Beating the Random Ordering is Hard: Inapproximability of Maximum Acyclic Subgraph

VENKATESAN GURUSWAMI * RAJSEKAR MANOKARAN † PRASAD RAGHAVENDRA ‡

August 18, 2008

Abstract

We prove that approximating the MAX ACYCLIC SUBGRAPH problem within a factor better than $1/2$ is Unique-Games hard. Specifically, for every constant $\varepsilon > 0$ the following holds: given a directed graph G that has an acyclic subgraph consisting of a fraction $(1-\varepsilon)$ of its edges, if one can efficiently find an acyclic subgraph of G with more than $(1/2+\varepsilon)$ of its edges, then the UGC is false. Note that it is trivial to find an acyclic subgraph with $1/2$ the edges, by taking either the forward or backward edges in an arbitrary ordering of the vertices of G. The existence of a ρ-approximation algorithm for $\rho > 1/2$ has been a basic open problem for a while.

Our result is the first tight inapproximability result for an *ordering* problem. The starting point of our reduction is a directed acyclic subgraph (DAG) in which every cut is nearly-balanced in the sense that the number of forward and backward edges crossing the cut are nearly equal; such DAGs were constructed in [3]. Using this, we are able to study MAX ACYCLIC SUBGRAPH, which is a constraint satisfaction problem (CSP) over an *unbounded* domain, by relating it to a proxy CSP over a *bounded* domain. The latter is then amenable to powerful techniques based on the invariance principle [13, 19].

Our results also give a super-constant factor inapproximability result for the MIN FEEDBACK ARC SET problem. Using our reductions, we also obtain SDP integrality gaps for both the problems.

1 Introduction

Given a directed acyclic graph G, one can efficiently order ("topological sort") its vertices so that all edges go forward from a lower ranked vertex to a higher ranked vertex. But what if a few, say fraction ε, of edges of G are reversed? Can we detect these "errors" and find an ordering with few back edges? Formally, given a directed graph whose vertices admit an ordering with many, i.e., $1-\varepsilon$ fraction, forward edges, can we find a good ordering with fraction α of forward edges (for some $\alpha \to 1$)? This is equivalent to finding a subgraph of G that is acyclic and has many edges, and hence this problem is called the MAX ACYCLIC SUBGRAPH (MAS) problem.

*Department of Computer Science & Engineering, University of Washington, Seattle, WA 98195. Work done while on leave at the Institute for Advanced Study, School of Mathematics, Princeton, NJ 08540. Research supported in part by a Packard Fellowship, and NSF grant CCR-0324906 to The IAS

†Supported by NSF grants MSPA-MCS 0528414, and ITR 0205594

‡Work done while the author was visiting Princeton University. Supported by NSF grant CCF-0343672

It is trivial to find an ordering with fraction 1/2 of forward edges: take the better of an arbitrary ordering and its reverse. This gives a factor 1/2 approximation algorithm for Max Acyclic Subgraph. (This is also achieved by picking a *random* ordering of the vertices.) Despite much effort, no efficient ρ-approximation algorithm for a constant $\rho > 1/2$ has been found for Max Acyclic Subgraph. The existence of such an algorithm has been a longstanding and central open problem in the theory of approximation algorithms. In this work, we prove a strong hardness result that rules out the existence of such an approximation algorithm assuming the Unique-Games conjecture. Our main result is the following.

Theorem 1.1. *Conditioned on the Unique Games conjecture, the following holds for every constant $\gamma > 0$. Given a directed graph G with m edges, it is NP-hard to distinguish between the following two cases:*

1. *There is an ordering of the vertices of G with at least $(1-\gamma)m$ forward edges (or equivalently, G has an acyclic subgraph with at least $(1-\gamma)m$ edges).*
2. *For every ordering of the vertices of G, there are at most $(1/2+\gamma)m$ forward edges (or equivalently, every subgraph of G with more than $(1/2+\gamma)m$ edges contains a directed cycle).*

To the best of our knowledge, the above is the first tight hardness of approximation result for an ordering/permutation problem. As an immediate consequence, we obtain the following hardness result for the complementary problem of Min Feedback Arc Set, where the objective is to minimize the number of back edges.

Corollary 1.2. *Conditioned on the Unique Games conjecture, for every $C > 0$, it is NP-hard to find a C-approximation to the* Min Feedback Arc Set *problem.*

Combining the unique game integrality gap instance of Khot-Vishnoi [10] along with the UG reduction, we obtain SDP integrality gaps for Max Acyclic Subgraph problem. Our integrality gap instances also apply to a related SDP relaxation studied by Newman [17]. This SDP relaxation was shown to obtain an approximation better than half on random graphs which were previously used to obtain integrality gaps for a natural linear program [15].

1.1 Related work

Max Acyclic Subgraph is a classic optimization problem, figuring in Karp's early list of NP-hard problems [7]; the problem remains NP-hard on graphs with maximum degree 3, when the in-degree plus out-degree of any vertex is at most 3. Max Acyclic Subgraph is also complete for the class of permutation optimization problems, MAX SNP[π], defined in [18], that can be approximated within a constant factor. It is shown in [15] that Max Acyclic Subgraph is NP-hard to approximate within a factor greater than $\frac{65}{66}$.

Turning to algorithmic results, the problem is known to be efficiently solvable on planar graphs [11, 6] and reducible flow graphs [20]. Berger and Shor [2] gave a polynomial time algorithm with approximation ratio $1/2+\Omega(1/\sqrt{d_{\max}})$ where $d_{\max}$ is the maximum vertex degree in the graph. When $d_{\max} = 3$, Newman [15] gave a factor 8/9 approximation algorithm.

The complementary objective of minimizing the number of back edges, or equivalently deleting the minimum number of edges in order to make the graph a DAG, leads to the Min Feedback Arc Set (FAS) problem. This problem admits a factor $O(\log n \log\log n)$ approximation algorithm [21] based on bounding the integrality gap of the natural covering linear program for FAS; see also [4]. Using this algorithm, one can get an approximation ratio of $\frac{1}{2} + \Omega(1/(\log n \log\log n))$ for Max Acyclic Subgraph.

Recently, Charikar, Makarychev, and Makarychev [3] gave a factor $(1/2+\Omega(1/\log n))$-approximation algorithm for Max Acyclic Subgraph, where n is the number of vertices. In fact, their algorithm is stronger: given a digraph with an acyclic subgraph consisting of a fraction $(1/2 + \delta)$ of edges, it finds a subgraph with at least a fraction $(1/2 + \Omega(\delta/\log n))$ of edges. This algorithm, and in particular an instance showing tightness of its analysis from [3], plays a crucial role in our work.

