Every Permutation CSP of arity 3 is Approximation Resistant

Moses Charikar[*]
Dept. of Computer Science
Princeton University
Princeton, NJ 08540.
moses@cs.princeton.edu

Venkatesan Guruswami[†]
Dept. of Computer Science & Engg.
University of Washington
Seattle, WA 98195.
venkat@cs.washington.edu

Rajsekar Manokaran[‡]
Dept. of Computer Science
Princeton University
Princeton, NJ 08540.
rajsekar@cs.princeton.edu

Abstract—**A permutation constraint satisfaction problem (permCSP) of arity k is specified by a subset $\Lambda \subseteq S_k$ of permutations on $\{1, 2, \ldots, k\}$. An instance of such a permCSP consists of a set of variables V and a collection of constraints each of which is an ordered k-tuple of V. The objective is to find a global ordering σ of the variables that maximizes the number of constraint tuples whose ordering (under σ) follows a permutation in Λ. This is just the natural extension of constraint satisfaction problems over finite domains (such as Boolean CSPs) to the world of ordering problems.**

The simplest permCSP corresponds to the case when Λ consists of the identity permutation on two variables. This is just the Maximum Acyclic Subgraph (MAS) problem. It was recently shown that the MAS problem is Unique-Games hard to approximate within a factor better than the trivial $1/2$ achieved by a random ordering [6]. Building on this work, in this paper we show that for *every* permCSP of arity 3, beating the random ordering is Unique-Games hard. The result is in fact stronger: we show that for every $\Lambda \subseteq \Pi \subseteq S_3$, given an instance of permCSP(Λ) that is almost-satisfiable, it is hard to find an ordering that satisfies more than $\frac{|\Pi|}{6} + \varepsilon$ of the constraints even under the relaxed constraint Π (for arbitrary $\varepsilon > 0$). A special case of our result is that the *Betweenness* problem is hard to approximate beyond a factor $1/3$. Interestingly, for *satisfiable* instances of Betweenness, a factor $1/2$ approximation algorithm is known.

Thus, every permutation CSP of arity up to 3 resists approximation beyond the trivial random ordering threshold. In contrast, for Boolean CSPs, there are both approximation resistant and non-trivially approximable CSPs of arity 3.

Keywords-hardness of approximation; betweenness; permutation constraint satisfaction problems; approximation resistance

I. INTRODUCTION

Constraint satisfaction problems (CSPs) are a rich class of optimization problems that arise naturally in many settings. A large number of well-studied problems such as Max 3SAT, Max Cut, Max k-colorable subgraph, Max k-set splitting, Max 3LIN, Unique Games, etc. are examples of CSPs. A CSP of arity k over domain $[q] = \{0, 1, \ldots, q-1\}$ (called a q-ary kCSP) is specified by a predicate $P : [q]^k \to \{0, 1\}$. An instance of such a CSP consists of a set of *variables* V and a *collection* $\mathcal{C}$ of constraints each of which applies P to a k-tuple of *literals* (which are variables or their "negations," i.e., translates modulo q). The objective is to find an assignment $A : V \to [q]$ of values to the variables that maximizes the number of constraints of $\mathcal{C}$ that are satisfied.

For most CSPs, it is NP-hard to find an optimal assignment with maximum number of satisfied constraints. Therefore, one settles for approximation algorithms that deliver provable guarantees. Ideally, for each CSP, we would like to know its *approximation threshold*, i.e., a value $\alpha < 1$ for which we can give a polynomial time algorithm that satisfies at least α times the optimum number of constraints on every instance of that CSP[1], together with a matching hardness result that rules out a factor $(\alpha + \varepsilon)$-approximation for arbitrary constant $\varepsilon > 0$. For a CSP, we define its *random assignment threshold* to be the fraction ρ of assignments that satisfy its predicate P. A trivial algorithm that ignores the structure of the instance, and simply picks an assignment to variables randomly and independently, satisfied an expected fraction ρ of the constraints. (This algorithm can be derandomized using standard methods.) The approximation threshold is thus always at least the random assignment threshold.

Approximation Resistance. The discovery of the PCP theorem and semi-definite programming based approximation algorithms in the early 90's has led to a rich body of work which has identified the approximation threshold for many important CSPs. The breakthrough paper of Håstad [8] proved the striking result that many important CSPs are *approximation resistant*: it is NP-hard to approximate them better than their random assignment threshold, and thus their approximation thresholds equals their random assignment threshold ! The list of such approximation resistant CSPs include Max 3SAT, Max 3LIN (whose predicate stipulates that the parity of 3 literals is 0), and in fact any binary

[1]For convenience, we are focusing on maximization problems only.

3CSP whose predicate is implied by the parity constraint $x \oplus y \oplus z = 0$, Max k-set splitting for $k \geqslant 4$, etc.

This motivates the natural question of classifying approximation resistant predicates. Posed in this generality, this is a very challenging goal. But we now know fairly broad classes of CSPs which are approximation resistant, as well as those that are not. We now discuss some of these results. Complementing Håstad's hardness result for 3-CSPs, Zwick [16] gave approximation algorithms outperforming a random assignment for every 3-ary predicate not implied by parity, thereby leading to a precise classification of approximation resistant binary 3CSPs. The situation for arity 4 and higher gets more complicated as one might imagine. Hast succeeds in characterizing 355 out of 400 different predicate types for binary 4CSPs [7]. Håstad [9] showed that *most* predicates (a fraction $1 - o_1(k)$ of k-ary predicates) are approximation resistant (this relies on the Unique Games conjecture).

It is known that every 2CSP, even over non-binary domains, can be approximated better than the random assignment threshold [5], [4], [8]. The approximation threshold of 2CSPs (such as Max Cut) remained a fascinating mystery until recent progress based on the Unique Games conjecture (UGC) tied it to the integrality gap of semi-definite programming (SDP) relaxations [11], [1], [15]. In fact, under the UGC, Raghavendra showed the striking result [15] that for *every* CSP, the approximation threshold equals the integrality gap of a natural SDP. Thus believing the UGC is equivalent to believing that SDPs are the "best algorithms out there" for approximating CSPs.

Note that this result does not pinpoint the value of the approximation threshold, but only ties it to the integrality gap. For various CSPs, identifying the integrality gap itself is a challenging task (this is especially so for arity 3 and higher). So this connection per se does not advance the classification of approximation resistant predicates.

Permutation CSPs. As discussed above, there have been significant advances on the approximability of CSPs, both in algorithms and hardness results. However, a rich class of CSP-like problems that have resisted such progress are *ordering* (or *permutation*) problems. At a high level, an ordering problem seeks a permutation of the variables of a CSP (or vertices of a graph) to optimize some objective function. The list of ordering problems that have been studied in the literature include Bandwidth, Minimum Linear Arrangement, Feedback arc set, Maximum acyclic subgraph, Betweenness, etc.

Our understanding of approximability of ordering problems lags that of CSPs significantly. In terms of algorithms, for CSPs, one typically rounds a SDP/linear program solution into a small number of values (0/1 for binary CSPs), whereas for ordering problems, the rounding algorithm has to pick one of an unbounded number of values. In terms of hardness results, due to the bounded range for CSPs,

powerful techniques from the analysis of Boolean functions can be directly employed to establish strong results, whereas permutations tend to resist such an analysis.

In a recent work [6], a tight inappproximability result was shown for the Maximum Acyclic Subgraph (MAS) problem (assuming the UGC). In the MAS problem, one is given a collection of $u < v$ constraints (for variables $u, v \in V$) and the goal is to permute V so that a maximum number of these local constraints are met. A random permutation will satisfy an expected fraction $1/2$ of constraints, and the result of [6] shows that beating this is Unique Games-hard. In other words, MAS is approximation resistant. This was the first such result, and indeed the first tight hardness of approximation result, for an ordering problem.

In this paper, we study problems with more general ordering constraints. A permutation constraint satisfaction problem (permCSP) of arity k is specified by a subset $\Lambda \subseteq S_k$ of permutations on $\{1, 2, \ldots, k\}$. An instance of such a permCSP consists of a set of variables V and a collection of constraints each of which is an ordered k-tuple of V. The objective is to find a global ordering σ of the variables that maximizes the number of constraint tuples whose ordering (under σ) follows a permutation in Λ. This is just the natural extension of CSPs to the world of ordering problems. As with CSPs, we say a permCSP Λ is *approximation resistant* if its approximation threshold equals $\frac{|\Lambda|}{k!}$, which is the expected fraction of constraints satisfied by a random permutation of the variables.

Note that in this language, MAS corresponds to the simplest permCSP: the arity 2 permCSP where Λ consists of the identity permutation on two variables. Also note that MAS is the only non-trivial permCSP of arity 2.

