. For the task of deciding satisfiability (i.e., finding an assignment that satisfies every constraint if there is one), the work of Schaefer [14] gave a complete characterization on the Boolean domain in 1978.

However, such a classification seems much harder when we study approximability of CSPs. Since the seminal work of Håstad [11], many natural problems have been proven to be approximation resistant. These examples include Max-3SAT / Max-3LIN (with negation) and Max-4-Set-Splitting (without negation), and for Boolean CSPs of arity 3, putting together the hardness results of [11] with the algorithmic results of Zwick [16], it is known that a CSP is approximation resistant if and only if it is implied by parity. However, characterizing approximation resistance of every CSP for larger arity k is a harder task. The Ph.D. thesis of Hast [10] is devoted to this task for $k = 4$, and succeeds to classify 354 out of 400 predicates.

The advent of the Unique Games Conjecture (UGC) [12], though it is not as widely believed as $\mathbf{P} \neq \mathbf{NP}$, revived the hope to classify every CSP according to its approximation resistance. For CSPs with negation, the work of Austrin and Mossel [4] gave a simple sufficient condition to be approximation resistant, namely the existence of a balanced pairwise independent distribution that is supported on the satisfying assignments of the predicate. The work of Austrin and Håstad [2] proved a similar sufficient condition for CSPs without negation, and that if this condition is not met, this predicate (both with / without negation) is *useful* for some polynomial optimization — for every such Q, there is a k-variate polynomial $p(y_1, \ldots, y_k)$ such that if we are given an instance of CSP(Q) that admits a $(1-\varepsilon)$-satisfying assignment, the altered problem, where we change each constraint C_j's payoff from $\mathbb{I}[(x_{j,1} \oplus b_{j,1}, \ldots, x_{j,k} \oplus b_{j,k}) \in Q]$ (where $\mathbb{I}[\cdot]$ is the indicator function) to $p(x_{j,1} \oplus b_{j,1}, \ldots, x_{j,k} \oplus b_{j,k})$, admits an approximation algorithm that does better than any random assignment.

Predicates that don't admit a pairwise independent distribution supported on their satisfying assignments can be expressed as the sign of a quadratic polynomial (see [2]). This motivates the study of the approximability of such predicates, though it is known that there are approximation resistant predicates that can be expressed as a quadratic threshold function and thus the sufficient condition of Austrin and Mossel [4] is not necessary for approximation resistance. Still this motivates the question of understanding which quadratic threshold functions can be approximated non-trivially.

Cheraghchi, Håstad, Isaksson, and Svensson [8] studied the simpler case of predicates which are the sign of a linear function with no constant term, obtaining algorithms beating the random assignment threshold of $1/2$ in some special cases. Austrin, Benabbas, and Magen [1] conjecture that every such predicate can be approximated better than a factor $1/2$ and is therefore not approximation resistant. They prove that predicates that are the sign of *symmetric* quadratic polynomials with no constant term are not approximation resistant.

Assuming the UGC, the work of Austrin and Khot [3] gave a characterization of approximation resistance for even k-partite CSPs, and Khot, Tulsiani, and Worah [13] gave a characterization of *strong approximation resistance* for general CSPs — strong approximation resistance roughly means hardness of finding an assignment that deviates from the performance of the random assignment in either direction (i.e., it is hard to also find an assignment saisfying a noticeably smaller fraction of constraints than the random assignment). These two works are notable in studying approximation resistance of general CSPs, but their characterizations become more complicated, which they suggest is necessary.

Without the UGC, even the existence of pairwise independent distribution supported on the predicate is not known to be sufficient for approximation resistance. Another line of work shows partial results either by using a stronger condition [7], or by using a restricted model of computation (e.g., Sherali-Adams or Lasserre hierarchy of convex relaxations) [15, 6, 5].

1.2 Our Results

Our work was initially motivated by a simple observation that for symmetric CSPs, the sufficient condition to be approximation resistant by Austrin and Håstad [2] admits a more compact and intuitive two-dimensional description in $\mathbb{R}^2$.

Fix a positive integer k and denote $[k] = \{1, 2, \ldots, k\}$. For $s \in [k] \cup \{0\}$, let $P(s) \in \mathbb{R}^2$ be the point defined by $P(s) := (\frac{s}{k}, \frac{s(s-1)}{k(k-1)})$. For any s, $P(s)$ lies on the curve $y = \frac{k}{k-1}x^2 - \frac{x}{k-1}$, which is slightly below the curve

$y = x^2$ for $x \in [0,1]$. Given a subset $S \subseteq [k] \cup \{0\}$, let $P_S := \{P(s) : s \in S\}$ and $\mathsf{conv}(P_S)$ be the convex hull of P_S. For symmetric CSPs, the condition of Austrin and Håstad depends on whether this convex hull intersects a certain curve or a point.

For SCSP(S) without negation, the condition becomes whether $\mathsf{conv}(P_S)$ intersects the curve $y = x^2$. If we let s_{min} and s_{max} be the minimum and maximum number in S respectively, by convexity of $y = \frac{k}{k-1}x^2 - \frac{x}{k-1}$, it is equivalent to that the line passing $P(s_{min})$ and $P(s_{max})$ and $y = x^2$ intersect, which is again equivalent to (see Lemma A.4)

$$\frac{(s_{max} + s_{min} - 1)^2}{k-1} \geqslant \frac{4 s_{max} s_{min}}{k}. \tag{1}$$

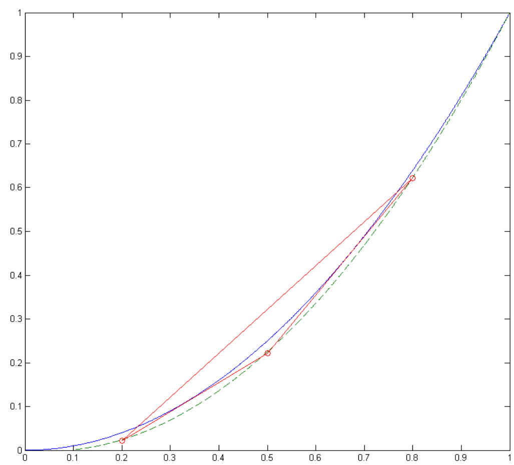

Figure 1: An example when $k = 10$ and $S = \{2,5,8\}$. The solid curve is $y = x^2$ and the dashed curve is $y = \frac{k}{k-1}x^2 - \frac{x}{k-1}$, where all $P(s)$ lie. In this case the triangle $\mathsf{conv}(P_S)$ intersects $y = x^2$, so SCSP(S) is approximation resistant.

A simple calculation shows that the above condition is implied by $(s_{max} - s_{min}) \geqslant \sqrt{2(s_{max} + s_{min})}$ which in turn holds if $(s_{max} - s_{min}) \geqslant 2\sqrt{k}$. This means that SCP($S$) is approximation resistant unless s_{min} and s_{max} are very close. See Figure 1 for an example. We conjecture that this simple condition completely characterizes approximation resistance of symmetric CSPs without negation. Note that we exclude the cases where S contains 0 or k, since without negation, a trivial deterministic strategy to give the same value to every variable satisfies every constraint.

Conjecture 1.1. *For $S \subseteq [k-1]$, SCSP(S) without negation is approximation resistant if and only if* (1) *holds.*

The hardness claim, the "if" part, is currently proved only under the UGC, but our focus is on the algorithmic claim that the violation of (1) leads to an approximation algorithm that outperforms the best random assignment. Even though we were not formally able to prove Conjecture 1.1, we explain the rationale behind the conjecture and we prove it for the following two natural special cases in Section 2:

1. S is an *interval*: S contains every integer from s_{min} to s_{max}.
2. S is *even*: $s \in S$ if and only if $k - s \in S$.