1.2 Organization

We begin with an outline of the key ideas of the proof in Section 2. In Section 3, we review the definitions of influences, noise operators and restate the unique games conjecture. The groundwork for the reduction is laid in Section 4 and Section 5, where we define influences for orderings, and multiscale gap instances respectively. We present the dictatorship test in Section 6, and convert it to a UG hardness result in Section 7. Finally, SDP integrality gaps for Max Acyclic Subgraph are presented in Section 8.

2 Proof Overview

In this section, we outline the central ideas of the proof. To keep the description concise, we will set up some basic notation. For sake of brevity, let us denote $[m] = \{1, \ldots, m\}$. Given an ordering $\mathcal{O}$ of the vertices of a directed graph $G = (V, E)$, let $\mathsf{Val}(\mathcal{O})$ refer to the fraction of the edges E that are oriented correctly in $\mathcal{O}$.

At the heart of all Unique Games based hardness results, lies a dictatorship testing result for an appropriate class of functions. A function $\mathcal{F} : [m]^R \to [m]$ is said to be a *dictator* if $\mathcal{F}(x) = x_i$ for some fixed i. A dictatorship test (DICT) is a randomized algorithm such that, given a function $\mathcal{F} : [m]^R \to [m]$, it makes a few queries to the values of $\mathcal{F}$ and distinguishes between whether $\mathcal{F}$ is a dictator or *far* from every dictator. While Completeness of the test refers to the probability of acceptance of a dictator function, Soundness is the maximum probability of acceptance of a function *far* from a dictator. The approximation problem one is showing UG hardness for, determines the nature of the dictatorship test needed for the purpose.

Now let us turn to the specific problem at hand : Max Acyclic Subgraph. Designing the appropriate dictatorship test for this problem amounts to the following: Construct a directed graph over the set of vertices $V = [m]^R$ such that :

- For a *Dictator* ordering $\mathcal{O}$ of the vertices V, $\mathsf{Val}(\mathcal{O}) \approx 1$
- For any ordering $\mathcal{O}$ which is *far from a dictator*, $\mathsf{Val}(\mathcal{O}) \approx \frac{1}{2}$.

Unlike the case of functions, it is unclear as to what is the right notion of *Dictators* for orderings. For every ordering $\mathcal{O}$ of $[m]^R$, define m^{2R} functions $\mathcal{F}^{[p,q]} : [m]^R \to \{0, 1\}$ as follows:

$$\mathcal{F}^{[p,q]}(x) = \begin{cases} 1 & \text{if } x \text{ appears in between the } p^{th} \text{ and the } q^{th} \text{positions in ordering } \mathcal{O} \\ 0 & \text{otherwise} \end{cases}$$

The i^{th} coordinate is said to be *influential* if it has a large influence ($> \tau$) on any of the functions $\mathcal{F}^{[p,q]}$. Here influence refers to the natural notion of influence for real valued functions on $[m]^R$ (see Section 3). An ordering $\mathcal{O}$ is said to be τ-pseudorandom (*far* from a dictator) if it has no influential coordinates ($> \tau$). For this notion to be useful, it is necessary that a given ordering $\mathcal{O}$ does not have too many *influential* coordinates. Towards this, in Lemma 4.3 we show that the number of influential coordinates is bounded (after certain smoothening). Further this notion of influence is well suited to deal with orderings of multiple long codes instead of one - a crucial requirement in translating dictatorship tests to UG hardness.

Armed with the notion of influential coordinates, we obtain a directed graph on $[m]^R$ (a dictatorship test) for which the following holds:

Theorem 2.1. *(Soundness) If $\mathcal{O}$ is any τ-pseudorandom ordering of $[m]^R$, then $\mathsf{Val}(\mathcal{O}) \leqslant \frac{1}{2} + o_\tau(1)$.*

This dictatorship test yields tight UG hardness for the MAX ACYCLIC SUBGRAPH problem. Using the Khot-Vishnoi [10] SDP gap gap instance for unique games, we obtain SDP integrality gap for the MAX ACYCLIC SUBGRAPH problem.

Now we describe the design of the dictatorship test in greater detail. At the outset, the approach is similar to recent work on Constraint Satisfaction Problems(CSPs) [19]. Fix a constraint satisfaction problem Λ. Starting with an integrality gap instance Φ for the natural semi-definite program for Λ, [19] constructs a dictatorship test DICT_Φ. The Completeness of DICT_Φ is equal to the SDP value $\mathsf{SDP}(\Phi)$, while the Soundness is close to the integral value $\mathsf{INT}(\Phi)$.

Since the result of [19] applies to arbitrary CSPs, a natural direction would be to pose the MAX ACYCLIC SUBGRAPH as a CSP. MAX ACYCLIC SUBGRAPH is fairly similar to a CSP, with each vertex being a variable taking values in domain $[n]$ and each directed edge a constraint between 2 variables. However, the domain, $[n]$, of the CSP is not fixed, but grows with input size. We stress here that this is not a superficial distinction but an essential characteristic of the problem. For instance, every 2-CSP over a domain of fixed size admits an approximation ratio better than a random assignment [5], while the MAX ACYCLIC SUBGRAPH problem has resisted such approximation algorithms.

Towards using techniques from the CSP result, we define the following variant of MAX ACYCLIC SUBGRAPH:

Definition 2.2. A t-ordering of a directed graph $G = (V, E)$ consists of a map $\mathcal{O} : V \to [t]$. The value of a t-ordering $\mathcal{O}$ is given by

$$\mathsf{Val}_t(\mathcal{O}) = \Pr_{(u,v)\in E}\Big(\mathcal{O}(u) < \mathcal{O}(v)\Big) + \frac{1}{2}\Pr_{(u,v)\in E}\Big(\mathcal{O}(u) = \mathcal{O}(v)\Big)$$

In the t-Order problem, the objective is to find an t-ordering of the input graph G with maximum value.

On the one hand, the t-Order problem is a CSP over a fixed domain that is similar to MAS. However, to the best of our knowledge, for the t-Order problem, there are no known SDP gaps, which constitute the starting point for results in [19]. For any fixed constant t, Charikar, Makarychev and Makarychev [3] construct directed acyclic graphs (i.e., with value of the best ordering equal to 1), while the value of any t-ordering of G is close to $\frac{1}{2}$. For the rest of the discussion, let us fix one such graph G on m vertices. Notice that the graph G does not serve as SDP gap example for either the MAS or the t-Order problem.