Our main result is that *every* permCSP of arity 3 is approximation resistant. Specifically, for every such permCSP outperforming the trivial approximation ratio achieved by random ordering is Unique Games-hard. The result is in fact stronger:

Theorem 1. *For every $\varepsilon > 0$, $\Lambda \subseteq \Pi \subseteq S_3$, the following is Unique Games hard: given an instance of* permCSP(Λ) *that is $(1 - \varepsilon)$-satisfiable, it is hard to find an ordering that satisfies more than $|\Pi|/6 + \varepsilon$ of the constraints even under the relaxed constraint Π.*

A special case of our result is that the *Betweenness* problem is hard to approximate beyond a factor $1/3$. The Betweenness problem consists of constraints of the form "j lies between i and k" corresponding to the subset $\{123, 321\}$ of S_3. Interestingly, for *satisfiable* instances of Betweenness, it *is* possible to outperform the random ordering: using a semidefinite programming relaxation, Chor and Sudan [3] gave a factor $1/2$ approximation algorithm.

Thus, every permutation CSP of arity up to 3 resists approximation beyond the trivial random ordering threshold. In contrast, as mentioned above for binary CSPs, there are

both approximation resistant and non-trivially approximable CSPs of arity 3 (and we even precisely know which ones are in each category).

Our work suggests the intriguing possibility that this approximation resistance holds for larger arities as well, and that *every* permutation CSP might be approximation resistant. Such a result would be striking for its generality — it would highlight the fundamental difficulty of satisfying ordering constraints in a powerful way, and be in sharp contrast to the situation for binary and other bounded domain CSPs.

II. Proof overview and relation to [15], [6]

In this section, we describe the high level ideas behind inapproximability results for ordering problems, and place it in context with two closely related works [15], [6]. The underlying theme in these works is to convert some kind of combinatorial "gap instance" into a reduction from Unique Games (UG) which preserves this gap. After describing this framework, we then highlight the main technical contribution that is new to this work.

As explained above, Raghavendra [15] showed how to convert an integrality gap instance of a natural SDP for an arbitrary CSP into a matching hardness result. Consider a permCSP problem specified by $\Lambda \subseteq S_k$. An instance of permCSP(Λ) with m variables can be viewed as an instance of a CSP with domain size m, and viewed this way admits a similar SDP relaxation. Suppose there is a "c vs s" gap instance for this SDP that has m variables. In other words, the SDP optimum is at least c but the permCSP instance admits no ordering satisfying more than fraction s of constraints. Using the techniques of [15], one can then get a UG-hardness for an m-*ordering* version of permCSP(Λ). In this version, the goal is to map the variables to m ordered buckets (the variables inside a bucket are not ordered), and the payoff of a constraint is the probability that it is satisfied if each bucket is randomly and independently permuted. For example, if $\Lambda = 123$, the payoff for a constraint $i < j < k$ when i, j are placed in the same bucket and k is placed in a later bucket equals $1/2$.

In fact, one can conclude something stronger. Using Gaussian noise stability bounds, and the notion of influential coordinates for orderings put forth in [6], one can show the following fact in the soundness case: For the instance of permCSP(Λ) produced by the reduction, the value of the best ordering cannot be much higher than the best m-ordering. Thus one can conclude a "c vs s" inapproximability result for permCSP(Λ).

So why are we not done? The problem is that we do not know strong integrality gaps to use as starting point for the above reduction. In fact, even for the simplest permCSP, the MAS problem, tight integrality gaps were obtained only from the UG-hardness in [6]!

In light of this difficulty, the approach taken in [6] is to relax the requirements on the gap instance. Start with a gap instance on m vertices that has good SDP value, but much smaller t-ordering value, *for $t \ll m$*. Thus, the instance is only required to perform poorly on t-orderings and not on normal orderings (i.e., m-orderings). In fact, given the difference in objectives (m vs. t-orderings), we can even hope for an instance with a "c vs. s" gap between optimum *integral* m-orderings (not SDP vectors) and t-orderings. We can plug in such a gap instance into the above-mentioned UG reduction. This will yield an instance of permCSP(Λ) with m-ordering value (and hence also optimum ordering value) at least $c - \varepsilon$ for Yes instances of UG, and t-ordering value at most $s + \varepsilon$ for No instances of UG. In the latter case, using Gaussian noise stability bounds, the optimum ordering value will also be close to s (for t chosen large enough). We thus get a "c vs s" inapproximability result for permCSP(Λ).

For the MAS problem, the authors of [6] showed how to use a construction of a directed acyclic graph with low cut-norm due to Charikar, Makarychev, and Makarychev [2] to obtain the required gap instance. The gap instance is a directed graph on m vertices which is nearly acyclic, and yet for $t \ll m$ the MAS value of any t-ordering is close to $1/2$. We will refer to this graph as the CMM instance in the sequel.

Our technical contribution in this work is to construct an appropriate gap instance for arity 3 permutation CSPs with the completeness and soundness properties listed below. Plugging this instance into the above UG-reduction framework of [6] then gives our inapproximability result. (Specifically, this gives a strong hardness for permCSP(123), or the 3-ary monotone ordering problem, where the constraints are of the form $i < j < k$. Our general result for all 3-ary permCSPs follows by a simple reduction.)

We now describe the properties of the gap instance. For integer parameters $t < m$, and $\eta > 0$, the gap instance consists of a distribution D on $[m] \times [m] \times [m]$ with the following properties:

- Completeness: $\Pr_{(i,j,k) \in D}[i < j < k] \geqslant 1 - \eta$.
- Soundness: For every permutation $\pi \in S_3$ and every t-ordering $\mathcal{O}_t$ of $[m]$, the probability over random linear extensions of $\mathcal{O}_t$ that a sample $(i, j, k) \in D$ is ordered according to π is at most $\frac{1}{6} + \eta$.

We obtain such a distribution over triples by first sampling 3-tuples from the circle S^1 according to a suitable continuous distribution, and then discretizing the distribution. We use Fourier analysis over the circle S^1 to prove the soundness property. The continuous view also gives an alternate, somewhat simpler, explanation of why the CMM instance works in the MAS setting. The extension to the arity 3 case presents additional challenges and technical difficulties compared to the CMM analysis, and working in

the continuous setting makes it possible to overcome these with a reasonable economy of analysis.

A gap instance with similar properties for larger arity k would show that every permutation CSP is approximation resistant (under the UGC) via the same reduction framework. However, we do not know whether such gap instances exist for arities larger than 3. Our proof technique for triples proceeds by some sort of "reduction" to the CMM construction for arity 2. It is not clear if such a method can be extended to the larger arity case.

III. PRELIMINARIES

For a positive integer t, Sim_t denotes the the t dimensional simplex. We will use boldface letters $\mathbf{z}$ to denote vectors $\mathbf{z} = (z^{(1)}, \ldots, z^{(R)})$. Let $o_\tau(1)$ denote a quantity that tends to zero as $\tau \to 0$, while keeping all other parameters fixed. For an integer m, we denote by $[m]$ the set $\{1, 2, \ldots, m\}$. For notational convenience, an ordering $\mathcal{O}$ of a set of points V is represented by a map $\mathcal{O} : V \to \mathbb{Z}$. On the other hand, a t-ordering $\mathcal{O}$ consists of a map $\mathcal{O} : V \to [t]$. For a t-ordering, the map need not be injective or surjective, but an ordering is required to be injective.

An instance of the 3-*ary Monotone Ordering (3-MO)* problem, G, is a set of points V along with a (weighted) set E of 3-tuples of V. A 3-tuple (u, v, w) in E is said to be satisfied by an ordering $\mathcal{O}$ if $\mathcal{O}(u) < \mathcal{O}(v) < \mathcal{O}(w)$. The quantity $\mathsf{Val}(G)$ refers to the maximum over orderings of V the (weighted) fraction of satisfied constraints. We will use $\mathsf{Val}(G, \mathcal{O})$ to denote the fraction satisfied by a particular ordering $\mathcal{O}$.

The problem can be naturally extended to t-orderings as follows. A t-ordering $\mathcal{O}$ can be extended to a (total) ordering, $\mathcal{O}'$, by ordering within the t partitions randomly, while retaining the natural order amongst the partitions. Define the payoff of a tuple (u, v, w) in E in the t-ordering $\mathcal{O}$ is to be the probability that $\mathcal{O}'(u) < \mathcal{O}'(v) < \mathcal{O}'(w)$. For example, if $\mathcal{O}(u) = \mathcal{O}(v) < \mathcal{O}(w)$, then the tuple is ordered correctly with probability half and hence the payoff is $1/2$. The quantity $\mathsf{Val}_t(G)$ refers to the maximum expected total payoff (where the expectation is taken over choice of a random tuple from G) over all t-orderings.

We will be interested in general arity 3 ordering problems where the constraints may accept an arbitrary subset of the permutations of 3 elements. For a permutation π of 3 elements, the payoff of a tuple (u, v, w) with respect to π in a t-ordering $\mathcal{O}$ is the probability that the random extension $\mathcal{O}'$ orders the tuple according to π. The quantity $\mathsf{Val}_t^\pi(G)$ refers to the maximum expected payoff with respect to π in a t-ordering. The quantity $\mathsf{Val}_t(G)$, then, is a short-form for $\mathsf{Val}_t^{id}(G)$ where id is the identity permutation. As before, we will use $\mathsf{Val}_t^\pi(G, \mathcal{O})$ and $\mathsf{Val}_t(G, \mathcal{O})$ to denote the value obtained by a particular ordering $\mathcal{O}$. Finally, the quantity $\mathsf{Val}^\pi(G)$ is simply the maximum expected payoff with respect to π in any ordering. The following observations

follow immediately from the definition of the quantities involved.