Theorem 1.2. *If $S \subseteq [k-1]$ and S is either an interval or even, SCSP(S) without negation is approximation resistant if and only if* (1) *holds (the hardness claim, i.e., the "if" part, is under the Unique Games conjecture).*

For SCSP(S) with negation, the analogous condition is whether $\mathsf{conv}(P_S)$ contains a single point $(\frac{1}{2},\frac{1}{4})$ or not. While it is tempting to pose a conjecture similar to Conjecture 1.1, we refrain from doing so due to the lack of evidence compared to the case without negation. However, we prove the following theorem which shows that the analogous characterization works at least for the two special cases introduced above.

Theorem 1.3. *If $S \subset [k] \cup \{0\}$ and S is either an interval or even, SCSP(S) with negation is approximation resistant if and only if $\mathsf{conv}(P_S)$ contains $(\frac{1}{2},\frac{1}{4})$ (the hardness claim, i.e., the "if" part, is under the Unique Games conjecture).*

1.3 Techniques

We mainly focus on SCSPs without negation, and briefly sketch why the violation of (1) might lead to an approximation algorithm that outperforms the best random assignment. Let α^* be the probability that the best random assignment uses, and ρ^* be the expected fraction of constraints satisfied by it. Our algorithms follow the following general framework: sample correlated random variables $X_1,\dots,X_n$, where each X_i lies in $[-\alpha^*, 1-\alpha^*]$, and independently round $x_i \leftarrow 1$ with probability $\alpha^* + X_i$.

Fix one constraint $C = (x_1,\dots,x_k)$ (for SCSPs with negation, additionally assume that offsets are all 0). Using symmetry, the probability that it is satisfied by the above strategy can be expressed as

$$\rho^* + \sum_{l=1}^{k} c_l \mathop{\mathbb{E}}_{I \in \binom{[k]}{l}} [\prod_{i \in I} X_i].$$

For some coefficients $\{c_l\}_{l\in[k]}$. These coefficients c_l can be expressed by the following two ways.

- Let $f(\alpha) : [0,1] \mapsto [0,1]$ the probability that a constraint is satisfied by a random assignment with probability α. c_l is proportional to $f^{(l)}(\alpha^*)$, the l'th derivative of f evaluated at α^*.
- Let $Q = \{(x_1,\dots,x_k) \in \{0,1\}^k : (x_1 + \cdots + x_k) \in S\}$ be the predicate associated with S. When $\alpha^* = \frac{1}{2}$, c_l is proportional to the Fourier coefficient $\hat{Q}(T)$ with $|T| = l$.

Given this observation, α^* for SCSPs without negation has nice properties since it should be a global maximum in the interval $[0,1]$. In particular, it should be a local maximum so that $c_1 = f'(\alpha) = 0$ and $c_2, f''(\alpha) \leqslant 0$. By modifying an algorithm by Austrin and Håstad [2], we prove that we can sample $X_1,\dots,X_n$ such that the average second moment $\mathbb{E}[X_i X_j]$ is strictly negative if (1) does not hold. By scaling X_i's so that the product of at least three X_i's becomes negligible, this idea results in an approximation algorithm that outperforms the best random assignment, except the degenerate case where $c_2 = f''(\alpha^*) = 0$ even though α^* is a local maximum. This is the main rationale behind Conjecture 1.1 and we elaborate this belief more in Section 2. It is notable that our conjectured characterization for the case without negation only depends on the minimum and the maximum number in S, while α^* depends on other elements.

For SCSPs with negation where α^* is fixed to be $\frac{1}{2}$, the situation becomes more complicated since c_1 and $f'(\alpha)$ are not necessarily zero and there are many ways that $\mathsf{conv}(P_S)$ does not contain $(\frac{1}{2},\frac{1}{4})$ (in the case of SCSPs without negation, the slope of the line separating $\mathsf{conv}(P_S)$ and $y = x^2$ is always positive, but it is not the case here). Therefore, a complete characterization requires understanding interactions among c_1, c_2, and the separating line. We found that the somewhat involved method of Austrin, Benabbas, and Magen [1] gives a way to sample these $X_1,\dots,X_n$ with desired first and second moments to prove our results when S exhibits additional special structures, but believe that a new set of ideas are required to give a complete characterization.

1.4 Organization

In Section 2, we study SCSPs without negation. We further elaborate our characterization in Section 2.1, and provide an algorithm in Section 2.2. We study SCSPs with negation in Section 3.

2 Symmetric CSPs without negation

2.1 A 2-dimensional characterization

Fix k and $S \subseteq [k-1]$. Our conjectured condition to be approximation resistant is that $\mathsf{conv}(P_S)$ intersects the curve $y = x^2$, which is equivalent to (1). Austrin and Håstad [2] proved that this simple condition is sufficient to be approximation resistant.

Theorem 2.1 ([2])**.** *Let $S \subseteq [k-1]$ be such that* (1) *holds. Then, assuming the Unique Games Conjecture, SCSP(S) without negation is approximation resistant.*

They studied general CSPs and their condition is more complicated than stated here. See Appendix A to see how it is simplified for SCSPs. We conjecture that for SCSPs, this condition is indeed equivalent to approximation resistance.

Conjecture 2.2 (Restatement of Conjecture 1.1)**.** *For $S \subseteq [k-1]$, SCSP(S) without negation is approximation resistant if and only if* (1) *holds.*

To provide our rationale behind the conjecture, we define the function $f : [0,1] \mapsto [0,1]$ to be the probability that one constraint is satisfied by the random assignment that gives $x_i \leftarrow 1$ independently with probability α.

$$f(\alpha) = \sum_{s \in S} \binom{k}{s} \alpha^s (1-\alpha)^{k-s}$$

Let $\alpha^* \in [0,1]$ be a value that maximizes $f(\alpha)$, and $\rho^* := f(\alpha^*)$. There might be more than one α with $f(\alpha) = \rho^*$. In Section 2.2, we prove that S is not approximation resistant if there exists one such α^* with a negative second derivative.

Theorem 2.3. *$S \subseteq [k-1]$ be such that* (1) *does not hold and there exists $\alpha^* \in [0,1]$ such that $f(\alpha^*) = \rho^*$ and $f''(\alpha^*) < 0$. Then, there is a randomized polynomial time algorithm for SCSP(S) that satisfies strictly more than ρ^* fraction of constraints in expectation.*

Since $f(0) = f(1) = 0 < \rho^*$, every $\alpha \in [0,1]$ with $f(\alpha) = \rho^*$ must be a local maximum, so it should have $f'(\alpha) = 0$ and $f''(\alpha) \leqslant 0$. If α is a local maximum, $f''(\alpha) = 0$ also implies $f'''(\alpha) = 0$, so ruling out this degeneracy at a global maximum gives the complete characterization!

Ruling out this degeneracy at a global maximum does not seem to be closely related to general shape of $f(\alpha)$ or S. It might still hold even if $f(\alpha)$ has multiple global maxima, or S satisfies (1) so that SCSP(S) is approximation resistant.

However, examples in Figure 2 led us to believe that the condition (1) is also related to general shape of f. When S contains two numbers l and r with $l + r = k$, as two numbers become far apart, f becomes unimodal to bimodal, and the transition happens exactly when (1) starts to hold. Furthermore, the degenerate case $f'(\alpha^*) = f''(\alpha^*) = 0$ happens when (1) holds with equality. Intuitively, when two numbers l and r are far apart, it is a better strategy to focus on only one of them (i.e. $\alpha^* \approx \frac{l}{k}$ or $\frac{r}{k}$) so f is bimodal, but if (1) does not hold and l and r are close enough, it is better to target in the middle to satisfy both l and r with reasonability probability so that f is unimodal with a large negative curvature at α^*.