As the graph G has only m vertices, and an ordering of value ≈ 1, it has a good t-ordering for $t = m$. Viewing G as an instance of the m-Order CSP (corresponding to predicate $<$), we obtain a directed graph, $\mathcal{G}$, on $[m]^R$. As a m-order CSP, the dictator m-orderings yield value ≈ 1 on $\mathcal{G}$. In turn, this implies that the *Dictator* orderings have value ≈ 1 on $\mathcal{G}$. Turning to the soundness proof, consider a τ-pseudorandom ordering $\mathcal{O}$. Obtain a t-ordering $\mathcal{O}^*$ by the following *coarsening* process : Divide the ordering $\mathcal{O}$ in to t equal blocks, and map the vertices in the i^{th} block to value i. The crucial observation relating $\mathcal{O}$ and $\mathcal{O}^*$ is as follows:

"For a τ-pseudorandom ordering $\mathcal{O}$, $\mathsf{Val}_t(\mathcal{O}^*) \approx \mathsf{Val}(\mathcal{O})$."

Clearly, $\mathsf{Val}(\mathcal{O}) - \mathsf{Val}_t(\mathcal{O}^*)$ is bounded by the fraction of edges whose both endpoints fall in the same block, during the coarsening. We use the Gaussian noise stability bounds of [13], to bound the fraction of such edges. From the above observation, in order to prove that $\mathsf{Val}(\mathcal{O}) \approx \frac{1}{2}$, it is enough to bound $\mathsf{Val}_t(\mathcal{O}^*)$. Notice that $\mathcal{O}^*$ is a solution to t-order problem - a CSP over finite domain. Consequently, the soundness analysis of [19] can be used to show that $\mathsf{Val}_t(\mathcal{O}^*)$ is at most the value of the best t-ordering for G, which is close to $\frac{1}{2}$.

Summarizing the key ideas, we define the notion of influential coordinates for orderings, and then use it to construct a dictatorship test for orderings. Using Gaussian noise stability bounds, we relate the value of a pseudorandom ordering to a related CSP, and then apply techniques from [19].

3 Preliminaries

For a positive integer t, Δ_t denotes the the t dimensional simplex. We will use bold face letters $\mathbf{z}$ to denote vectors $\mathbf{z} = (z^{(1)}, \ldots, z^{(R)})$. A t-ordering $\mathcal{O}$ of the graph G consists of a map $\mathcal{O} : V \to [t]$. Note that the map $\mathcal{O}$ need not be injective or surjective. If the map $\mathcal{O}$ is a bijection, then it corresponds to an ordering of the vertices V. In a t-ordering $\mathcal{O}$, an edge $e = (u, v)$ is a *forward* edge if $\mathcal{O}(u) < \mathcal{O}(v)$.

Observation 3.1. *For all directed graphs G, and integers $t \leqslant t'$, $\mathsf{Val}_t(G) \leqslant \mathsf{Val}_{t'}(G) \leqslant \mathsf{Val}(G)$*

3.1 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 3.2. 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 k^{th} coordinate $z^{(k)}$.

Definition 3.3. 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 $\tilde{\mathbf{z}} = (\tilde{z}^{(1)}, \ldots, \tilde{z}^{(R)})$ is equal to $z^{(k)}$ with probability ρ and with the remaining probability, $\tilde{z}^{(k)}$ is a random element from the distribution Ω.

We will need the following simple facts.

Lemma 3.4. *Given a function $\mathcal{F} : [m]^R \to [0,1]$, if $\mathcal{H} = T_{1-\epsilon}\mathcal{F}$ then $\sum_{k=1}^{R} \mathrm{Inf}_k(\mathcal{H}) \leqslant \frac{1}{e \ln 1/(1-\epsilon)} \leqslant \frac{1}{\epsilon}$*

Proof. Let $\mathcal{F}(x) = \sum_{\sigma \in [m]^R} \hat{\mathcal{F}}_\sigma \chi_\sigma(x)$ denote the expansion of $\mathcal{F}$. The function $\mathcal{H}$ is given by $\mathcal{H}(x) = \sum_\sigma (1-\epsilon)^{|\sigma|} \hat{\mathcal{F}}_\sigma \chi_\sigma(x)$. Hence we get,

$$\sum_{k=1}^{R} \mathrm{Inf}_k(\mathcal{H}) = \sum_{k=1}^{R} \sum_{\sigma, \sigma_k \neq 0} (1-\epsilon)^{2|\sigma|} \hat{\mathcal{F}}_\sigma^2 = \sum_{\sigma \in [m]^R} (1-\epsilon)^{2|\sigma|} |\sigma| \hat{\mathcal{F}}_\sigma^2$$
$$\leqslant \max_{\sigma \in [m]^R} \left((1-\epsilon)^{2|\sigma|} |\sigma| \right) \cdot \sum_\sigma \hat{\mathcal{F}}_\sigma^2 \leqslant \max_{\sigma \in [m]^R} (1-\epsilon)^{2|\sigma|} |\sigma|$$

The function $\phi(x) = x(1-\epsilon)^{2x}$ achieves a maximum at $x = -1/2 \ln(1-\epsilon)$. Substituting we get $\sum_{k=1}^{R} \mathrm{Inf}_k(\mathcal{H}) \leqslant \frac{1}{e \ln 1/(1-\epsilon)}$.

□

Lemma 3.5. *Consider two functions $\mathcal{F}, \mathcal{G} : [m]^R \to [0,1]$ with $\mathbf{E}[\mathcal{F}] = \mathbf{E}[\mathcal{G}] = \mu$, such that for all k, $\mathrm{Inf}_k(T_{1-\epsilon}\mathcal{F}), \mathrm{Inf}_k(T_{1-\epsilon}\mathcal{G}) \leqslant \tau$. Let $\mathbf{x}, \mathbf{y}$ be random vectors in $[m]^R$ whose marginal distributions are uniform over $[m]^R$ but are arbitrarily correlated. For small enough μ, we have*

$$\mathbf{E}_{\mathbf{x},\mathbf{y}}[T_{1-2\epsilon}\mathcal{F}(\mathbf{x}) T_{1-2\epsilon}\mathcal{G}(\mathbf{y})] \leqslant \mu^{\frac{2}{2-\epsilon}}$$