Observation 2. *For all 3-MO instances G, and integers $t \leqslant t'$ and all π, $\mathsf{Val}_t^\pi(G) \leqslant \mathsf{Val}_{t'}^\pi(G) \leqslant \mathsf{Val}^\pi(G)$.*

Observation 3. *For all 3-MO instances G and integers t, for any ordering $\mathcal{O}$ of the points in G, $\sum_{\pi \in S_3} \mathsf{Val}_t^\pi(G, \mathcal{O}) = 1$.*

A. Fourier analysis over S^1

We review the basis facts that we will use about Fourier analysis of real valued functions defined on the circle $S^1 = \mathbb{R}/\mathbb{Z}$. A great reference for this topic is Körner's book [13]. For a Riemann integrable function $f : S^1 \to \mathbb{R}$, define the Fourier coefficients of f by

$$\hat{f}(n) = \int_0^1 f(t) e^{-2\pi int} dt \; ,$$

for integers $n \in \mathbb{Z}$. Under fairly general conditions, for example if f is continuous everywhere and has a continuous bounded derivative except at finitely many points, then the Fourier series $\sum_{n=-M}^{M} \hat{f}(n) e^{2\pi inx}$ converges uniformly to $f(x)$ as $M \to \infty$, and we can write

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{2\pi inx} \; .$$

We will only encounter such functions in our application. In this case, for such functions f and g, Parseval's identity says that

$$\int_{S^1} f(x) g(x) dx = \sum_{n=-\infty}^{\infty} \hat{f}(n) \overline{\hat{g}(n)} \; , \qquad (1)$$

which in particular implies that if $|f(x)| \leqslant 1$ for every x, then $\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 \leqslant 1$. If $f, g : S^1 \to \mathbb{R}$ are continuous functions, then their convolution $f * g$ is a continuous function given by

$$(f * g)(r) = \int_{S^1} f(r - u) g(u) \; du \; .$$

Its Fourier coefficients are given by $\widehat{f * g}(n) = \hat{f}(n) \hat{g}(n)$. The Fourier coefficients of the function $h(x) = f(-x)$ are given by $\hat{h}(n) = \hat{f}(-n)$.

B. Noise Operators and Influences

Let Ω denote the finite probability space corresponding to the uniform distribution over $[m]$. Let $\{\chi_0 = 1, \chi_1, \chi_2, \ldots, \chi_{m-1}\}$ be an orthonormal basis for the space $L_2(\Omega)$. For $\sigma \in [m]^R$, define $\chi_\sigma(\mathbf{z}) = \prod_{k \in [R]} \chi_{\sigma_i}(z^{(k)})$. Every function $\mathcal{F} : \Omega^R \to \mathbb{R}$ can be expressed as a multilinear polynomial as $\mathcal{F}(\mathbf{z}) = \sum_\sigma \hat{\mathcal{F}}(\sigma) \chi_\sigma(\mathbf{z})$. The L_2 norm of $\mathcal{F}$ in terms of the coefficients of the multilinear polynomial is $||\mathcal{F}||_2^2 = \sum_\sigma \hat{\mathcal{F}}^2(\sigma)$.

Definition 4. For a function $\mathcal{F} : \Omega^R \to \mathbb{R}$, define $\mathrm{Inf}_k(\mathcal{F}) = \mathbf{E}_\mathbf{z}[\mathbf{Var}_{z^{(k)}}[\mathcal{F}]] = \sum_{\sigma_k \neq 0} \hat{\mathcal{F}}^2(\sigma)$.

Here $\mathbf{Var}_{z^{(k)}}[\mathcal{F}]$ denotes the variance of $\mathcal{F}(\mathbf{z})$ over the choice of the kth coordinate $z^{(k)}$.

Definition 5. For a function $\mathcal{F} : \Omega^R \to \mathbb{R}$, define the function $T_\rho\mathcal{F}$ as follows:

$$T_\rho\mathcal{F}(\mathbf{z}) = \mathbf{E}[\mathcal{F}(\tilde{\mathbf{z}}) \mid \mathbf{z}] = \sum_{\sigma\in[m]^R} \rho^{|\sigma|}\hat{\mathcal{F}}(\sigma)\chi_\sigma(\mathbf{z})$$

where each coordinate $\tilde{z}^{(k)}$ of $\tilde{\mathbf{z}} = (\tilde{z}^{(1)}, \dots, \tilde{z}^{(R)})$ is equal to $z^{(k)}$ with probability ρ and with the remaining probability, $\tilde{z}^{(k)}$ is a random element from the distribution Ω.

The following result is established in [6] using the Majority is Stablest theorem (see Theorem 4.4 in [14]).

Lemma 6. *For every $\varepsilon > 0$, there exists a $\mu_0 > 0$ for which the following holds for all $\mu \leqslant \mu_0$. Suppose $\mathcal{F}, \mathcal{G} : [m]^R \to [0,1]$ are two functions with $\mathbf{E}[\mathcal{F}] = \mathbf{E}[\mathcal{G}] = \mu$, and $\mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{F}) \leqslant \tau$, $\mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{G}) \leqslant \tau$ for all $k \in [R]$. Let $\mathbf{x}, \mathbf{y}$ be random vectors in $[m]^R$ whose marginal distributions are uniform over $[m]^R$ but are arbitrarily correlated. Then*

$$\mathbf{E}_{\mathbf{x},\mathbf{y}}[T_{1-2\varepsilon}\mathcal{F}(\mathbf{x})T_{1-2\varepsilon}\mathcal{G}(\mathbf{y})] \leqslant \mu^{1+\varepsilon/2} + o_\tau(1) \ .$$

Lemma 7. *Given a function $\mathcal{F} : [m]^R \to [0,1]$, then $\sum_{k=1}^{R} \mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{F}) \leqslant \frac{1}{e \ln 1/(1-\varepsilon)} \leqslant \frac{1}{\varepsilon}$*

C. Unique Games

Definition 8. An instance of Unique Games represented as $\Upsilon = (\mathcal{A} \cup \mathcal{B}, E, \Pi, [R])$, consists of a bipartite graph over node sets $\mathcal{A}, \mathcal{B}$ with the edges E between them. Also part of the instance is a set of labels $[R] = \{1, \dots, R\}$, and a set of permutations $\pi_{ab} : [R] \to [R]$ for each edge $e = (a, b) \in E$. An assignment Λ of labels to vertices is said to satisfy an edge $e = (a, b)$, if $\pi_{ab}(\Lambda(a)) = \Lambda(b)$. The objective is to find an assignment Λ of labels that satisfies the maximum number of edges.

For a vertex $a \in \mathcal{A} \cup \mathcal{B}$, we shall use $N(a)$ to denote its neighborhood. For the sake of convenience, we shall use the following version of the Unique Games Conjecture [12] which is equivalent to the original conjecture [10].

Conjecture 9. *(Unique Games Conjecture [12], [10]) For all constants $\delta > 0$, there exists large enough constant R such that given a bipartite unique games instance $\Upsilon = (\mathcal{A} \cup \mathcal{B}, E, \Pi = \{\pi_e : [R] \to [R] \ : \ e \in E\}, [R])$ with number of labels R, it is NP-hard to distinguish between the following two cases:*

- *$(1-\delta)$-satisfiable instances: There exists an assignment Λ of labels such that for $1-\delta$ fraction of vertices $a \in \mathcal{A}$, all the edges incident at a are satisfied.*
- *Instances that are not δ-satisfiable: No assignment satisfies more than a δ-fraction of the constraints Π.*

D. Influential variables for orderings

We will use the notion of influence for orderings first introduced in [6].

Definition 10. Given an ordering $\mathcal{O}$ of vertices V, its t-coarsening is a t-ordering $\mathcal{O}^*$ obtained by dividing $\mathcal{O}$ into t contiguous blocks, and assigning label i to vertices in the ith block. Formally, if $M = |V|/t$ then

$$\mathcal{O}^*(u) = \left\lfloor \frac{|\{v|\mathcal{O}(v) < \mathcal{O}(u)\}|}{M} \right\rfloor + 1$$

For an ordering $\mathcal{O}$ of points in $[m]^R$, Define functions $\mathcal{F}_\mathcal{O}^{[p,q]} : [m]^R \to \{0,1\}$ for integers p, q as follows:

$$\mathcal{F}_\mathcal{O}^{[p,q]}(x) = \begin{cases} 1 & \text{if } \mathcal{O}(x) \in [p, q] \\ 0 & \text{otherwise} \end{cases}$$

We will omit the subscript and write $\mathcal{F}^{[p,q]}$ instead of $\mathcal{F}_\mathcal{O}^{[p,q]}$, when it is clear.

Definition 11. For an ordering $\mathcal{O}$ of $[m]^R$, define the set of influential coordinates $S_\tau(\mathcal{O})$ as follows:

$$S_\tau(\mathcal{O}) = \{k \mid \mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{F}^{[p,q]}) \leqslant \tau \text{ for some } p, q \in \mathbb{Z}\}$$

An ordering $\mathcal{O}$ is said to be τ-pseudo-random if $S_\tau(\mathcal{O})$ is empty.