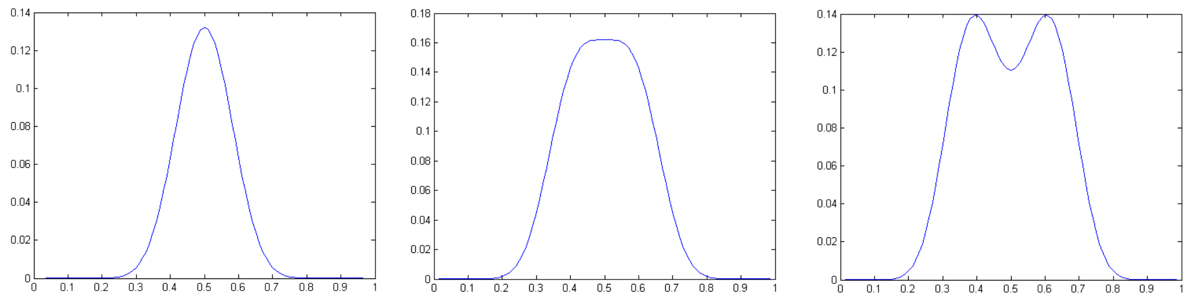

Figure 2: Examples for $k = 36$. **Left**: $S = \{18\}$, (1) is not satisfied, unimodal with $\alpha^* = \frac{1}{2}$, $f''(\frac{1}{2}) < 0$. **Middle**: $S = \{15, 21\}$, (1) is satisfied with equality, unimodal with $\alpha^* = \frac{1}{2}$, but $f''(\frac{1}{2}) = 0$. **Right**: $S = \{14, 22\}$, (1) is satisfied with slack, bimodal with two α^*, but $f''(\alpha^*) < 0$.

Having more points between l and r seems to strengthen the above intuition, and removing the assumption that $l + r = k$ only seems to add algebraic complication without hurting the intuition. Thus, we propose the following stronger conjecture that implies Conjecture 1.1.

Conjecture 2.4. *If* (1) *does not hold, $f(\alpha)$ is unimodal in* $[0,1]$ *with the unique maximum at α^*, and $f''(\alpha^*) < 0$.*

While we are unable to formally prove Conjecture 2.4 for every S, we establish it for the case when S is either an interval (Section 2.3) or even (Section 2.4), thus proving Theorem 1.2.

2.2 Algorithm

Let $\alpha^* \in [0,1]$ be such that $f(\alpha^*) = \rho^*$ and $f''(\alpha^*) < 0$. Furthermore, suppose that S does not satisfy (1). We give a randomized approximation algorithm which is guaranteed to satisfy strictly more than ρ^* fraction of constraints in expectation, proving Theorem 2.3. Let $D := D(k)$ be a large constant determined later. Our strategy is the following.

1. Sample $X_1, \ldots, X_n$ from some correlated multivariate normal distribution where each X_i has mean 0 and variance at most σ^2 for some $\sigma := \sigma(k)$.

2. For each $i \in [n]$, set
$$X'_i = \begin{cases} -D\alpha^* & \text{if } X_i < -D\alpha^* \\ D(1-\alpha^*) & \text{if } X_i > D(1-\alpha^*) \\ X_i & \text{otherwise} \end{cases}$$
so that $\alpha^* + \frac{X'_i}{D}$ is always in $[0,1]$.

3. Set $x_i \leftarrow 1$ independently with probability $\alpha^* + \frac{X'_i}{D}$.

Fix one constraint C and suppose that $C = (x_1, \ldots, x_k)$. We consider a multivariate polynomial
$$g(y_1, \ldots, y_k) := \sum_{T \subseteq [k], |T| \in S} \prod_{i \in T} (\alpha^* + \frac{y_i}{D}) \prod_{i \in [k] \setminus T} (1 - \alpha^* - \frac{y_i}{D}).$$

$g(X'_1, \ldots, X'_k)$ is equal to the probability that the constraint C is satisfied. By symmetry, for any $1 \leqslant i_1 < \cdots < i_l \leqslant k$, the coefficient of a monomial $y_{i_1} y_{i_2} \ldots y_{i_l}$ only depends on l. Let c_l be this coefficient.

Lemma 2.5. $c_l = \frac{(k-l)!}{k!D^l} f^{(l)}(\alpha^*)$.

Proof. Note that $g(y,y,\ldots,y) = f(\alpha^* + \frac{y}{D})$, which has the Taylor expansion

$$\sum_{l=0}^{k} \frac{f^{(l)}(\alpha^*)}{l!} (\frac{y}{D})^l.$$

Since g is multilinear, by symmetry, the coefficient of a monomial $y_{i_1} y_{i_2} \ldots y_{i_l}$ in $g(y_1,\ldots,y_k)$ is equal to the coefficient of y^l in $f(\alpha^* + \frac{y}{D})$ divided by $\binom{k}{l}$, which is $c_l = \frac{(k-l)!}{k!D^l} f^{(l)}(\alpha^*)$. □

We analyze the overall performance of this algorithm. Let $\mathscr{D}_l$ be the distribution on $\binom{[n]}{l}$ where we sample a constraint C uniformly at random, sample l distinct variables from $\binom{C}{l}$, and output their indices. We prove the following lemma, which implies that by taking large D, the effect of truncation from X_i to X_i' and the contribution of monomials of degree greater than two become small.

Lemma 2.6. *The expected fraction of constraints satisfied by the above algorithm is at least*

$$\rho^* + c_2 \binom{k}{2} \mathop{\mathbb{E}}_{(i,j)\sim\mathscr{D}_2} [X_i X_j] - O_k(\frac{1}{D^3}) = \rho^* + \frac{f''(\alpha^*)}{2D^2} \mathop{\mathbb{E}}_{(i,j)\sim\mathscr{D}_2} [X_i X_j] - O_k(\frac{1}{D^3}),$$

where $O_k(\cdot)$ is hiding constants depending on k.

Proof. Note that as long as S does not contain 0 or k, $\alpha^* \in [\frac{1}{k}, 1-\frac{1}{k}]$. For any $1 \leqslant l \leqslant k$ and $1 \leqslant i_1 < \cdots < i_l \leqslant k$, we apply Lemma B.1 (set $D \leftarrow \frac{D}{k}$),

$$|\mathbb{E}[\prod_{j=1}^{l} X_{i_j}] - \mathbb{E}[\prod_{j=1}^{l} X_{i_j}]| \leqslant 2^l \cdot \sigma^l \cdot l! \cdot e^{-D/kl}.$$

If we expand $f(\alpha) = \sum_{l=0}^{k} a_l \alpha^l$, each coefficient a_l has magnitude at most 2^k, which means that $|f^{(l)}(\alpha^*)|$ is bounded by $k2^k k!$. Therefore, any $|c_l|$ is at most $k2^k k!$. Let c_{max} be this quantity. Summing over this error for all monomials, the probability that a fixed constraint $C = \{x_1,\ldots,x_k\}$ is satisfied is

$$\begin{aligned}
\mathbb{E}[g(X_1',\ldots,X_k')] \geqslant\ & \mathbb{E}[g(X_1,\ldots,X_k)] - c_{max} \cdot 2^{2k} \cdot \sigma^k \cdot k! \cdot e^{-D/k^2} \\
=\ & \rho^* + \sum_{l=1}^{k} c_l \sum_{1\leqslant i_1<\cdots<i_l\leqslant k} X_{i_1} X_{i_2} \ldots X_{i_l} - O_k(e^{-D/k^2}) \\
=\ & \rho^* + \sum_{l=1}^{k} c_l \sum_{1\leqslant i_1<\cdots<i_l\leqslant k} X_{i_1} X_{i_2} \ldots X_{i_l} - O_k(e^{-D/k^2})
\end{aligned}$$