Proof. The lemma essentially follows from the Majority is Stablest theorem (see Theorem 4.4 in [14]). We bound each factor individually as follows:

$$||T_{1-2\epsilon}\mathcal{F}||_2^2 = \sum_{\sigma \in [k]^R} (1-2\epsilon)^{2|\sigma|} \hat{\mathcal{F}}^2(\sigma)$$
$$\leqslant \sum_{\sigma \in [k]^R} (1-\epsilon)^{|\sigma|} \hat{\mathcal{F}}^2(\sigma) (1-\epsilon)^{2|\sigma|} \hat{\mathcal{F}}^2(\sigma) \leqslant \mathbf{E}[(T_{1-\epsilon}\mathcal{F}) T_{1-\epsilon}(T_{1-\epsilon}\mathcal{F})]$$

Now, since the influences of $\mathcal{F}$ are low, the last expression can be bounded by the noise stability in Gaussian space, $\Gamma_{(1-\epsilon)}(\mu)$. This can now be bounded using standard estimates (see Theorem B.2 in [14]).

$$\mathbf{E}[(T_{1-\epsilon}\mathcal{F}) T_{1-\epsilon}(T_{1-\epsilon}\mathcal{F})] \leqslant \Gamma_{(1-\epsilon)}(\mu) + o_\tau(1) \leqslant \mu^{\frac{2}{2-\epsilon}} + o_\tau(1)$$

Applying a similar bound for $\mathcal{G}$ and applying Cauchy-Schwartz gives the result:

$$\mathbf{E}_x[T_{1-2\epsilon}\mathcal{F}(x) T_{1-2\epsilon}\mathcal{G}(y)] \leqslant \sqrt{||T_{1-2\epsilon}\mathcal{F}||_2^2 ||T_{1-2\epsilon}\mathcal{G}||_2^2}$$
$$\leqslant \mu^{\frac{2}{2-\epsilon}} + o_\tau(1) \qquad \text{(for } \mu \text{ small enough)}$$

□

3.2 Semidefinite Program

We use the following natural SDP relaxation of the MAX ACYCLIC SUBGRAPH problem. Given a directed graph $G = (V, E)$ with $|V| = n$, the program has n variables $\{u_1, \ldots, u_n\}$ for each vertex $u \in V$. In the intended solution, the variable $u_i = 1$ and $u_j = 0$ for all $j \neq i$ if and only if u is assigned position i.

$$\begin{array}{lll}
\text{Maximize} & \mathbf{E}_{e=(u,v)}\Big[\sum_{i<j} u_i \cdot v_j + \frac{1}{2}\sum_i u_i \cdot v_i\Big] & \text{(MAS-SDP)} \\
\text{Subject to} & u_i \cdot v_j \geq 0,\ u_i \cdot u_j = 0 & \forall u, v \in V, i, j \in [n] \\
& \sum_{i\in[n]} |u_i|^2 = 1 & \forall u \in V \\
& \Big|\sum_{i\in[n]} u_i - \sum_{i\in[n]} v_i\Big|^2 = 0 & \forall u, v \in V
\end{array}$$

As shown in [16], this relaxation is equivalent to the relaxation in [17].

3.3 Unique Games

Definition 3.6. An instance of Unique Games represented as $\Gamma = (\mathcal{A} \cup \mathcal{B}, E, \Pi, [R])$, 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 $[R] = \{1, \ldots, R\}$, 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[9] which was shown to be equivalent to the original conjecture by [8].

Conjecture 3.7. *(Unique Games Conjecture [9, 8]) For all constants $\delta > 0$, there exists large enough constant R such that given a bipartite unique games instance $\Gamma = (\mathcal{A} \cup \mathcal{B}, E, \Pi = \{\pi_e : [R] \to [R] \ : \ e \in E\}, [R])$ 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.*

4 Orderings

In this section, we develop the notions of influences for orderings and prove some basic results about it.

Definition 4.1. Given an ordering $\mathcal{O}$ of vertices V, its t-coarsening is a t-ordering $\mathcal{O}^*$ obtained by dividing $\mathcal{O}$ in to t-contiguous blocks, and assigning label i to vertices in the i^{th} block. Formally, if $M = |V|/t$ then

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

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

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

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

Lemma 4.3. *(Few Influential Coordinates) For any ordering $\mathcal{O}$ of $[m]^R$, we have $|S_\tau(\mathcal{O})| \leqslant \frac{400}{\epsilon\tau^3}$*

Proof. For integers p, q, δ_1, δ_2 such that $|\delta_i| < \frac{\tau}{8}m^R$, let $f = T_{1-\epsilon}\mathcal{F}^{[p,q]}$ and $g = T_{1-\epsilon}\mathcal{F}^{[p+\delta_1,q+\delta_2]}$. Now,

$$\mathrm{Inf}_k(f-g) \leqslant ||f-g||_2^2 \leqslant ||\mathcal{F}^{[p,q]} - \mathcal{F}^{[p+\delta_1,q+\delta_2]}||_2^2 = \Pr_{\mathbf{z}}[\mathcal{F}^{[p,q]}(\mathbf{z}) \neq \mathcal{F}^{[p+\delta_1,q+\delta_2]}(\mathbf{z})] \leqslant \tau/4$$

Hence,

$$\begin{aligned}\mathrm{Inf}_k(f) = \sum_{\sigma_k \neq 0} \hat{f}^2(\sigma) &\leqslant 2\left[\sum_{\sigma_k\neq 0} \hat{g}^2(\sigma) + \sum_{\sigma_k \neq 0}\left(\hat{f}(\sigma) - \hat{g}(\sigma)\right)^2\right] \quad (\text{Using } a^2 \leqslant 2(b^2 + (a-b)^2)) \\ &\leqslant 2\mathrm{Inf}_k(g) + \tau/2\end{aligned}$$

Thus, if $\mathrm{Inf}_k(f) \geqslant \tau$, then $\mathrm{Inf}_k(g) \geqslant \tau/4$. It is easy to see that there is a set $N = \{\mathcal{F}^{[p,q]}\}$ of size at most $100/\tau^2$ such that for every $\mathcal{F}^{[p,q]}$ there is a $\mathcal{F}^{[r,s]} \in N$ such that $\max |p-r|, |q-s| < \frac{\tau m^R}{8}$. Further, by Lemma 3.4, the functions $T_{1-\epsilon}\mathcal{F}^{[p,q]}$ have at most $\frac{4}{\epsilon\tau}$ coordinates with influence more than $\tau/4$. Hence, $|S_\tau(\mathcal{O})| \leqslant \frac{400}{\epsilon\tau^3}$. □

Claim 4.4. *For any τ-pseudorandom ordering $\mathcal{O}$ of $[m]^R$, its t-coarsening $\mathcal{O}^*$ is also τ-pseudorandom.*

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

5 Multiscale Gap Instances

In this section, we will construct acyclic directed graphs with no good t-ordering. These graphs will be crucial in designing the dictatorship test (Section 6).