Lemma 12 (Few Influential Coordinates). *[6] For any ordering $\mathcal{O}$ of $[m]^R$, we have $|S_\tau(\mathcal{O})| \leqslant \frac{400}{\varepsilon\tau^3}$*

Claim 13. *For any τ-pseudo-random ordering $\mathcal{O}$ of $[m]^R$, its t-coarsening $\mathcal{O}^*$ is also τ-pseudo-random.*

Proof: Since the functions $\{\mathcal{F}_{\mathcal{O}^*}^{[\cdot,\cdot]}\}$ are a subset of the functions $\{\mathcal{F}_\mathcal{O}^{[\cdot,\cdot]}\}$, $S_\tau(\mathcal{O}^*) \subseteq S_\tau(\mathcal{O})$. ∎

IV. COARSENING GAP INSTANCES FOR 3-MONOTONE ORDERING

Definition 14 ((η, t)-Coarsening Gap). A (weighted) 3-MO instance over $[m]$, $G = ([m], E)$ is a (η, t)-coarsening gap if:

- Completeness: $\Pr_{(i,j,k)\in E}[i < j < k] \geqslant 1 - \eta$. In other words, $\mathsf{Val}(G) \geqslant 1 - \eta$ and in particular is obtained by the obvious ordering of $[m]$.
- Soundness: For all permutations π of 3 elements,
$$|\mathsf{Val}_t^\pi(G) - \tfrac{1}{6}| \leqslant \eta$$

Note that it is easy to modify the construction to obtain a coarsening gap instance with perfect completeness—where $\mathsf{Val}(G)$ is 1—by simply throwing away the η fraction of bad tuples. This however does not help in the hardness reduction as the unique games instance has imperfect completeness. Moreover, the construction is more natural if allowed to have imperfect completeness.

The main result of this section is the following theorem which constructs (η, t)-coarsening gap for arbitrarily large t and arbitrarily small η.

Theorem 15. *For every integer $t > 0$ and $\eta > 0$, there exists an $m = m(\eta, t)$ such that there is a (η, t)-coarsening gap instance for 3-MO over $[m]$.*

A. Continuous gap Instances over S^1

We will prove Theorem 15 by first constructing a continuous analogue over the unit circle $S^1 = \mathbb{R}/\mathbb{Z}$. We will later discretize the instance to obtain the necessary result. In particular, we will analyze the following distribution over 3-tuples of points on S^1.

Definition 16. *For every $T > 0$, let D_T be the following distribution over $S^1 \times S^1 \times S^1$:*

- *Pick $x \in S^1$ uniformly at random*
- *Pick $\Delta \in [-T, 0]$ uniformly at random*
- *Pick $s \in [-\frac{\ln T}{2}, 0]$ uniformly at random*
- *Output $(x, x + e^{\Delta - Te^s}, x + e^{\Delta})$*

Remark 17 (Relation to CMM instance). A directed graph with m vertices with properties similar to the CMM instance, namely a large gap between the MAS value for m-orderings and t-orderings for $t \ll m$, can be obtained by a suitable discretization (see Section IV-B below) of the following distribution on $S^1 \times S^1$: Pick $x \in S^1$ uniformly at random, pick $s \in [-\ln T, 0]$ at random, and output $(x, y = x + e^s)$. This instance has roughly the same mass in all (geometric) "scales" of the jump between x and y. The fact that t-orderings have MAS value close to $1/2$ follows from the bound

$$\left| \mathbf{E}_{x,s}\Big(g(x)h(x + e^s) - h(x + e^s)g(x) \Big) \right| \leqslant O\left(\frac{1}{\ln T}\right) \quad (2)$$

for continuous functions $g, h : S^1 \to [0, 1]$, via an argument similar to (and in fact simpler than) Lemma 23 below. A bound similar to (2) is at the heart of the proof of Lemma 19 below.

The distribution in Definition 16 behaves similar to a CMM instance even on fixing x (see Lemma 19). Intuitively, this means that any t-ordering of the discretized instance cannot order the tuples significantly more or less often as $i < j < k$ than as $i < k < j$. Interestingly, as shown in the harder Lemma 20, a similar property is true even when fixing z. These two properties are enough to argue about every permutation of the tuples (i, j, k) as is shown in Lemma 23.

The proof of the following integral is in Appendix A.

Lemma 18. *Let m be an integer and a, b be such that $0 < a < b < 1$. Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a continuous, non-increasing function that is at least 1 in the interval $[a, b]$. Then,*

$$I = \left| \int_a^b \frac{\sin(2\pi m x)}{\psi(x)x} dx \right| \leqslant 14 .$$

The following lemma and Lemma 20 are at heart of arguing the soundness property (Definition 16) of the coarsening gap instance for t-orderings. This may not be apparent, but will be in the proof of Lemma 23.

Lemma 19. *For large enough T, for all continuous functions $f, g, h : S^1 \to [0, 1]$ with continuous derivatives except at finitely many points:*

$$\left| \mathbf{E}_{(x,y,z) \in D_T} \Big[f(x)\Big(g(y)h(z) - h(y)g(z)\Big) \Big] \right| \leqslant O\left(\frac{1}{\ln T}\right)$$

Proof: We have

$$\mathbf{E}[f(x)\{g(y)h(z) - g(z)h(y)\}]$$
$$= \int_0^1 f(x) \left[\int_{-T}^0 \int_{-\frac{\ln T}{2}}^0 g(x + e^{\Delta - Te^s})h(x + e^{\Delta}) \right.$$
$$\left. - h(x + e^{\Delta - Te^s})g(x + e^{\Delta} \frac{d\Delta}{T} \frac{2ds}{\ln T} \right] dx$$

We will bound the inner integrals for every x. Define the functions $g_1, h_1 : \mathbb{R} \to [0, 1]$ as

$$g_1(y) = g(x + e^{T(y-1)}) \text{ and } h_1(z) = h(x + e^{T(z-1)}) .$$

(Note these are defined with domain $\mathbb{R}$ and not S^1.) Now, setting $u = 1 + \Delta/T$ and $r = e^s$, for each x the inner integral can be written as:

$$I_x = \int_0^1 \int_{-\frac{1}{\sqrt{T}}}^1 \Big(g_1(u-r)h_1(u) - g_1(u)h_1(u-r) \Big) du \frac{2dr}{r \ln T}$$

Define functions $g_0, h_0 : S^1 \to [0, 1]$ as $g_0(\theta) = g_1(\theta)$ and $h_0(\theta) = h_1(\theta)$ where the argument θ in the right hand side (i.e., for g_1, h_1) is treated as a real in $[0, 1)$. Let

$$I'_x = \int_{u \in S^1} \int_{-\frac{1}{\sqrt{T}}}^1 [g_0(u-r)h_0(u) - g_0(u)h_0(u-r)] du \frac{2dr}{r \ln T} .$$

In other words, I'_x is the integral of the same integrand as in I_x but with operations over the unit circle S^1. Note that the expressions are different exactly when $u - r < 0$. Since each integrand is at most 1 in absolute value, we can bound

$$|I_x - I'_x| \leqslant 2 \int_0^1 \int_{\max\{-\frac{1}{\sqrt{T}}, u\}}^1 \frac{2dr}{r \ln T} du$$
$$\leqslant 2 \int_0^{1/\sqrt{T}} du + \frac{4}{\ln T} \int_{1/\sqrt{T}}^1 - \ln u \, du$$
$$\leqslant \frac{2}{\sqrt{T}} + \frac{4}{\ln T} \leqslant \frac{5}{\ln T} \quad (3)$$

for large enough T. To bound I'_x, we use Fourier analysis

over S^1 and proceed as follows:

$$I'_x = \int \int_{\frac{1}{\sqrt{T}}}^1 \Big(g_0(u-r)h_0(u) - g_0(u)h_0(u-r) \Big) \, du \frac{2dr}{r \ln T}$$

$$= \int_{\frac{1}{\sqrt{T}}}^1 \sum_{n=-\infty}^{\infty} \widehat{g_0}(-n)\widehat{h_0}(n)\big(e^{2\pi i n r} - e^{-2\pi i n r}\big) \frac{2dr}{r \ln T}$$

(using transform identity for convolution)

$$= \frac{4i}{\ln T} \sum_n \widehat{g_0}(-n)\widehat{h_0}(n) \int_{\frac{1}{\sqrt{T}}}^1 \frac{\sin(2\pi n r)}{r} \, dr \ .$$

Using Cauchy Schwarz, Parseval's identity, and Lemma 18, we get

$$|I'_x| \leqslant \frac{56}{\ln T} \sqrt{\sum_n |\widehat{g_0}(n)|^2 \sum_n |\widehat{h_0}(n)|^2} \leqslant O\left(\frac{1}{\ln T}\right) \quad (4)$$

Combining (3) and (4), we have the desired bound:

$$\Big| \mathbf{E}\big[f(x)\big(g(y)h(z) - g(z)h(y)\big)\big] \Big| = \left| \int_0^1 f(x) \cdot I_x \, dx \right|$$

$$\leqslant \int_0^1 |f(x)||I_x| dx \leqslant O\left(\frac{1}{\ln T}\right)$$

∎

The proof of the following lemma is more involved than the above proof, and is deferred to Appendix A.