Averaging over m constraints, the expected fraction of satisfied constraints is at least

$$\begin{aligned}
& \rho^* + \sum_{l=1}^{k} c_l \binom{k}{l} \mathop{\mathbb{E}}_{(i_1,\ldots,i_l)\sim\mathscr{D}_l} [X_{i_1} \ldots X_{i_l}] - O_k(e^{-D/k^2}) \\
=\ & \rho^* + c_2 \binom{k}{2} \mathop{\mathbb{E}}_{(i_1,i_2)\sim\mathscr{D}_2} [X_{i_1} X_{i_2}] + \sum_{l=3}^{k} c_l \binom{k}{l} \mathop{\mathbb{E}}_{(i_1,\ldots,i_l)\sim\mathscr{D}_l} [X_{i_1} \ldots X_{i_l}] - O_k(e^{-D/k^2}) \\
=\ & \rho^* + c_2 \binom{k}{2} \mathop{\mathbb{E}}_{(i_1,i_2)\sim\mathscr{D}_2} [X_{i_1} X_{i_2}] - O_k(\frac{1}{D^3}),
\end{aligned}$$

where the first equality follows from the fact that $\mathbb{E}[X_i] = 0$ for all i. Recall that $c_l = \frac{(k-l)!}{k!D^l} f^{(l)}(\alpha^*)$ so that $|c_l| = O_k(\frac{1}{D^l})$. □

Therefore, if we have a way to sample $X_1,\dots,X_n$ such that each X_i has mean 0 and variance at most σ^2, and $\mathbb{E}_{(i,j)\sim\mathscr{D}_2}[X_iX_j] < -\delta$ for some $\delta := \delta(k) > 0$, taking D large enough ensures that the algorithm satisfies strictly more than ρ^* fraction of constraints. We now show how to do such a sampling.

We assume that for some $\varepsilon := \varepsilon(k) > 0$, the given instance admits a solution that satisfies $(1-\varepsilon)$ fraction of constraints. Otherwise, the random assignment with probability α^* guarantees the approximation ratio of $\frac{\rho^*}{1-\varepsilon}$. The following lemma completes the proof of Theorem 2.3.

Lemma 2.7. *Suppose that S does not satisfy* (1). *For sufficiently small $\varepsilon,\delta > 0$ and sufficiently large σ all depending only on k, given an instance of SCSP(S) where $(1-\varepsilon)$ fraction of constraints are simultaneously satisfiable, it is possible to sample $X_1,\dots,X_n$ from a multivariate normal distribution such that each X_i has mean 0 and variance bounded by σ^2, and $\mathbb{E}_{(i,j)\sim\mathscr{D}_2}[X_iX_j] < -\delta$.*

Proof. Recall that (1) is equivalent to the fact that the line ℓ passing $P(s_{min})$ and $P(s_{max})$ intersects the curve $y = x^2$. Let a be the value that the vector $(a,-1)$ is orthogonal to ℓ. a is strictly positive since ℓ has a positive slope. If ℓ and $y = x^2$ do not intersect, there is a line with the same slope as ℓ that strictly separates $y = x^2$ and $\{P(s) : s \in S\}$ — in other words, there exists $c \in \mathbb{R}$ such that

- $ax - y + c > \gamma > 0$ for $(x,y) \in \{P(s) : s \in S\}$.
- $ax - x^2 + c < 0$ for any $x \in \mathbb{R} \Rightarrow c < \frac{-a^2}{4}$.

Consider a constraint $C = (x_1,\dots,x_k)$. Since $(\mathbb{E}_{i\in[k]}[x_i], \mathbb{E}_{i\neq j\in[k]}[x_ix_j]) = P(x_1+\cdots+x_k)$, if C is satisfied,

$$a \mathop{\mathbb{E}}_{i\in[k]}[x_i] - \mathop{\mathbb{E}}_{i\neq j\in[k]}[x_ix_j] + c > \gamma.$$

Let

$$\eta := -\min_{x_1,\dots,x_k\in\{0,1\}} \left(a \mathop{\mathbb{E}}_{i\in[k]}[x_i] - \mathop{\mathbb{E}}_{i\neq j\in[k]}[x_ix_j] + c \right).$$

We solve the following semidefinite programm (SDP):

$$\begin{aligned}
\text{maximize} \quad & a \mathop{\mathbb{E}}_{i\in\mathscr{D}_1}[\langle v_0, v_i\rangle] - \mathop{\mathbb{E}}_{i,j\in\mathscr{D}_2}[\langle v_i, v_j\rangle] + c \\
\text{subject to} \quad & ||v_0|| = 1 \\
& \langle v_i, v_0\rangle = ||v_i||^2 \qquad \text{for all } i \in [n]
\end{aligned}$$

Note that $\langle v_i, v_0\rangle = ||v_i||^2$ implies $||v_i|| \leqslant 1$. For any assignment to $x_1,\dots,x_n$, setting $v_i = x_iv_0$ satisfies that $x_i = \langle v_0, v_i\rangle$ and $x_ix_j = \langle v_i, v_j\rangle$. Since at least $(1-\varepsilon)$ fraction of constraints can be simultaneously satisfied, the optimum of the above SDP is at least $(1-\varepsilon)\gamma - \varepsilon\eta$. Given $\gamma > 0$ and η, take sufficiently small $\varepsilon,\delta > 0$ such that $(1-\varepsilon)\gamma - \varepsilon\eta = \delta$. There are finitely many S (thus γ and η) for each k, so ε and δ can be taken to depend only on k. Given vectors $v_0, v_1, \dots, v_n$, we sample $X_1,\dots,X_n$ by the following simple procedure:

1. Sample a vector g whose coordinates are independent standard normal.
2. Let $X_i = \langle g, v_i - \frac{a}{2}v_0\rangle$.

It is clear that $\mathbb{E}[X_i] = 0$ for each i, and $\mathbb{E}[X_i^2] = ||v_i - \frac{a}{2}v_0||^2 \leqslant (a+1)^2 + 1$, so taking $\sigma := \sigma(k)$ large enough ensures that the variance of each X_i is bounded by σ^2. We now compute the second moment.

$$\begin{aligned}
& \mathop{\mathbb{E}}_{i,j\sim\mathscr{D}_2}[X_i X_j] \\
= \quad & \mathop{\mathbb{E}}_{i,j\sim\mathscr{D}_2}[\langle v_i - \frac{a}{2}v_0, v_j - \frac{a}{2}v_0\rangle] \\
= \quad & \mathop{\mathbb{E}}_{i,j\sim\mathscr{D}_2}[\langle v_i, v_j\rangle] - a \mathop{\mathbb{E}}_{i\in\mathscr{D}_1}[\langle v_i, v_0\rangle] + \frac{a^2}{4} \\
< \quad & \mathop{\mathbb{E}}_{i,j\sim\mathscr{D}_2}[\langle v_i, v_j\rangle] - a \mathop{\mathbb{E}}_{i\in\mathscr{D}_1}[\langle v_i, v_0\rangle] - c \\
\leqslant \quad & -((1-\varepsilon)\gamma - \varepsilon\eta) = -\delta,
\end{aligned}$$

where the first inequality follows from $c < -\frac{a^2}{4}$ and the second follows from the optimality of our SDP. □

2.3 Case of Interval S

We study properties of $f(\alpha)$ when S is an interval — $S = \{s_{min}, s_{min}+1, \ldots, s_{max}-1, s_{max}\}$, and prove Conjecture 2.4 for this case. One notable fact is that as long as S is interval, the conclusion of Conjecture 2.4 is true even if S does satisfy (1) and becomes approximation resistant.