Definition 5.1. For $\eta > 0$ and a positive integer t, a (η, t)-*Multiscale Gap* instance is a weighted directed graph $G = (V, E)$ with the following properties:

- $\mathsf{Val}(G) = 1$ and $\mathsf{Val}_t(G) \leqslant \frac{1}{2} + \eta$
- There exists a solution $\{u_i \mid u \in V, 1 \leqslant i \leqslant |V|\}$ to SDP with objective value at least $1 - \eta$ such that for all $u, v \in V$ and $1 \leqslant i, j \leqslant |V|$, we have $|u_i|^2 = \frac{1}{|V|}$.

The cut norm of a directed graph, G, represented by a skew-symmetric matrix W is:

$$||G||_C = \max_{x_i, y_j \in \{0,1\}} \sum_{ij} x_i y_j w_{ij}$$

We will need the following theorem from [3] relating the cut norm of a directed graph G to $\mathsf{Val}(G)$.

Theorem 5.2 (Theorem 3.1, [3])**.** *If a directed graph G on n vertices has a maximum acyclic subgraph with at least a $\frac{1}{2} + \delta$ fraction of the edges, then, $||G||_C \geqslant \Omega\left(\frac{\delta}{\log n}\right)$.*

The following lemma and its corollary construct Multiscale Gap instances starting from graphs that are the "tight cases" of the above theorem.

Lemma 5.3. *For every $\eta > 0$ and a positive integer t, there exists directed graph $G = (V, E)$ such that $\mathsf{Val}(G) = 1$ and $\mathsf{Val}_t(G) \leqslant \frac{1}{2} + \eta$.*

Proof. Charikar et al (Section 4, [3]) construct a directed graph, $G = (V, E)$, on n vertices whose cut norm is bounded by $O(1/\log n)$. The graph is represented by the skew-symmetric matrix W, where $w_{ij} = \sum_k \sin \frac{\pi(j-i)k}{n+1}$. It is easy to verify that for every $0 < t < n$, $\sum_k \sin\left(\frac{\pi tk}{n+1}\right) \geqslant 0$. Thus, $w_{ij} \geqslant 0$ whenever $i < j$, implying that the graph is acyclic (in other words, $\mathsf{Val}(G) = 1$).

We bound $\mathsf{Val}_t(G)$ as follows. Let $\mathsf{Val}_t(G) = \frac{1}{2} + \delta$ and let $\mathcal{O} : V \to [t]$ be the optimal t-ordering. Construct a graph H on t vertices with a directed edge from $\mathcal{O}(u)$ to $\mathcal{O}(v)$ for every edge $(u, v) \in E$ with $\mathcal{O}(u) \neq \mathcal{O}(v)$. Now, using Theorem 5.2, the cut norm of H is bounded from below by $\Omega\left(\frac{\delta}{\log t}\right)$. Moreover, since $\mathcal{O}$ is a partition of V, the cut norm of G is at least the cut norm of H. Thus, we have the following:

$$\Omega\left(\frac{\delta}{\log t}\right) \leqslant ||H||_C \leqslant ||G||_C \leqslant O(1/\log n)$$

Thus, $\delta \leqslant O\left(\frac{\log t}{\log n}\right)$ implying that $\mathsf{Val}_t(G) \leqslant \frac{1}{2} + O\left(\frac{\log t}{\log n}\right)$. Choosing n large enough gives the required result. ☐

Corollary 5.4. *For every $\eta > 0$ and positive integer t, there exists a Multiscale Gap instance with a corresponding SDP solution $\{u_i | u \in V, 1 \leqslant i \leqslant |V|\}$.*

Proof. Let $G = (V, E)$ be the graph obtained by taking $\lceil 1/\eta \rceil$ disjoint copies of the graph guaranteed by Lemma 5.3 and let $m = |V|$. Note that the graph still satisfies the required properties: $\mathsf{Val}(G) = 1$, $\mathsf{Val}_t(G) \leqslant \frac{1}{2} + \eta$. Let $\mathcal{O}$ be the ordering of $[m]$ that satisfies every edge of G. Let D denote the distribution over labellings obtained by shifting $\mathcal{O}$ by a random offset cyclically. For every $u \in V, i \in [m]$, $\Pr[D(u) = i] = 1/m$. Further, every directed edge is satisfied with probability at least $1 - \eta$. Being a distribution over integral labellings, D gives raise to a set of vectors satisfying the constraints in Definition 5.1. G along with these vectors form the required (η, t)-multiscale gap instance. ☐

6 Dictatorship Test

Definition 6.1. For an edge $e = (u, v) \in E$ in a (η, t)-multiscale gap instance G, define the local integral distribution P_e over $[m]^2$ as follows:

$$\mathsf{P}_e(i, j) = u_i \cdot v_j$$

The details of the dictatorship test DICT_G are described below:

DICT_G Test

- Pick an edge $e = (u, v) \in E$ at random from the Multiscale gap instance G.
- Sample $\mathbf{z}_e = \{\mathbf{z}_u, \mathbf{z}_v\}$ from the product distribution P_e^R, i.e. For each $1 \leqslant k \leqslant R$, $z_e^{(k)} = \{z_u^{(k)}, z_v^{(k)}\}$ is sampled using the distribution $\mathsf{P}_e(i, j) = u_i \cdot v_j$.
- Obtain $\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v$ by perturbing each coordinate of $\mathbf{z}_u$ and $\mathbf{z}_v$ independently. Specifically, sample the k^{th} coordinates $\tilde{z}_u^{(k)}, \tilde{z}_v^{(k)}$ as follows: With probability $(1 - 2\epsilon)$, $\tilde{z}_u^{(k)} = z_u^{(k)}$, and with the remaining probability $\tilde{z}_u^{(k)}$ is a new sample from Ω.
- Introduce a directed edge $\tilde{\mathbf{z}}_u \to \tilde{\mathbf{z}}_v$. (alternatively test if $\mathcal{O}(\tilde{\mathbf{z}}_u) < \mathcal{O}(\tilde{\mathbf{z}}_v)$)

Theorem 6.2. *(Soundness Analysis) For any τ-pseudorandom ordering $\mathcal{O}$ of $[m]^R$, $\mathsf{Val}(\mathcal{O}) \leqslant \mathsf{Val}_t(G) + O(t^{-\frac{\epsilon}{2-\epsilon}}) + o_\tau(1)$.*

Let $\mathcal{F}^{[p,q]} : [m]^R \to \{0, 1\}$ 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 6.4 and Lemma 6.3 shown below.