Lemma 20. *For all large enough T, for all continuous functions $f, g, h : S^1 \to [0,1]$ with continuous derivatives except at finitely many points:*

$$\left| \mathbf{E}_{(x,y,z)\in D_T} \Big[h(z)\big(f(x)g(y) - f(y)g(x) \big) \Big] \right| \leqslant O\left(\frac{1}{\ln T}\right) \ .$$

B. Discrete Coarsening Gap Instance

Although the continuous setting is more amenable to analytic tools, we will need the instance to be over a finite set for constructing the dictatorship test. We will discretize the distribution D_T to obtain the coarsening gap instance.

Definition 21 (3-MO Gap Instance, G_T^m). Partition S^1 into m intervals $[0, \frac{1}{m}), [\frac{1}{m}, \frac{2}{m}) \ldots [\frac{m-1}{m}, 1)$, identified with $[m]$ in the obvious way. G_T^m is a weighted 3-MO instance over $[m]$, where the weight of a tuple (i,j,k) is the probability that D_T outputs points x, y, and z in the partitions corresponding i, j, and k respectively.

Lemma 22. *For $m > e^{2T}$, the weight of tuples (i,j,k) such that $i < j < k$ is at least $1 - \frac{1}{T}$.*

Proof: Suppose the tuple (x,y,z) did not wrap around the unit interval (in other words, $x < y < z$). Then, the minimum distance between points, $\min_{(x,y,z)\leftarrow D_T}\{y - x, z - y, z - x\}$ is at least e^{-2T}. For $m > e^{2T}$, the tuples (i,j,k) obtained by discretizing (x,y,z) are over distinct and increasing i, j and k. Thus, unless the points (x,y,z)

wrap-around the unit interval, the tuple (i,j,k) will be ordered as $i < j < k$. Hence,

$$\mathsf{Wt}(i < j < k) = \mathsf{Wt}(i < k) \geqslant \Pr_{(x,y,z)\leftarrow D_T}[x < z]$$

(As $i < k$ implies $i < j < k$ in the construction)

$$\geqslant 1 - \int_{-T}^0 e^\Delta \frac{d\Delta}{T} = 1 - \frac{1 - e^{-T}}{T} \geqslant 1 - \frac{1}{T}$$

∎

Lemma 23. *For every positive integer t, for large enough m, T, the following holds for every $\pi \in S_3$:*

$$\left| \mathsf{Val}_t^\pi(G_T^m) - \tfrac{1}{6} \right| \leqslant O\left(\frac{t^3}{\ln T}\right)$$

Proof: Fix a t-ordering $\mathcal{O}$. We will bound the difference between $\mathsf{Val}_t^{\pi_1}(G_T^m, \mathcal{O})$ and $\mathsf{Val}_t^{\pi_2}(G_T^m, \mathcal{O})$ for every two permutations π_1, π_2.

Let $P = \{P_1, P_2 \ldots P_t\}$ be the partition of $[m]$ into t pieces induced by $\mathcal{O}$. Abusing notation, we will use $P(i)$ to denote the partition to which i belongs and P_a to denote the indicator of the part P_a.

$$\mathsf{Val}_t(G_T^m, \mathcal{O}) = \mathsf{Val}_t^{123}(G_T^m, \mathcal{O}) = \Pr[P(i) < P(j) < P(k)]$$
$$+ \frac{1}{2}\big(\Pr[P(i) = P(j) < P(k)] + \Pr[P(i) < P(j) = P(k)] \big)$$
$$+ \frac{1}{6} \Pr[P(i) = P(j) = P(k)]$$
$$= \sum_{a<b<c} \mathbf{E}[P_a(i)P_b(j)P_c(k)] + \frac{1}{2} \sum_{a<c} \mathbf{E}[P_a(i)P_a(j)P_c(k)]$$
$$+ \frac{1}{2} \sum_{a<c} \mathbf{E}[P_a(i)P_c(j)P_c(k)] + \frac{1}{6} \sum_a \mathbf{E}[P_a(i)P_a(j)P_a(k)]$$

Similarly,

$$\mathsf{Val}_t^{132}(G_T^m, \mathcal{O}) = \Pr[P(i) < P(k) < P(j)]$$
$$+ \frac{1}{2}\big(\Pr[P(i) = P(k) < P(j)] + \frac{1}{2} \Pr[P(i) < P(k) = P(j)] \big)$$
$$+ \frac{1}{6} \Pr[P(i) = P(k) = P(j)]$$
$$= \sum_{a<b<c} \mathbf{E}[P_a(i)P_b(k)P_c(j)] + \frac{1}{2}\Big(\sum_{a<c} \mathbf{E}[P_a(i)P_a(k)P_c(j)]$$
$$+ \sum_{a<c} \mathbf{E}[P_a(i)P_c(k)P_c(j)] \Big) + \frac{1}{6} \sum_a \mathbf{E}[P_a(i)P_a(k)P_a(j)]$$

From the above expressions,

$$\left| \mathsf{Val}_t^{123}(G_T^m, \mathcal{O}) - \mathsf{Val}_t^{132}(G_T^m, \mathcal{O}) \right|$$
$$\leqslant \sum_{a<b<c} \mathbf{E}_{(i,j,k)} \left| \Big[P_a(i)P_b(j)P_c(k) - P_a(i)P_b(k)P_c(j) \Big] \right|$$

Defining function $f_a : S^1 \to [0,1]$ to be the indicator of the union of the intervals corresponding to the partition P_a, made continuous is by connecting endpoints of the intervals

by a "steep" line. We can bound the difference in expectation using Lemma 19 as follows:

$$\mathbf{E}_{(i,j,k)}\left|\left[P_a(i)P_b(j)P_c(k) - P_a(i)P_b(k)P_c(j)\right]\right|$$

$$= \mathbf{E}_{(x,y,z)\leftarrow D_T}\left|\left[f_a(x)\Big(f_b(y)f_c(z) - f_c(y)f_b(z)\Big)\right]\right|$$

$$\leqslant O\left(\frac{1}{\ln T}\right)$$

We therefore conclude $|\mathsf{Val}_t^{123}(G_T^m, \mathcal{O}) - \mathsf{Val}_t^{132}(G_T^m, \mathcal{O})| \leqslant O\left(\frac{t^3}{\ln T}\right)$. An almost identical proof bounds the differences $|\mathsf{Val}_t^{213}(G_T^m, \mathcal{O}) - \mathsf{Val}_t^{312}(G_T^m, \mathcal{O})|$ and $|\mathsf{Val}_t^{231}(G_T^m, \mathcal{O}) - \mathsf{Val}_t^{321}(G_T^m, \mathcal{O})|$. Using a similar argument and employing Lemma 20 instead of Lemma 19, the differences $|\mathsf{Val}_t^{123}(G_T^m, \mathcal{O}) - \mathsf{Val}_t^{213}(G_T^m, \mathcal{O})|$, $|\mathsf{Val}_t^{132}(G_T^m, \mathcal{O}) - \mathsf{Val}_t^{231}(G_T^m, \mathcal{O})|$, $|\mathsf{Val}_t^{312}(G_T^m, \mathcal{O}) - \mathsf{Val}_t^{321}(G_T^m, \mathcal{O})|$ can also be bounded by $O\left(\frac{t^3}{\ln T}\right)$ from above.

Thus, for all π_1, π_2, $|\mathsf{Val}_t^{\pi_1}(G_T^m, \mathcal{O}) - \mathsf{Val}_t^{\pi_2}(G_T^m, \mathcal{O})| \leqslant O\left(\frac{t^3}{\ln T}\right)$. Now, Observation 3 gives the required result. ∎

Given (η, t), setting $T = \exp(\Omega(\frac{t^3}{\eta}))$ and setting $m = \exp(\exp(\Omega(\frac{t^3}{\eta})))$ and using Lemma 22 and Lemma 23 proves Theorem 15.

V. DICTATORSHIP TEST

Let $G = (V, E)$ be a (η, t)-coarsening gap instance for the 3-MO. Without loss of generality, $V = [m]$ for an m divisible by t. Using G, construct a (weighted) 3-MO dictatorship test instance DICT_G on orderings $\mathcal{O}$ of $[m]^R$ as follows:

DICT_G **Test:**
- Pick a tuple $e = (u, v, w) \in E$ at random.
- Pick $\mathbf{z}_e = (\mathbf{z}_u, \mathbf{z}_v, \mathbf{z}_w)$ to be k random shifts of e. In other words, pick k random integers $r_1, r_2 \ldots r_k$ from $[m]$ and set $z_u^i = u + r_i$, $z_v^i = v + r_i$, $z_w^i = w + r_i \pmod{m}$.
- Obtain $\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v, \tilde{\mathbf{z}}_w$ by perturbing each coordinate of $\mathbf{z}_u$, $\mathbf{z}_v$, and $\mathbf{z}_w$ independently. Specifically, sample the kth coordinates $\tilde{z}_u^{(k)}, \tilde{z}_v^{(k)}$ as follows: With probability $(1 - 2\varepsilon)$, $\tilde{z}_u^{(k)} = z_u^{(k)}$, and otherwise $\tilde{z}_u^{(k)}$ is a new sample from V.
- Output tuple $(\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v, \tilde{\mathbf{z}}_w)$.