Lemma 2.8. *Suppose $S \subseteq [k-1]$ is an interval. Then, $f(\alpha)$ is unimodal in $[0,1]$ with the unique maximum at α^* and $f''(\alpha^*) < 0$.*

Proof. Let $l := s_{min}$ and $r = s_{max}$. Given

$$f(\alpha) = \sum_{s=l}^{r} \binom{k}{s} \alpha^s (1-\alpha)^{k-s}$$

and

$$f'(\alpha) = \sum_{s=l}^{r} \binom{k}{s} \left(s\alpha^{s-1}(1-\alpha)^{k-s} - (k-s)\alpha^s(1-\alpha)^{k-s-1} \right),$$

since $\binom{k}{s}(k-s) = \binom{k}{s+1}(s+1)$, we have

$$f'(\alpha) = \binom{k}{l} l\alpha^{l-1}(1-\alpha)^{k-l} - \binom{k}{r}(k-r)\alpha^r(1-\alpha)^{k-r-1}.$$

If $0 < \alpha < 1$, setting $\beta := \frac{\alpha}{1-\alpha}$ gives a unique non-zero solution to $f'(\beta) = 0$. This proves the unimodality. For the second derivative,

$$\begin{aligned}
f''(\alpha) = & \binom{k}{l} l(l-1)\alpha^{l-2}(1-\alpha)^{k-l} - \binom{k}{l} l(k-l)\alpha^{l-1}(1-\alpha)^{k-l-1} + \\
& \binom{k}{r}(k-r)(k-r-1)\alpha^r(1-\alpha)^{k-r-2} - \binom{k}{r} r(k-r)\alpha^{r-1}(1-\alpha)^{k-r-1} \\
= & \binom{k}{l} l\alpha^{l-2}(1-\alpha)^{k-l-1}\Big((l-1)(1-\alpha) - (k-l)\alpha\Big) + \\
& \binom{k}{r}(k-r)\alpha^{r-1}(1-\alpha)^{k-r-2}\Big((k-r-1)\alpha - r(1-\alpha)\Big).
\end{aligned}$$

Since $\frac{l-1}{k-1} < \frac{l}{k} \leqslant \alpha^* \leqslant \frac{r}{k} < \frac{r}{k-1}$,

$$(l-1)(1-\alpha^*) - (k-l)\alpha^* = (l-1) - (k-1)\alpha^* < 0$$

and

$$(k-r-1)\alpha^* - r(1-\alpha^*) = (k-1)\alpha^* - r < 0,$$

so that $f''(\alpha^*) < 0$. □

2.4 Case of Even S

We study properties of $f(\alpha)$ when S is even — $s \in S$ if and only if $k-s \in S$, and prove Conjecture 2.4 for this case. We first simplify (1) for this setting. If we let $l := s_{min}$ and $r := s_{max} = k-l$, (1) is equivalent to

$$\frac{(l+r-1)^2}{k-1} \geqslant \frac{4lr}{k} \quad \Leftrightarrow \quad k(k-1) \geqslant 4lr \quad \Leftrightarrow \quad (r-l)^2 \geqslant k.$$

Therefore, (1) is equivalent to

$$r - l \geqslant \sqrt{k} \tag{2}$$

Lemma 2.9. *Suppose $S \subseteq [k-1]$ is even. If (2) does not hold, $f(\alpha)$ is unimodal in $[0,1]$ with the unique maximum at $\alpha^* = \frac{1}{2}$ and $f''(\alpha^*) < 0$.*

Proof. Given a even S, let $S_1 = \{s \in S : s \leqslant k/2\}$. When we write f_S to denote the dependence of f on S, we can decompose $f_S(\alpha) = \sum_{s \in S_1} f_{\{s,k-s\}}(\alpha)$, so the following claim proves the lemma. □

Claim 2.10. *Let $l \leqslant \frac{k}{2}$ and $r = k-l$ such that $r - l < \sqrt{k} \Leftrightarrow k(k-1) < 4lr$. Let $S = \{l, r\}$. f is unimodal with the unique maximum at $\frac{1}{2}$, and $f''(\frac{1}{2}) < 0$.*

Proof. Note that f is symmetric around $\alpha = 1/2$. If there exists a local maximum at $\alpha' \in (0, 1/2)$, f also has a local maximum at $(1-\alpha')$ with the same value, so there must exist a local minimum in $(\alpha', 1-\alpha')$. In particular, there is $\alpha \in (\alpha', 1-\alpha')$ such that $f'(\alpha) = 0$ and $f''(\alpha) \geqslant 0$. We prove that such α cannot exist.

$$\begin{aligned}
& f'(\alpha) = 0 \\
\Leftrightarrow \quad & \binom{k}{l}\alpha^{l-1}(1-\alpha)^{r-1}(l-k\alpha) + \binom{k}{r}\alpha^{r-1}(1-\alpha)^{l-1}(r-k\alpha) = 0 \\
\Leftrightarrow \quad & \frac{\binom{k}{l}\alpha^{l-1}(1-\alpha)^{r-1}}{\binom{k}{r}\alpha^{r-1}(1-\alpha)^{l-1}} = \frac{\binom{k}{l}(1-\alpha)^{r-l}}{\binom{k}{r}\alpha^{r-l}} = -\frac{(k\alpha-r)}{(k\alpha-l)}
\end{aligned}$$

Similarly,

$$\begin{aligned}
& f''(\alpha) \geqslant 0 \\
\Leftrightarrow \quad & \frac{\binom{k}{l}(1-\alpha)^{r-l}}{\binom{k}{r}\alpha^{r-l}} \geqslant -\frac{r(r-1)(1-\alpha)^2 - 2rl\alpha(1-\alpha) + l(l-1)\alpha^2}{l(l-1)(1-\alpha)^2 - 2rl\alpha(1-\alpha) + r(r-1)\alpha^2}.
\end{aligned}$$

By symmetry, we can assume $\alpha \geqslant \frac{1}{2}$, so that $(k\alpha - l) \geqslant 0$ and $l(l-1)(1-\alpha)^2 - 2rl\alpha(1-\alpha) + r(r-1)\alpha^2 \geqslant 0$.

$$
\begin{aligned}
& \frac{(k\alpha - r)}{(k\alpha - l)} \leqslant \frac{r(r-1)(1-\alpha)^2 - 2rl\alpha(1-\alpha) + l(l-1)\alpha^2}{l(l-1)(1-\alpha)^2 - 2rl\alpha(1-\alpha) + r(r-1)\alpha^2} \\
\Leftrightarrow \quad & (k\alpha - r)(l(l-1)(1-\alpha)^2 - 2rl\alpha(1-\alpha) + r(r-1)\alpha^2) \\
& \leqslant (k\alpha - l)(r(r-1)(1-\alpha)^2 - 2rl\alpha(1-\alpha) + l(l-1)\alpha^2) \\
\Leftrightarrow \quad & \alpha^2(l^3 - r^3 - (l^2 - r^2) + rl(l-r) - 2k(l^2 - r^2) + 2k(l-r)) + \alpha(k(l^2 - r^2) - k(l-r)) - rl(l-r) \leqslant 0 \\
\Leftrightarrow \quad & \alpha^2(-k^2 + k) + \alpha(k^2 - k) - rl \geqslant 0 \qquad \text{divide by } (l-r) \text{ and use } l + r = k
\end{aligned}
$$

However, $\alpha^2(-k^2+k) + \alpha(k^2-k) - rl$ has a negative leading coefficient and its discriminant is

$$(k^2-k)^2 - 4rl(k^2-k) = (k^2-k)(k^2-k-4rl) < 0$$

by the assumption of the claim. $\square$

We do not formally prove the converse, but Figure 2 shows examples where it is tight. When (2) holds with equality, f still has the unique local maximum at $\frac{1}{2}$ but $f''(\frac{1}{2}) = 0$, and even when (2) holds with small slack, two local maxima start to appear. This phenomenon is one of the main reasons that we pose Conjecture 2.4. Though we were not able to formally prove for the general case, we believe that the violation of (1) not only allows us to sample random variables with desired second moments but also ensures that $f(\alpha)$ is a nice unimodal curve.

3 Approximability of symmetric CSPs with negation

Fix k and $S \subset [k] \cup \{0\}$. In this section, we consider SCSP(S) with negation and prove Theorem 1.2. Note that in this section we allow S to contain 0 or k. For example, famous Max-3SAT is 3-SCSP($\{1,2,3\}$). We still exclude the trivial case $S = [k] \cup \{0\}$.