Lemma 6.3. *For any τ-pseudorandom ordering $\mathcal{O}$ of $[m]^R$*

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

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

Proof. As $\mathcal{O}^*$ is a coarsening of $\mathcal{O}$, clearly $\mathsf{Val}(\mathcal{O}) \geq \mathsf{Val}_t(\mathcal{O}^*)$. Note that the loss due to coarsening, is because for some edges $e = (\mathbf{z}, \mathbf{z}')$ which are oriented correctly in $\mathcal{O}$, fall in to same block during coarsening, i.e $\mathcal{O}^*(z) = \mathcal{O}^*(z')$. Thus we can write

$$\mathsf{Val}(\mathcal{O}) \leqslant \mathsf{Val}_t(\mathcal{O}^*) + \frac{1}{2} \Pr\Big(\mathcal{O}^*(\tilde{\mathbf{z}}_u) = \mathcal{O}^*(\tilde{\mathbf{z}}_v)\Big)$$

$$\begin{aligned}\Pr\Big(\mathcal{O}^*(\tilde{\mathbf{z}}_u) = \mathcal{O}^*(\tilde{\mathbf{z}}_v)\Big) &= \sum_{i\in[t]} \mathbf{E}_{e=(u,v)} \mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v} \mathbf{E}_{\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v} \Big[\mathcal{F}^i(\tilde{\mathbf{z}}_u) \cdot \mathcal{F}^i(\tilde{\mathbf{z}}_v)\Big] \\ &= \sum_{i\in[t]} \mathbf{E}_{e=(u,v)} \mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v} \Big[T_{1-2\epsilon}\mathcal{F}_u^i(\mathbf{z}_u) \cdot T_{1-2\epsilon}\mathcal{F}_v^i(\mathbf{z}_v)\Big]\end{aligned}$$

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 $i \in [t]$, $\mathbf{E}_{\mathbf{z}}[\mathcal{F}_u^i(\mathbf{z}) = \frac{1}{t}]$. Further, since the ordering $\mathcal{O}^*$ is τ-pseudorandom, for every $k \in [R]$ and $i \in [t]$, $\mathrm{Inf}_k(T_{1-\epsilon}\mathcal{F}_a^i) \leqslant \tau$. Hence using Lemma 3.5, the above probability is bounded by $t \cdot t^{-\frac{2}{2-\epsilon}} + t \cdot o_\tau(1) = O(t^{-\frac{\epsilon}{2-\epsilon}}) + o_\tau(1)$. ∎

Lemma 6.4. *For any τ-pseudorandom t-ordering $\mathcal{O}^*$ of $[m]^R$, $\mathsf{Val}_t(\mathcal{O}^*) \leqslant \mathsf{Val}_t(G) + o_\tau(1)$.*

Proof. The t-ordering problem is a CSP over a finite domain, and is thus amenable to techniques of [19]. In fact, we shall use the soundness analysis of [19] to infer the result. The details of the proof are below.

Consider the payoff function $P : [t]^2 \to [0,1]$ defined by: $P(i,j) = 1$ for $i < j$, $P(i,j) = 0$ for $i > j$ and $P(i,j) = \frac{1}{2}$ otherwise. The t-ordering problem is a Generalized CSP(see Definition 3.1, [19]) with the payoff function P. Let DICT_G denote the dictatorship test obtained by running the reduction of [19] on the t-ordering instance G. DICT_G is a dictatorship test on functions $\mathcal{F} : [m]^R \to \Delta_t$, with the following soundness condition:

(Corollary 2.2, [19]) $\quad \mathsf{Soundness}_{\gamma,\tau}(\mathsf{DICT}_G) \leqslant \mathsf{Val}_t(G) + o_{\tau,\epsilon,\alpha,\gamma}(1)$

In other words, for any function $\mathcal{F} : [m]^R \to \Delta_t$ that is (γ, τ) pseudorandom, the expected payoff obtained in DICT_G is close to $\mathsf{Val}_t(G)$.

Consider the function $\mathcal{F} = (\mathcal{F}^1, \ldots, \mathcal{F}^t)$ over the domain $[m]^R$. For each point $\mathbf{z} \in [m]^R$, exactly one of these functions take value 1 while others are zero. Thus the range of $\mathcal{F}$ is the t-dimensional simplex Δ_t.

The dictatorship test described in this section is equivalent to DICT_G. Specifically, it produces a distribution of queries identical to DICT_G. In fact, $\mathsf{Val}_t(\mathcal{O}^*)$ is exactly equal to the expected payoff obtained by the function $\mathcal{F}$ on the dictatorship test DICT_G.

With the ordering $\mathcal{O}$ being τ-pseudorandom, for every $k \in [R]$ and $i \in [t]$, we have $\mathrm{Inf}_k(T_{1-\epsilon}\mathcal{F}^i) \leqslant \tau$. Consequently, the function $\mathcal{F} = (\mathcal{F}^1, \ldots, \mathcal{F}^t)$ is "(γ, τ)-pseudorandom" with $\gamma = 0$, as per Definition 4.1, and Definition 4.2 in [19]. This implies that the expected payoff of $\mathcal{F} = (\mathcal{F}^1, \ldots, \mathcal{F}^t)$ is at most $\mathsf{Val}_t(G)$

In [19], the corollary 2.2 is shown for a dictatorship test DICT_Φ for functions over domain $[m]^R$ taking values in the m-dimensional simplex Δ_m. However, the proof extends without any modifications for functions $\mathcal{F} : [m]^R \to \Delta_t$.

□

In terms of the functions $\mathcal{F}^i$, the expression for $\mathsf{Val}_t(\mathcal{O}^*)$ is as follows:

$$\begin{aligned}\mathsf{Val}_t(\mathcal{O}^*) &= \mathbf{E}_{e=(u,v)}\mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v}\mathbf{E}_{\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v}\left[\frac{1}{2}\sum_{i=j}\mathcal{F}^i(\tilde{\mathbf{z}}_u)\cdot\mathcal{F}^j(\tilde{\mathbf{z}}_v) + \sum_{i<j}\mathcal{F}^i(\tilde{\mathbf{z}}_u)\cdot\mathcal{F}^j(\tilde{\mathbf{z}}_v)\right] \\ &= \mathbf{E}_{e=(u,v)}\mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v}\mathbf{E}_{\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v}\left[\frac{1}{2}\sum_{i=j}\mathcal{F}^i(\tilde{\mathbf{z}}_u)\cdot\mathcal{F}^j(\tilde{\mathbf{z}}_v) + \sum_{i<j}\mathcal{F}^i(\tilde{\mathbf{z}}_u)\cdot\mathcal{F}^j(\tilde{\mathbf{z}}_v)\right]\end{aligned}$$

Here we restate the above lemma, in terms of the function $\mathcal{F}$.