Theorem 24. *(Soundness Analysis) For every $\varepsilon > 0$, for any τ-pseudo-random ordering $\mathcal{O}$ of $[m]^R$,*

$$\mathsf{Val}^{\pi}(\mathcal{O}) \leqslant \mathsf{Val}_t^{\pi}(G) + O(t^{-\frac{\varepsilon}{2}}) + o_{\tau}(1)$$

for all permutations π, where $o_{\tau}(1) \rightarrow 0$ as $\tau \rightarrow 0$ keeping all other parameters fixed.

Let $\mathcal{F}^{[p,q]} : [m]^R \rightarrow \{0, 1\}$ denote the functions associated with the t-ordering $\mathcal{O}^*$. For the sake of brevity, we shall write $\mathcal{F}^i$ for $\mathcal{F}^{[i,i]}$. The result follows from Lemma 25 and Lemma 26 below, whose proofs parallel those of similar statements in [6].

Lemma 25. *For every $\varepsilon > 0$, there exists sufficiently large m, t such that : For any τ-pseudo-random ordering $\mathcal{O}$ of $[m]^R$ and permutation π,*

$$\mathsf{Val}^{\pi}(\mathcal{O}) \leqslant \mathsf{Val}_t^{\pi}(\mathcal{O}^*) + O(t^{-\frac{\varepsilon}{2}}) + o_{\tau}(1)$$

where $\mathcal{O}^*$ is the t-coarsening of $\mathcal{O}$.

Proof: As $\mathcal{O}^*$ is a coarsening of $\mathcal{O}$, clearly $\mathsf{Val}_t^{\pi}(\mathcal{O}^*) \leqslant \mathsf{Val}_t^{\pi}(\mathcal{O})$. Note that the loss due to coarsening, is because some tuples $e = (\mathbf{u}, \mathbf{v}, \mathbf{w})$ which are oriented correctly in $\mathcal{O}$, fall into the same block during coarsening, i.e $\mathcal{O}^*(\mathbf{u})$, $\mathcal{O}^*(\mathbf{v})$, $\mathcal{O}^*(\mathbf{w})$ are not all distinct. Thus we can write

$$\mathsf{Val}(\mathcal{O}) \leqslant \mathsf{Val}_t(\mathcal{O}^*) + \Pr\left(\mathcal{O}^*(\tilde{\mathbf{z}}_u) = \mathcal{O}^*(\tilde{\mathbf{z}}_v)\right)$$
$$+ \Pr\left(\mathcal{O}^*(\tilde{\mathbf{z}}_u) = \mathcal{O}^*(\tilde{\mathbf{z}}_w)\right) + \Pr\left(\mathcal{O}^*(\tilde{\mathbf{z}}_v) = \mathcal{O}^*(\tilde{\mathbf{z}}_w)\right)$$

$$\Pr\left(\mathcal{O}^*(\tilde{\mathbf{z}}_u) = \mathcal{O}^*(\tilde{\mathbf{z}}_v)\right)$$
$$= \sum_{i\in[t]} \mathbf{E}_{e=(u,v)}\mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v}\mathbf{E}_{\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v}\left[\mathcal{F}^i(\tilde{\mathbf{z}}_u) \cdot \mathcal{F}^i(\tilde{\mathbf{z}}_v)\right]$$
$$= \sum_{i\in[t]} \mathbf{E}_{e=(u,v)}\mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v}\left[T_{1-2\varepsilon}\mathcal{F}_u^i(\mathbf{z}_u) \cdot T_{1-2\varepsilon}\mathcal{F}_v^i(\mathbf{z}_v)\right]$$

As $\mathcal{O}$ is a t-coarsening of $\mathcal{O}$, for each value $i \in [t]$, there are exactly $\frac{1}{t}$ fraction of $\mathbf{z}$ for which $\mathcal{O}^*(\mathbf{z}) = i$. Hence for each $i \in [t]$, $\mathbf{E}_{\mathbf{z}}[\mathcal{F}^i(\mathbf{z}) = \frac{1}{t}]$. Further, since the ordering $\mathcal{O}^*$ is τ-pseudo-random, for every $k \in [R]$ and $i \in [t]$, $\mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{F}_a^i) \leqslant \tau$. Hence using for sufficiently large t, the above probability is bounded by $t \cdot t^{-1-\frac{\varepsilon}{2}} + t \cdot o_{\tau}(1) = O(t^{-\frac{\varepsilon}{2}}) + o_{\tau}(1)$. The same bound holds for the other two probabilities too, hence giving the required result. ∎

Lemma 26. *For every choice of m, t, ε, and any τ-pseudo-random t-ordering $\mathcal{O}^*$ of $[m]^R$, $\mathsf{Val}_t^{\pi}(\mathcal{O}^*) \leqslant \mathsf{Val}_t^{\pi}(G) + o_{\tau}(1)$.*

Proof: Deferred to Appendix A. ∎

VI. HARDNESS REDUCTION

Let $G = (V, E)$ be a (η, t)-coarsening gap instance, and let $m = |V|$. Let $\Upsilon = (\mathcal{A} \cup \mathcal{B}, E, \Pi = \{\pi_e : [R] \rightarrow [R] | e \in E\}, [R])$ be a bipartite unique games instance. Towards constructing a 3-MO instance $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ from Υ, we shall introduce a long code for each vertex in $\mathcal{B}$. Specifically, the set of vertices $\mathcal{V}$ of $\mathcal{G}$ is indexed by $\mathcal{B} \times [m]^R$.

Hardness Reduction:

Input : Unique games instance $\Upsilon = (\mathcal{A} \cup \mathcal{B}, E, \Pi = \{\pi_e : [R] \rightarrow [R] | e \in E\}, [R])$ and a (η, t) coarsening gap instance $G = (V, E)$.

Output : 3-MO instance $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with set of vertices : $\mathcal{V} = \mathcal{B} \times [m]^R$ and edges $\mathcal{E}$ given by the following verifier:

- Pick a random vertex $a \in \mathcal{A}$. Choose three neighbors $b, b', b'' \in \mathcal{B}$ independently at random. Let π, π', π'' denote the permutations on the edges (a, b), (a, b') and (a, b'').
- Pick a tuple $e = (u, v, w) \in E$ at random from the coarsening gap instance G.
- Sample $\mathbf{z}_e = \{\mathbf{z}_u, \mathbf{z}_v, \mathbf{z}_w\}$ using k random shifts of $\mathbf{z}_e$ as in the dictatorship test.
- Obtain $\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v, \tilde{\mathbf{z}}_w$ by perturbing each coordinate of $\mathbf{z}_u$, $\mathbf{z}_v$, and $\mathbf{z}_w$ independently.
- Output tuple $((b, \pi(\tilde{\mathbf{z}}_u)), (b', \pi'(\tilde{\mathbf{z}}_v)), (b'', \pi''(\tilde{\mathbf{z}}_w)))$.

The proof of the following theorem follows from the soundness analysis (Theorem 24) of the dictatorship test and using techniques similar to [6] and is omitted.

Theorem 27. *For every $\gamma > 0$, there exists choice of parameters $\varepsilon, \eta, t, \delta$ such that:*

- COMPLETENESS: *If Υ is a $(1 - \delta)$-satisfiable instance of Unique Games, then there is an ordering $\mathcal{O}$ for the graph $\mathcal{G}$ with value at least $(1 - \gamma)$. i.e. $\mathsf{Val}(\mathcal{G}) \leqslant 1 - \gamma$*
- SOUNDNESS: *If Υ is not δ-satisfiable, then no ordering to $\mathcal{G}$ has value more than $\frac{1}{6} + \gamma$ with respect to any permutation π, i.e, for all permutations π, $\mathsf{Val}^\pi(\mathcal{G}) \leqslant \frac{1}{6} + \gamma$*

Given a hard instance $\mathcal{G}$ with the above properties, we can give an easy reduction to a hard instance for $\mathsf{permCSP}(\Lambda)$. In fact, in the soundness case we can even allow a more relaxed set of constraints $\Pi \supseteq \Lambda$, and it will still be hard to outperform a random ordering.