The condition we are interested in is whether $\mathsf{conv}(P_S)$ contains $(\frac{1}{2}, \frac{1}{4})$. In SCSPs with negation, the sufficient condition of Austrin and Mossel on general CSPs to be approximation resistant becomes equivalent to it. See Appendix A to see the equivalence.

Theorem 3.1 ([2]). *Fix k and let $S \subset [k] \cup \{0\}$ be such that $\mathsf{conv}(P_S)$ contains $(\frac{1}{2}, \frac{1}{4})$. Then, assuming the Unique Games Conjecture, SCSP(S) with negation is approximation resistant.*

On the other hand, we now show that the algorithm of Austrin et al. [1], which is inspired by Hast [10], can be used to show that if S is an interval or even and $\mathsf{conv}(P_S)$ does not contain $(\frac{1}{2}, \frac{1}{4})$, SCSP($S$) is not approximation resistant.

Let $f : \{0,1\}^k \mapsto \{0,1\}$ be the function such that $f(x_1, \ldots, x_k) = 1$ if and only if $(x_1 + \cdots + x_k) \in S$. Define the inner product of two functions as $\langle f, g \rangle = \mathbb{E}_{x \in \{0,1\}^k}[f(x)g(x)]$, and for $T \subseteq [k]$, let $\chi_T(x_1, \ldots, x_k) = \prod_{i \in T}(-1)^{x_i}$. It is well known that $\{\chi_T\}_{T \subseteq [k]}$ form an orthonormal basis and every function has a unique *Fourier expansion* with respect to this basis,

$$f = \sum_{T \subseteq [k]} \hat{f}(T)\chi_T, \qquad \hat{f}(T) := \langle f, \chi_T \rangle.$$

Define

$$f^{=d}(x) = \sum_{|T|=d} \hat{f}(S)\chi_T(x).$$

The main theorem of Austrin et al. [1] is

Theorem 3.2 ([1])**.** *Suppose that there exists $\eta \in \mathbb{R}$ such that*

$$\frac{2\eta}{\sqrt{2\pi}} f^{=1}(x) + \frac{2}{\pi} f^{=2}(x) > 0 \tag{3}$$

for every $x \in f^{-1}(1)$. Then there is a randomized polynomial time algorithm that approximates SCSP(S) better than the random assignment in expectation.

We compute $f^{=1}$ and $f^{=2}$.

$$\hat{f}(\{1\}) = \langle f, \chi_{\{1\}} \rangle = \frac{1}{2^k} \sum_{s \in S} \left(\binom{k-1}{s} - \binom{k-1}{s-1} \right)$$
$$\hat{f}(\{1,2\}) = \langle f, \chi_{\{1,2\}} \rangle = \frac{1}{2^k} \sum_{s \in S} \left(\binom{k-2}{s} - 2\binom{k-2}{s-1} + \binom{k-2}{s-2} \right)$$

By symmetry, $\hat{f}_T =: \hat{f}_1$ is the same for all $|T| = 1$ and $\hat{f}_T =: \hat{f}_2$ is the same for all $|T| = 2$. If we let $s = x_1 + \cdots + x_k$,

$$f^{=1}(x) = \hat{f}_1 \sum_{i \in [k]} (-1)^{x_i} = k\hat{f}_1 \mathop{\mathbb{E}}_{i \in [k]}[-2x_i + 1] = k\hat{f}_1(-2\frac{s}{k} + 1)$$
$$f^{=2}(x) = \hat{f}_2 \sum_{i \neq j} (-1)^{x_i + x_j} = \binom{k}{2} \hat{f}_2 \mathop{\mathbb{E}}_{i \neq j}[(-2x_i + 1)(-2x_j + 1)] = \binom{k}{2} \hat{f}_2 (4\frac{s(s-1)}{k(k-1)} - 4\frac{s}{k} + 1).$$

When S is an interval. Let $S = \{l, l+1, \ldots, r-1, r\}$. If $r \leqslant \frac{k}{2}$, we have $(\frac{-2s}{k} + 1) \leqslant 0$ for all $s \in S$, so choosing η either large enough or small enough ensures (3). Similarly, if $l \geqslant \frac{k}{2}$, (3) holds. Therefore, we assume that $l < \frac{k}{2}$ and $r > \frac{k}{2}$, and compute $\hat{f}_2$.

$$\begin{aligned} \hat{f}_2 &= \frac{1}{2^k} \sum_{s=l}^{r} \left(\binom{k-2}{s} - 2\binom{k-2}{s-1} + \binom{k-2}{s-2} \right) \\ &= \frac{1}{2^k} \left(\binom{k-2}{l-2} - \binom{k-2}{l-1} + \binom{k-2}{r} - \binom{k-2}{r-1} \right) \end{aligned}$$

Since $\binom{k-2}{l-1} > \binom{k-2}{l-2}$ for $0 < l < \frac{k}{2}$ and $\binom{k-2}{r-1} > \binom{k-2}{r}$ for $\frac{k}{2} < r < k$, $\hat{f}_2 < 0$ except when $l = 0$ and $r = k$ (i.e., $S = [k] \cup \{0\}$).

If $\mathsf{conv}(P_S)$ does not contain $(\frac{1}{2}, \frac{1}{4})$, there exist $\alpha, \beta \in \mathbb{R}$ such that for any $(a, b) \in \mathsf{conv}(P_S)$,

$$\alpha(a - \frac{1}{2}) + \beta(b - \frac{1}{4}) > 0.$$

If k is even, $s := \frac{k}{2} \in S$ and $P(s) = (\frac{1}{2}, \frac{s-1}{2(k-1)})$ where $\frac{s-1}{2(k-1)} < \frac{1}{4}$, which implies $\beta < 0$ since the above inequality should hold for all $s \in S$. When k is odd (let $k = 2s+1$), s and $s+1$ should be in S and

$$\frac{1}{2}\Big(P(s) + P(s+1)\Big) = (\frac{1}{2}, \frac{s^2}{k(k-1)}),$$

where $\frac{s^2}{k(k-1)} < \frac{1}{4}$. Therefore, we can conclude $\beta < 0$ in any case. For any $x \in f^{-1}(1)$ with $s = x_1 + \cdots + x_k$ and $P(s) = (a,b)$,

$$\begin{aligned}
&\frac{2\eta}{\sqrt{2\pi}} f^{=1}(x) + \frac{2}{\pi} f^{=2}(x) \\
= \quad &\frac{2\eta}{\sqrt{2\pi}} k\hat{f}_1(-2a+1) + \frac{2}{\pi}\binom{k}{2}\hat{f}_2(4b-4a+1) \\
= \quad &\frac{8}{\beta\pi}\binom{k}{2}\hat{f}_2\left(\frac{\frac{2\eta}{\sqrt{2\pi}}k\hat{f}_1}{\frac{8}{\beta\pi}\binom{k}{2}\hat{f}_2}(-2a+1) + \beta(b-a+\frac{1}{4})\right) \\
= \quad &\frac{8}{\beta\pi}\binom{k}{2}\hat{f}_2\left((-\frac{\alpha+\beta}{2})(-2a+1) + \beta(b-a+\frac{1}{4})\right) \quad \text{by adjusting } \eta \text{ so that } \frac{\frac{2\eta}{\sqrt{2\pi}}k\hat{f}_1}{\frac{8}{\beta\pi}\binom{k}{2}\hat{f}_2} = -\frac{\alpha+\beta}{2} \\
= \quad &\frac{8}{\beta\pi}\binom{k}{2}\hat{f}_2\left(\alpha(a-\frac{1}{2}) + \beta(b-\frac{1}{4})\right) \\
&> 0.
\end{aligned}$$

Therefore, (3) is satisfied if S is an interval and $\mathsf{conv}(S)$ does not contain $(\frac{1}{2},\frac{1}{4})$.