Claim 6.5. *For a function $\mathcal{F} : [m]^R \to \Delta_t$ satisfying $\mathrm{Inf}_k(T_{1-\epsilon}\mathcal{F}) \leqslant \tau$ for all $k \in [R]$,*

$$\mathbf{E}_{e=(u,v)}\mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v}\mathbf{E}_{\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v}\left[\frac{1}{2}\sum_{i=j}\mathcal{F}^i(\tilde{\mathbf{z}}_u)\mathcal{F}^j(\tilde{\mathbf{z}}_u) + \sum_{i<j}\mathcal{F}^i(\tilde{\mathbf{z}}_u)\mathcal{F}^j(\tilde{\mathbf{z}}_u)\right] \leqslant \mathsf{Val}_t(G) + o_\tau(1)$$

7 Hardness Reduction

Let $G = (V, E)$ be a (η, t)-Multiscale gap instance, and let $m = |V|$. Further let $\{u_i | u \in V, i \in [m]\}$ denote the corresponding SDP solution. Let $\Gamma = (\mathcal{A} \cup \mathcal{B}, E, \Pi = \{\pi_e : [R] \to [R] | e \in E\}, [R])$ be a bipartite unique games instance. Towards constructing a MAS instance $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ from Γ, we shall introduce a long code for each vertex in $\mathcal{B}$. Specifically, the set of vertices $\mathcal{V}$ of the directed graph $\mathcal{G}$ is indexed by $\mathcal{B} \times [m]^R$.

Hardness Reduction
Input : Unique games instance $\Gamma = (\mathcal{A} \cup \mathcal{B}, E, \Pi = \{\pi_e : [R] \to [R] | e \in E\}, [R])$ and a (η, t) Multiscale gap instance $G = (V, E)$.
Output : Directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ 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 two neighbours $b, b' \in \mathcal{B}$ independently at random. Let π, π' denote the permutations on the edges (a, b) and (a, b').
- Pick an edge $e = (u, v) \in E$ at random from the Multiscale gap instance G.
- Sample $\mathbf{z}_e = \{\mathbf{z}_u, \mathbf{z}_v\}$ from the product distribution P_e^R, i.e. For each $1 \leqslant k \leqslant R$, $z_e^{(k)} = \{z_u^{(k)}, z_v^{(k)}\}$ is sampled using the distribution $\mathsf{P}_e(i, j) = u_i \cdot v_j$.
- Obtain $\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v$ by perturbing each coordinate of $\mathbf{z}_u$ and $\mathbf{z}_v$ independently. Specifically, sample the k^{th} coordinates $\tilde{z}_u^{(k)}, \tilde{z}_v^{(k)}$ as follows: With probability $(1 - 2\epsilon)$, $\tilde{z}_u^{(k)} = z_u^{(k)}$, and with the remaining probability $\tilde{z}_u^{(k)}$ is a new sample from Ω.
- Introduce a directed edge $(b, \pi(\tilde{\mathbf{z}}_u)) \to (b', \pi'(\tilde{\mathbf{z}}_v))$.

Theorem 7.1. *For every $\gamma > 0$, there exists choice of parameters $\epsilon, \eta, t, \delta$ such that:*

- Completeness: *If Γ 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. $\mathsf{Val}(\mathcal{G}) \geq 1 - \gamma$.*
- Soundness: *If Γ is not δ-satisfiable, then no ordering to $\mathcal{G}$ has value more than $\frac{1}{2} + \gamma$, i.e $\mathsf{Val}(\mathcal{G}) \leqslant \frac{1}{2} + \gamma$.*

In the rest of the section, we will present the proof of the above theorem. To begin with, we fix the parameters of the reduction.
Parameters : Fix $\epsilon = \gamma/8$ and $\eta = \gamma/4$. Let τ, t be the constants obtained from Theorem 7.5. Finally, let us choose $\delta = \min\{\gamma/4, \gamma\epsilon^2\tau^8/10^9\}$.

7.1 Completeness

In order to show that $\mathsf{Val}(\mathcal{G}) \geq 1 - \gamma$, we will instead show that $\mathsf{Val}_m(\mathcal{G}) \geq 1 - \gamma$. From Observation 3.1, this will imply the required result.

By assumption, there exists labelings to the Unique Game instance Γ such that for $1 - \delta$ fraction of the vertices $a \in \mathcal{A}$ all the edges (a, b) are satisfied. Let $\Lambda : \mathcal{X} \cup \mathcal{Y} \to [R]$ denote one such labelling. Define an m-ordering of $\mathcal{G}$ as follows:

$$\mathcal{O}(a, \mathbf{z}) = z^{(\Lambda(a))} \qquad \forall a \in \mathcal{A}, \mathbf{z} \in [m]^R$$

Clearly the mapping $\mathcal{O} : \mathcal{V} \to [m]$ defines an m-ordering of the vertices $\mathcal{V} = \mathcal{B} \times [m]^R$. To determine $\mathsf{Val}_m(\mathcal{O})$, let us compute the probability of acceptance of a verifier that follows the above procedure to generate an edge in $\mathcal{E}$ and then checks if the edge is satisfied. Arithmetizing this probability, we can write

$$\mathsf{Val}_m(\mathcal{O}) = \frac{1}{2}\Pr\Big(\mathcal{O}(b,\pi(\tilde{\mathbf{z}}_u)) = \mathcal{O}(b',\pi'(\tilde{\mathbf{z}}_v))\Big) + \Pr\Big(\mathcal{O}(b,\pi(\tilde{\mathbf{z}}_u)) < \mathcal{O}(b',\pi'(\tilde{\mathbf{z}}_v))\Big)$$