Lemma 28. *For every $\Lambda \subseteq \Pi \subseteq S_3$, there is a polynomial time reduction from 3-MO instance, G to $\mathsf{permCSP}(\Lambda)$ instance H such that:*

- COMPLETENESS: $\mathsf{Val}^\Lambda(H) \geqslant \mathsf{Val}(G)$.
- SOUNDNESS: *If, for some γ, $\mathsf{Val}^\pi(G) \leqslant \frac{1}{6} + \frac{\gamma}{6}$ for all permutations π, then $\mathsf{Val}^\Pi(H) \leqslant \frac{|\Pi|}{6} + \gamma$.*

Proof: Given $\Lambda \subseteq \Pi \subseteq S_3$ and a 3-MO instance G, pick an arbitrary permutation $\pi \in \Lambda$ and permute the tuples of G by π to obtain the instance H. It is easy to see that for every permutation π', $\mathsf{Val}^{\pi \circ \pi'}(H) = \mathsf{Val}^{\pi'}(G)$. In particular, since $\pi \in \Lambda$, $\mathsf{Val}^\Lambda(H) \geqslant \mathsf{Val}^\pi(H) \geqslant \mathsf{Val}(G)$. Similarly, if $\mathsf{Val}^{\pi'}(G) \leqslant \frac{1}{6} + \frac{\gamma}{6}$ for all permutations π', then $\mathsf{Val}^\Pi(H) = \sum_{\pi' \in \Pi} \mathsf{Val}^{\pi'}(H) \leqslant \frac{|\Pi|}{6} + \gamma$ ∎

Our final result follows immediately follows from Theorem 27 and Lemma 28.

Theorem 29 (Main). *Conditioned on the Unique Games conjecture, for every $\Lambda \subseteq \Pi \subseteq S_3$, given an instance of $\mathsf{permCSP}(\Lambda)$ that is $1 - \varepsilon$ satisfiable, it is hard to find an ordering that satisfies more than $|\Pi|/6 + \varepsilon$ of the constraints even under the relaxed constraint Π (for arbitrary $\varepsilon > 0$).*

REFERENCES

[1] P. Austrin. Towards sharp inapproximability for any 2-csp. In *FOCS*, pages 307–317. IEEE Computer Society, 2007.

[2] M. Charikar, K. Makarychev, and Y. Makarychev. On the advantage over random for maximum acyclic subgraph. In *FOCS*, pages 625–633. IEEE Computer Society, 2007.

[3] B. Chor and M. Sudan. A geometric approach to betweenness. *SIAM Journal on Discrete Mathematics*, 11(4):511–523, Nov. 1998.

[4] L. Engebretsen and V. Guruswami. Is constraint satisfaction over two variables always easy? *Random Struct. Algorithms*, 25(2):150–178, 2004.

[5] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995.

[6] V. Guruswami, R. Manokaran, and P. Raghavendra. Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph. In *FOCS*, pages 573–582. IEEE Computer Society, 2008.

[7] G. Hast. *Beating a Random Assignment*. PhD thesis, Royal Institute of Technology, Stockholm, Sweden, 2005.

[8] J. Håstad. Every 2-csp allows nontrivial approximation. In H. N. Gabow and R. Fagin, editors, *STOC*, pages 740–746. ACM, 2005.

[9] J. Håstad. On the approximation resistance of a random predicate. In M. Charikar, K. Jansen, O. Reingold, and J. D. P. Rolim, editors, *APPROX-RANDOM*, volume 4627 of *Lecture Notes in Computer Science*, pages 149–163. Springer, 2007.

[10] S. Khot. On the power of unique 2-prover 1-round games. In *IEEE Conference on Computational Complexity*, page 25, 2002.

[11] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for max-cut and other 2-variable csps? *SIAM J. Comput.*, 37(1):319–357, 2007.

[12] S. Khot and O. Regev. Vertex cover might be hard to approximate to within 2-epsilon. *J. Comput. Syst. Sci.*, 74(3):335–349, 2008.

[13] T. Körner. *Fourier Analysis*. Cambridge University Press, 1988.

[14] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. In *FOCS*, pages 21–30. IEEE Computer Society, 2005.

[15] P. Raghavendra. Optimal algorithms and inapproximability results for every csp? *In STOC '08: Proceedings of the 40th ACM Symposium on Theory of Computing*, 2008.

[16] U. Zwick. Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint. In *SODA*, pages 201–210, 1998.

APPENDIX

A. Omitted proofs

Lemma 30 (Lemma 18). *Let m be an integer and a, b be such that $0 < a < b < 1$. Let $\psi : \mathbb{R}^+ \to \mathbb{R}^+$ be a continuous, non-increasing function that is at least 1 in the interval $[a, b]$. Then,*

$$I = \left| \int_a^b \frac{\sin(2\pi m x)}{\psi(x)x} dx \right| \leqslant 14 \ .$$

Proof: w.l.o.g., we can assume m is positive. Set $l = \lceil ma \rceil$ and $n = \lfloor mb \rfloor$. If $l > n$, the interval $[a, b]$ does not contain any multiple of $1/m$ and thus $b - a \leqslant 1/m$. In this case the integral is at most

$$I = \left| \int_a^b \frac{\sin(2\pi m x)}{x \psi(x)} \, dx \right| \leqslant \int_a^b \left| \frac{\sin(2\pi m x)}{x \psi(x)} \right| \, dx$$

$$\leqslant \int_a^b \frac{|2\pi m x|}{x \psi(x)} \, dx \leqslant 2\pi m (b - a) \leqslant 2\pi$$

Therefore, assume $l \leqslant n$. The absolute value of the integral can be bounded as:

$$I \leqslant \left| \int_a^{\frac{l}{m}} \frac{\sin(2\pi m x)}{\psi(x)x} dx \right| + \left| \sum_{k=l}^{n-1} \int_{\frac{k}{m}}^{\frac{k+1}{m}} \frac{\sin(2\pi m x)}{\psi(x)x} dx \right|$$

$$+ \left| \int_{\frac{n}{m}}^b \frac{\sin(2\pi m x)}{\psi(x)x} dx \right|$$

$$\leqslant 4\pi + \sum_{k=l}^{n-1} \left| \int_{\frac{k}{m}}^{\frac{k+1}{m}} \frac{\sin(2\pi m x)}{\psi(x)x} dx \right|$$

Using the identity $\sin(\pi + x) = -\sin(x)$, we can split the interval $[k/m, (k+1)/m]$ into two halves and group terms

as follows:

$$\left| \int_{\frac{k}{m}}^{\frac{k+1}{m}} \frac{\sin(2\pi m x)}{x \psi(x)} \, dx \right| = \left| \int_{\frac{k}{m}}^{\frac{k+\frac{1}{2}}{m}} \sin(2\pi m x) \left(\frac{1}{x \psi(x)} \right. \right.$$

$$\left. \left. - \frac{1}{\psi(x + \frac{1}{2m})(x + \frac{1}{2m})} \right) dx \right|$$

$$\leqslant \int_{\frac{k}{m}}^{\frac{k+\frac{1}{2}}{m}} \left| \frac{1}{x \psi(x)} - \frac{1}{\psi(x + \frac{1}{2m})(x + \frac{1}{2m})} \right| \, dx$$

$$\leqslant \int_{\frac{k}{m}}^{\frac{k+\frac{1}{2}}{m}} \frac{dx}{2mx^2} \max \left\{ \frac{1}{\psi(x)}, \frac{1}{\psi(x + \frac{1}{2m})} \right\}$$

$$\leqslant \int_{\frac{k}{m}}^{\frac{k+\frac{1}{2}}{m}} \frac{dx}{2mx^2} \leqslant \frac{1}{4k^2}$$

We conclude that $I \leqslant 4\pi + \frac{1}{4} \sum_{k=1}^\infty \frac{1}{k^2} \leqslant 4\pi + \frac{\pi^2}{24} \leqslant 14$. ∎

Lemma 31 (Lemma 20). *For all large enough T, for all continuous functions $f, g, h : S^1 \to [0, 1]$ with continuous derivatives except at finitely many points:*

$$\left| \mathbf{E}_{(x,y,z) \in D_T} \left[h(z) \Big(f(x)g(y) - f(y)g(x) \Big) \right] \right| \leqslant O \left(\frac{1}{\ln T} \right) \ .$$

Proof: The distribution D_m is equivalent to $(z - e^\Delta, z - e^\Delta + e^{\Delta - Te^s}, z)$ where $z \in S^1$ is chosen uniformly at random, and Δ, s are chosen as in D_m. Thus, the expectation can be written as follows:

$$\mathbf{E}\left[h(z) \Big(f(x)g(y) - g(x)f(y) \Big) \right]$$

$$= \int_0^1 h(z) \left[\int_{-T}^0 \int_{-\frac{\ln T}{2}}^0 f(z - e^\Delta) g(z - e^\Delta + e^{\Delta - Te^s}) \right.$$

$$\left. - g(z - e^\Delta) f(z - e^\Delta + e^{\Delta - Te^s}) \frac{d\Delta}{T} \frac{2ds}{\ln T} \right] dz$$

As in the previous lemma, we will bound the inner integral for every z. Write $e^\Delta (1 - e^{-Te^s})$ as $e^{\Delta - rT}$. For large T, as s varies from 0 to $-\frac{\ln T}{2}$, r varies from $r_0 \approx e^{-T}/T$ to $r_1 \approx e^{-\sqrt{T}}/T$.

Differentiating $e^{-rT} = 1 - e^{-Te^s}$, we can compute the probability density function of r, and it is given by

$$-p(r) \, dr = \frac{2Te^{-rT} dr}{\ln T (1 - e^{-rT}) \ln(\frac{1}{1 - e^{-rT}})} = \frac{2ds}{\ln T}$$

(The negative sign in front of $p(r)dr$ is because r decreases when s increases.)