When S is even. Given S, let $Q \in \{0,1\}^k$ be the predicate associated with S and $f : \{0,1\}^k \mapsto \{0,1\}$ be the indicator function of Q. We want to show that when S is even,

$$\frac{2\eta}{\sqrt{2\pi}} f^{=1}(x) + \frac{2}{\pi} f^{=2}(x) > 0$$

is satisfied for any $x \in f^{-1}(1)$. When S is even,

$$\hat{f}_1 = \frac{1}{2^{k+1}} \sum_{s\in S} \left(\binom{k-1}{s} - \binom{k-1}{s-1} + \binom{k-1}{k-s} - \binom{k-1}{k-s-1} \right) = 0.$$

We compute the sign of the contribution of each s to $\hat{f}_2$.

$$\begin{aligned}
&\binom{k-2}{s} - 2\binom{k-2}{s-1} + \binom{k-2}{s-2} \geqslant 0 \\
\Leftrightarrow \quad &(k-s)(k-s-1) - 2s(k-s) + s(s-1) \geqslant 0 \\
\Leftrightarrow \quad &4s^2 - 4sk + k^2 - k \geqslant 0 \\
\Leftrightarrow \quad &s \leqslant \frac{k-\sqrt{k}}{2} \text{ or } s \geqslant \frac{k+\sqrt{k}}{2}
\end{aligned}$$

We also consider the line passing $P(s)$ and $P(k-s)$. If we denote $t = k-s$, Its slope is

$$\frac{\frac{t(t-1)-s(s-1)}{k(k-1)}}{\frac{t-s}{k}} = \frac{t^2-s^2-(t-s)}{(k-1)(t-s)} = 1,$$

and the value of this line at $\frac{1}{2}$ is at least $\frac{1}{4}$ when

$$\begin{aligned}
&\frac{s(s-1)+(k-s)(k-s-1)}{2k(k-1)} \geqslant \frac{1}{4} \\
\Leftrightarrow \quad &2s(s-1) + 2(k-s)(k-s-1) \geqslant k(k-1) \\
\Leftrightarrow \quad &s \leqslant \frac{k-\sqrt{k}}{2} \text{ or } s \geqslant \frac{k+\sqrt{k}}{2}.
\end{aligned}$$

Intuitively, if we consider the line of slope 1 that passes $(\frac{1}{2}, \frac{1}{4})$, $P(s)$ is below this line if $s \in (\frac{k-\sqrt{k}}{2}, \frac{k+\sqrt{k}}{2})$. Let $S_1 = S \cap \{0, 1, \ldots, \lceil \frac{k}{2} \rceil\}$. If S_1 contains a value $s_1 \leqslant \frac{k-\sqrt{k}}{2}$ and a value $s_2 \geqslant \frac{k-\sqrt{k}}{2}$ (including the case $s_1 = s_2 = \frac{k-\sqrt{k}}{2}$ is an integer in S_1), the line passing $P(s_1)$ and $P(k-s_1)$ passes a point $(\frac{1}{2}, t_1)$ for some $t_1 \geqslant \frac{1}{4}$ and the line passing $P(s_2)$ and $P(k-s_2)$ passes a point $(\frac{1}{2}, t_2)$ for some $t_2 \leqslant \frac{1}{4}$. Therefore, $\mathsf{conv}(P_S)$ contains a point $(\frac{1}{2}, \frac{1}{4})$ and S becomes balanced pairwise independent. We consider the remaining two cases.

1. $s < \frac{k-\sqrt{k}}{2}$ for all $s \in S_1$: $\hat{f}_2 > 0$ and for all $s \in S$, $-(\frac{s}{k} - \frac{1}{2}) + (\frac{s(s-1)}{k(k-1)} - \frac{1}{4}) > 0$. Therefore, for any $x \in f^{-1}$ with $s = x_1 + \cdots + x_k$,

$$\begin{aligned}
& \frac{2\eta}{\sqrt{2\pi}} f^{=1}(x) + \frac{2}{\pi} f^{=2}(x) \\
= \quad & \frac{2}{\pi} f^{=2}(x) \\
= \quad & \frac{2}{\pi} \binom{k}{2} \hat{f}_2 (4 \frac{s(s-1)}{k(k-1)} - 4 \frac{s}{k} + 1) \\
> \quad & 0.
\end{aligned}$$

2. $s > \frac{k-\sqrt{k}}{2}$ for all $s \in S_1$: $\hat{f}_2 < 0$ and for all $s \in S$, $-(\frac{s}{k} - \frac{1}{2}) + (\frac{s(s-1)}{k(k-1)} - \frac{1}{4}) < 0$. Similarly as above, for any $x \in f^{-1}$ with $s = x_1 + \cdots + x_k$, (3) is satisfied.

References

[1] P. Austrin, S. Benabbas, and A. Magen. On quadratic threshold CSPs. *Discrete Mathematics & Theoretical Computer Science*, 14(2):205–228, 2012. 2, 4, 11, 12

[2] P. Austrin and J. Håstad. On the usefulness of predicates. *ACM Transactions on Computation Theory*, 5(1):1:1–1:24, May 2013. 1, 2, 4, 5, 11, 15

[3] P. Austrin and S. Khot. A characterization of approximation resistance for even k-partite CSPs. In *Proceedings of the 4th Conference on Innovations in Theoretical Computer Science*, ITCS '13, pages 187–196, 2013. 2

[4] P. Austrin and E. Mossel. Approximation resistant predicates from pairwise independence. *Computational Complexity*, 18(2):249–271, 2009. 1, 2

[5] B. Barak, S. O. Chan, and P. Kothari. Sum of squares lower bounds from pairwise independence. In *Proceedings of the 47th Annual ACM Symposium on Theory of Computing*, STOC '15, 2015. To appear. 2

[6] S. Benabbas, K. Georgiou, A. Magen, and M. Tulsiani. SDP gaps from pairwise independence. *Theory of Computing*, 8(1):269–289, 2012. 2

[7] S. O. Chan. Approximation resistance from pairwise independent subgroups. In *Proceedings of the 45th annual ACM Symposium on Theory of Computing*, STOC '13, pages 447–456, 2013. 2

[8] M. Cheraghchi, J. Håstad, M. Isaksson, and O. Svensson. Approximating linear threshold predicates. *ACM Trans. Comput. Theory*, 4(1):2:1–2:31, Mar. 2012. 2

[9] W.-S. Cheung. Generalizations of Hölders inequality. *International Journal of Mathematics and Mathematical Sciences*, 26(1):7–10, 2001. 17

[10] G. Hast. *Beating a random assignment. KTH, Stockholm.* PhD thesis, Ph. D Thesis, 2005. 2, 11

[11] J. Håstad. Some optimal inapproximability results. *Journal of the ACM*, 48(4):798–859, July 2001. 1, 2

[12] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the 34th annual ACM Symposium on Theory of Computing*, STOC '02, pages 767–775, 2002. 2

[13] S. Khot, M. Tulsiani, and P. Worah. A characterization of strong approximation resistance. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing*, STOC '14, pages 634–643, 2014. 2

[14] T. Schaefer. The complexity of satisfiability problems. In *Proceedings of the 10th annual ACM Symposium on Theory of Computing*, STOC '78, pages 216–226, 1978. 1

[15] M. Tulsiani. CSP gaps and reductions in the Lasserre hierarchy. In *Proceedings of the 41st annual ACM Symposium on Theory of Computing*, STOC '09, pages 303–312, 2009. 2

[16] U. Zwick. Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint. In *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '98, pages 201–210, 1998. 2

A Austrin-Håstad Condition for Symmetric CSPs

This section explains how the condition of Austrin-Håstad [2] is simplified for SCSPs. They studied general CSPs where a predicate Q is a subset of $\{0,1\}^k$. Note that given $S \subseteq [k] \cup \{0\}$, SCSP($S$) is equivalent to CSP($Q$) where

$$Q = \{(x_1, \dots, x_k) \in \{0,1\}^k : (x_1 + \cdots + x_k) \in S\} \tag{4}$$

Given Q, their general definition of *pairwise independence* and *positive correlation* is given below.

Definition A.1. *Q is* balanced pairwise independent *if there is a distribution μ supported on Q such that* $\Pr_\mu[x_i = 1] = \frac{1}{2}$ *for every $i \in [k]$ and* $\Pr_\mu[x_i = x_j = 1] = \frac{1}{4}$ *for every $1 \leqslant i < j \leqslant k$.*

Definition A.2. *Q is* positively correlated *if there is a distribution μ supported on Q and $p, \rho \in [0,1]$ with $\rho \geqslant p^2$ such that* $\Pr_\mu[x_i = 1] = p$ *for every $i \in [k]$ and* $\Pr_\mu[x_i = x_j = 1] = \rho$ *for every $1 \leqslant i < j \leqslant k$.*

We formally prove that their definitions have simpler descriptions in $\mathbb{R}^2$ for symmetric CSPs. Recall that given $s \in [k] \cup \{0\}$,

$$P(s) = (\frac{s}{k}, \frac{s(s-1)}{k(k-1)}) \in \mathbb{R}^2 \qquad \text{and} \quad P_S := \{P(s) : s \in S\} .$$

Lemma A.3. *Let $S \subseteq [k] \cup \{0\}$ and Q be obtained by* (4)*. Q is pairwise independent if and only if $\mathsf{conv}(P_S)$ contains $(\frac{1}{2}, \frac{1}{4})$, and Q is positively correlated if and only if $\mathsf{conv}(P_S)$ intersects the curve $y = x^2$.*

Proof. We first prove the second claim of the lemma. Let Q be positively correlated with parameters p, ρ ($\rho \geqslant p^2$) and the distribution μ such that $\Pr_\mu[x_i = 1] = p$ for all i, $\Pr_\mu[x_i = x_j = 1] = \rho$ and for all $i < j$. Let ν be the distribution of $x_1 + \cdots + x_k$ where $(x_1, \ldots, x_k)$ are sampled from μ.

$$(p, \rho) = (\mathop{\mathbb{E}}_i[x_i], \mathop{\mathbb{E}}_{i<j}[x_i x_j]) = (\mathop{\mathbb{E}}_{s\sim\nu}[\frac{s}{k}], \mathop{\mathbb{E}}_{s\sim\nu}[\frac{s(s-1)}{k(k-1)}]) = \mathop{\mathbb{E}}_{s\sim\nu}[P(s)],$$

proving that positive correlation of Q implies $(p, \rho) \in \mathsf{conv}(P_S)$. Since $P(s)$ is strictly below the curve $y = x^2$ for any $s \in [k-1]$ and (p, ρ) is on or above this curve, $\mathsf{conv}(P_S)$ must intersect $y = x^2$.

Suppose that $\mathsf{conv}(P_S)$ intersects the curve $y = x^2$. There exists a distribution ν on S such that $\mathbb{E}_{s\sim\nu}[P(s)] = (p, p^2)$. Let μ_s be the distribution on $\{0,1\}^k$ that uniformly samples a string with exactly s 1's. Let μ be the distribution where s is sampled from ν and $(x_1, \ldots, x_k)$ is sampled from μ_s. By definition, $\Pr_\mu[x_i = 1]$ and $\Pr_\mu[x_i = x_j = 1]$ do not depend on choice of indices,

$$\Pr_\mu[x_1 = 1] = \mathop{\mathbb{E}}_\mu[x_1] = \mathop{\mathbb{E}}_{s\sim\nu}\mathop{\mathbb{E}}_{x\sim\mu_s}[x_1] = \mathop{\mathbb{E}}_{s\sim\nu}[\frac{s}{k}] = p$$

$$\Pr_\mu[x_1 = x_2 = 1] = \mathop{\mathbb{E}}_\mu[x_1 x_2] = \mathop{\mathbb{E}}_{s\sim\nu}\mathop{\mathbb{E}}_{x\sim\mu_s}[x_1 x_2] = \mathop{\mathbb{E}}_{s\sim\nu}[\frac{s(s-1)}{k(k-1)}] = p^2,$$

implying that $(p, p^2) \in \mathsf{conv}(P_S)$.

The proof of the first claim is similar except that the curve $y = x^2$ is replaced by $(\frac{1}{2}, \frac{1}{4})$. $\square$

Lemma A.4. $\mathsf{conv}(P_S)$ *intersects the curve* $x = y^2$ *if and only if*

$$\frac{(s_{max} + s_{min} - 1)^2}{k-1} \geqslant \frac{4 s_{max} s_{min}}{k} .$$

Proof. Let $l = s_{min}$ and $r = s_{max}$. The line passing $P(l)$ and $P(r)$ has a slope $\frac{\frac{r(r-1)-l(l-1)}{k(k-1)}}{\frac{r-l}{k}} = \frac{r+l-1}{k-1}$ and a y-intercept b such that

$$\frac{l(l-1)}{k(k-1)} = \frac{r+l-1}{k-1} \cdot \frac{l}{k} + b \Leftrightarrow b = \frac{l(l-1) - l(r+l-1)}{k(k-1)} = \frac{-lr}{k(k-1)}.$$

This line intersects $y = x^2$ if and only if

$$x^2 = \frac{r+l-1}{k-1} x - \frac{lr}{k(k-1)}$$

has a real root, which is equivalent to

$$(\frac{r+l-1}{k-1})^2 - \frac{4lr}{k(k-1)} \geqslant 0 \Leftrightarrow \frac{(r+l-1)^2}{k-1} \geqslant \frac{4lr}{k}. \qquad \square$$

B Technical Proof

Lemma B.1. *Let* $Y_1, \ldots, Y_l$ *be sampled from a multivariate normal distribution where each* Y_i *has mean 0 and variance at most* σ^2. *Let* $Y'_1, \ldots, Y'_l$ *be such that*

$$Y'_i = \begin{cases} Y_i & \text{if } |Y_i| \leqslant D \\ D & \text{if } Y_i > D \\ -D & \text{if } Y_i < -D \end{cases}$$

Then, for large enough D,

$$|\mathbb{E}[\prod_{i=1}^{l} Y_i] - \mathbb{E}[\prod_{i=1}^{l} Y_i']| \leqslant 2^l \cdot \sigma^l \cdot l! \cdot e^{-D/l}.$$

Proof. For each $i \in [l]$, let $Y_i'' = Y_i' - Y_i$. Take D large enough so that

$$\mathbb{E}[|Y_i''|^l] = 2\int_{y=D}^{\infty} (y-D)^l \phi(y) \leqslant 2\int_{y=D}^{\infty} y^l \phi(y) \leqslant e^{-D}.$$

Also each Y_i, a normal random variable with mean 0 and variance σ, satisfies $\mathbb{E}[|Y_i|^l] \leqslant \sigma^l \cdot l!$. We have

$$\begin{aligned}
|\mathbb{E}[\prod_{i=1}^{l} Y_i] - \mathbb{E}[\prod_{i=1}^{l} Y_i']| &= \quad |\sum_{T\subseteq[l], T\neq[l]} \mathbb{E}[\prod_{i\in T} Y_i \prod_{i\notin T} Y_i'']| \\
&\leqslant \quad \sum_{T\subseteq[l], T\neq[l]} \prod_{i\in T} (\mathbb{E}[|Y_i|^l])^{1/l} \prod_{i\notin T} (\mathbb{E}[|Y_i''|^l])^{1/l} \quad \text{By Generalized Hölder's inequality [9]} \\
&\leqslant \quad 2^l \cdot \sigma^l \cdot l! \cdot e^{-D/l}\ .
\end{aligned}$$

□

ECCC ISSN 1433-8092

http://eccc.hpi-web.de