We will bound the inner integral above for every z. Now, as in the proof of Lemma 19, define functions $f_1, g_1 : \mathbb{R} \to [0, 1]$ by

$$f_1(x) = f(z - e^{T(x-1)}) \text{ and } g_1(y) = g(z - e^{T(y-1)})$$

(Note these are defined with domain $\mathbb{R}$ and not S^1.) Setting $u = 1 + t/T$, the inner integral can be written as

$$I_z = \int_0^1 \int_{r_0}^{r_1} \Big(f_1(u)g_1(u-r) - g_1(u)f_1(u-r) \Big) du \; p(r) dr$$

Further, r varies from $r_0 \approx e^{-T}/T$ to $r_1 \approx e^{-\sqrt{T}}/T$; hence for large enough T, $1 - e^{-rT} \simeq rT$. More formally, we will approximate $p(r)$ by a simpler function for the purpose of the analysis. Setting $h = rT$, we see that,

$$\left| \frac{e^{-h}}{(1-e^{-h})\ln(\frac{1}{1-e^{-h}})} - \frac{1}{h\ln(\frac{1}{h})} \right|$$

$$\leq \left| \frac{e^{-h}}{(1-e^{-h})\ln(\frac{1}{1-e^{-h}})} - \frac{e^{-h}}{h\ln(\frac{1}{1-e^{-h}})} \right|$$

$$+ \left| \frac{e^{-h}}{h\ln(\frac{1}{1-e^{-h}})} - \frac{e^{-h}}{h\ln(\frac{1}{h})} \right| + \left| \frac{e^{-h}}{h\ln(\frac{1}{h})} - \frac{1}{h\ln(\frac{1}{h})} \right|$$

$$\leq \left| \frac{e^{-h}}{\ln(\frac{1}{1-e^{-h}})} \left(\frac{1}{1-e^{-h}} - \frac{1}{h} \right) \right|$$

$$+ \left| \frac{e^{-h}}{h} \left(\frac{1}{\ln(\frac{1}{1-e^{-h}})} - \frac{1}{\ln(\frac{1}{h})} \right) \right| + \left| \frac{e^{-h}-1}{h\ln(\frac{1}{h})} \right|$$

For our setting of parameters, we have $h/2 \leq 1 - e^{-h} \leq h$ and $|\ln(\frac{1-e^{-h}}{h})| \leq 5h$. Hence,

$$\left| \frac{e^{-h}}{(1-e^{-h})\ln(\frac{1}{1-e^{-h}})} - \frac{1}{h\ln(\frac{1}{h})} \right|$$

$$\leq 5 \left| \frac{e^{-h}}{\ln(\frac{1}{h})} \left(\frac{h-1-e^{-h}}{h^2} \right) \right|$$

$$+ \left| \frac{e^{-h}}{h} \left(\frac{\ln(\frac{1-e^{-h}}{h})}{\ln^2(\frac{1}{h})} \right) \right| + \left| \frac{e^{-h}-1}{h\ln(\frac{1}{h})} \right| \leq \frac{10}{\ln(\frac{1}{h})}$$

Thus,

$$\left| p(r) - \frac{2}{\ln(T)r\ln(\frac{1}{rT})} \right| \leq \frac{40T}{\ln T \ln(\frac{1}{rT})} \leq \frac{40T}{\ln T} \; .$$

Hence, it is enough to bound the simpler integral

$$I_z' = \int_0^1 \int \frac{2\Big(f_1(u)g_1(u-r) - g_1(u)f_1(u-r) \Big) du dr}{\ln(T)r\ln(\frac{1}{rT})}$$

More precisely, the difference between the two integrals is at most:

$$|I_z - I_z'| \leq \frac{200}{\ln T} e^{-\sqrt{T}} \leq O\left(\frac{1}{\ln T} \right)$$

As in the proof of the previous lemma, define functions $f_0, g_0 : S^1 \to [0,1]$ as $f_0(\theta) = f_1(\theta)$ and $g_0(\theta) = g_1(\theta)$

where the argument θ in the right hand side (i.e., for f_1, g_1) is treated as a real between 0 and 1. Now let

$$I_z'' = \int \int_{r_0}^{r_1} \frac{2\Big(f_0(u)g_0(u-r) - g_0(u)f_0(u-r) \Big) du dr}{\ln(T)r\ln(\frac{1}{rT})}$$

To bound I_z'', as in the proof of Lemma 19, we use Fourier analysis over S^1, transform identity for convolutions, Lemma 18, Cauchy-Schwarz, and Parseval's theorem, and proceed as follows:

$$|I_z''| = \left| \int_{S^1} \int_{r_0}^{r_1} [f_0(u)g_0(u-r) - g_0(u)f_0(u-r)] du \right.$$

$$\left. \frac{2}{\ln(T)r\ln(\frac{1}{rT})} dr \right|$$

$$= \left| \frac{4i}{\ln T} \sum_{n=-\infty}^{\infty} \widehat{g_0}(-n)\widehat{f_0}(n) \int_{r_0}^{r_1} \frac{\sin(2\pi n r)}{r \ln(\frac{1}{rT})} dr \right|$$

$$\leq \frac{56}{\ln T} \sqrt{ \sum_n |\widehat{f_0}(n)|^2 \sum_n |\widehat{g_0}(n)|^2 } \leq O\left(\frac{1}{\ln T} \right)$$

Finally, since r is at most $e^{-\sqrt{\Delta}}/T$, the difference between the integrals I_z'' and I_z' is also at most $e^{-\sqrt{\Delta}}/T$ as for a fixed r, $u < r$ with probability at most r. We therefore reach our desired conclusion

$$\mathbb{E}\left[h(z)\Big(f(x)g(y) - g(x)f(y) \Big) \right] = \left| \int_0^1 h(z) \; I_z \; dz \right|$$

$$\leq \int_0^1 |h(z)| \; |I_z| \; dz \leq O\left(\frac{1}{\ln T} \right)$$

■

Lemma 32 (Lemma 26). *For every choice of m, t, ε, and any τ-pseudo-random t-ordering $\mathcal{O}^*$ of $[m]^R$, $\mathsf{Val}_t^\pi(\mathcal{O}^*) \leq \mathsf{Val}_t^\pi(G) + o_\tau(1)$.*

Proof: We will look at the case where π is the identity permutation. The proof for the other orderings is almost identical. The t-ordering problem is a CSP over a finite domain, and is thus amenable to techniques of [15]. Specifically, consider the payoff function $P : [t]^3 \to [0,1]$ defined by: $P(i,j,k) = 1$ for $i < j < k$, $P(i,j,k) = 1/2$ for $i = j < k$ and $i < j = k$, $P(i,j,k) = 1/6$ for $i = j = k$ and $P(i,j) = 0$ otherwise. The t-ordering problem (with respect to the identity permutation) is a Generalized CSP(see Definition 3.1, [15]) with the payoff function P.

For the sake of exposition, let us pretend that $t = m$. The dictatorship test we construct essentially picks a random integral solution from the m different integral solutions (corresponding to the m cyclic shifts of $1 \ldots m$) to perform the test. This is exactly the same as the t-ordering CSP dictatorship DICT_G obtained by running the reduction of [15] on the trivial SDP solution corresponding to the convex combination of the integral solutions. Further, all the SDP constraints are satisfied as the SDP solution is integral.

A t-ordering solution $\mathcal{O}^*$ for the dictaDICT$_G$ corresponds naturally to a function $\mathcal{F} : [t]^R \to \mathsf{Sim}_t$. Thus, we have the following observations:

- The t-ordering instance DICT$_G$ is identical to the dictatorship test described in this section when $t = m$.
- For a τ-pseudo-random t-ordering $\mathcal{O}^*$, for every $k \in [R]$ and $i \in [t]$, the corresponding function $\mathcal{F}$ satisfies $\mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{F}^i) \leqslant \tau$. In the terminology of [15](Definitions 4.1 and 4.2), this is equivalent to the function $\mathcal{F} = (\mathcal{F}^1, \ldots, \mathcal{F}^t)$ being "(γ, τ)-pseudo-random" with $\gamma = 0$.
- By Corollary 2.2 in [15], for a (γ, τ)-pseudo-random function $\mathcal{F}$, its probability of acceptance on the dictatorship test is at most $\mathsf{Val}_t(G) + o_{\gamma,\tau}(1)$.

Hence the above lemma is just a restatement of Corollary 2.2 of [15] for the specific generalized CSP – t-Ordering – albeit in the language of τ-pseudo-random orderings.

Recall that the actual case of interest here satisfies $t < m$. Unfortunately, in this case, a black box application of the result from [15] does not suffice. However, the proof in [15] can be easily adopted without any new technical ideas. In fact, many of the technical difficulties encountered in [15] can be avoided here. For instance, the SDP solution associates with each vertex u, the uniform probability distribution over $\{1 \ldots m\}$ (due to the cyclic shifts), unlike [15] where there are several arbitrary probability distributions to deal with. With the value of m fixed, this removes the need for smoothing the SDP solution (Lemma 3.4 in [15]). ∎