With probability at least $(1-\delta)$, the verifier picks a vertex $a \in \mathcal{A}$ such that the assignment Λ satisfies all the edges (a,b). In this case, for all choices of $b, b' \in N(a)$, $\pi(\Lambda(a)) = \Lambda(b)$ and $\pi'(\Lambda(a)) = \Lambda(b')$. Let us denote $\Lambda(a) = l$. By definition of the m-ordering $\mathcal{O}$, we get $\mathcal{O}(b,\pi(\mathbf{z})) = (\pi(\mathbf{z}))^{(\Lambda(b))} = z^{(\pi^{-1}(\Lambda(b)))} = z^{(l)}$ for all $\mathbf{z} \in [m]^R$. Similarly for b', $\mathcal{O}(b',\pi'(\mathbf{z})) = z^{(l)}$ for all $\mathbf{z} \in [m]^R$. Thus we get

$$\mathsf{Val}_m(\mathcal{O}) \geq (1-\delta)\cdot\Big(\frac{1}{2}\Pr\big(\tilde{z}_u^{(l)} = \tilde{z}_v^{(l)}\big) + \Pr\big(\tilde{z}_u^{(l)} < \tilde{z}_v^{(l)}\big)\Big)$$

With probability at least $(1-2\epsilon)^2$, for both $\tilde{\mathbf{z}}_u$ and $\tilde{\mathbf{z}}_v$ we have $\tilde{z}_u^{(l)} = z_u^{(l)}$ and $\tilde{z}_v^{(l)} = z_v^{(l)}$. Hence,

$$\mathsf{Val}_m(\mathcal{O}) \geq (1-\delta)(1-2\epsilon)^2\cdot\Big(\frac{1}{2}\Pr\big(z_u^{(l)} = z_v^{(l)}\big) + \Pr\big(z_u^{(l)} < z_v^{(l)}\big)\Big)$$

Note that each coordinate $z_u^{(l)}, z_v^{(l)}$ is generated according to the local distribution P_e for the edge $e = (u,v)$. For the local distribution P_e corresponding to an edge $e = (u,v) \in E$,

$$\Pr\Big(z_u^{(l)} = z_v^{(l)}\Big) = \sum_{i=j} u_i \cdot v_j \qquad\qquad \Pr\Big(z_u^{(l)} < z_v^{(l)}\Big) = \sum_{i<j} u_i \cdot v_j$$

Substituting in the expression for $\mathsf{Val}_m(\mathcal{O})$ we get,

$$\mathsf{Val}_m(\mathcal{O}) \geq (1-\delta)(1-2\epsilon)^2\mathbf{E}_{e=(u,v)}\Big[\frac{1}{2}\sum_{i=j} u_i \cdot v_j + \sum_{i<j} u_i \cdot v_j\Big]$$

Recall that the SDP vectors $\{u_i\}$ have an objective value at least $(1-\eta)$. Thus for small enough choice of δ, ϵ and η, we have $\mathsf{Val}_m(\mathcal{O}) \geq 1-\gamma$.

7.2 Soundness

Let $\mathcal{O}$ be an ordering of $\mathcal{G}$ with $\mathsf{Val}(\mathcal{O}) \geq \frac{1}{2}+\gamma$. Using the ordering, we will obtain a labelling Λ for the unique games instance Γ. Towards this, we shall build machinery to deal with multiple long codes. For $b \in \mathcal{B}$, define $\mathcal{O}_b$ as the restriction of the map $\mathcal{O}$ to vertices corresponding to the long code of b. Formally, $\mathcal{O}_b$ is a map $\mathcal{O}_b : [m]^R \to \mathbb{Z}$ given by $\mathcal{O}_b(\mathbf{z}) = \mathcal{O}(b,\mathbf{z})$. Similarly, for a vertex $a \in \mathcal{A}$, let $\mathcal{O}_a$ denote the restriction of the map $\mathcal{O}$ to the vertices $N(a)\times[m]^R$, i.e $\mathcal{O}_a(b,\mathbf{z}) = \mathcal{O}(b,\mathbf{z})$.

7.2.1 Multiple Long Codes

Throughout this section, we shall fix a vertex $a \in \mathcal{A}$ and analyze the long codes corresponding to all neighbours of a. For a neighbour $b \in N(a)$, we shall use π_b to denote the permutation along the edge (a,b). Let $\mathcal{F}_b^{[p,q]}$ denote the functions associated with the ordering $\mathcal{O}_b$. Define functions $\mathcal{F}_a^{[p,q]} : [m]^R \to \mathbb{R}$ as follows:

$$\mathcal{F}_a^{[p,q]}(\mathbf{z}) = \Pr_{b\in N(a)}\Big(\mathcal{O}_a(b,\pi_b(\mathbf{z})) \in [p,q]\Big) = \mathbf{E}_{b\in N(a)}[\mathcal{F}_b^{[p,q]}(\pi_b(\mathbf{z}))]$$

Definition 7.2. Define the set of influential coordinates $S_\tau(\mathcal{O}_a)$ as follows:

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

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

Lemma 7.3. *For any influential coordinate $k \in S_\tau(\mathcal{O}_a)$, for at least $\frac{\tau}{2}$ fraction of $b \in N(a)$, $\pi_b(k)$ is influential on $\mathcal{O}_b$. More precisely, $\pi_b(k) \in S_{\tau/2}(\mathcal{O}_b)$.*

Proof. As the coordinate k is influential on $\mathcal{O}_a$, there exists p, q such that $\mathrm{Inf}_k(\mathcal{F}_a^{[p,q]}) \geq \tau$. Recall that $\mathcal{F}_a^{[p,q]}(\mathbf{z}) = \mathbf{E}_{b \in N(a)}[\mathcal{F}_b^{[p,q]}(\pi_b(\mathbf{z}))]$. Using convexity of Inf this implies,

$$\mathbf{E}_{b \in N(a)}[\mathrm{Inf}_{\pi_b(k)}(\mathcal{F}_b^{[p,q]})] \geq \tau$$

All the influences $\mathrm{Inf}_{\pi_b(k)}(\mathcal{F}_b^{[p,q]})$ are bounded by 1, since each of the functions $\mathcal{F}_b^{[p,q]}$ take values in the range $[0, 1]$. Therefore for at least $\tau/2$ fraction of vertices $b \in N(a)$, we have $\mathrm{Inf}_{\pi_b(k)}(\mathcal{F}_b^{[p,q]}) \geq \tau/2$. This concludes the proof. ☐

Lemma 7.4. *For any vertex $a \in \mathcal{A}$, $|S_\tau(\mathcal{O}_a)| \leqslant 800/\epsilon\tau^4$.*

Proof. From Lemma 7.3, for each coordinate $k \in S_\tau(\mathcal{O}_a)$ there is a corresponding coordinate $\pi_b(k)$ in $S_{\tau/2}(\mathcal{O}_b)$ for at least $\tau/2$ fraction of the neighbours b. Further from Lemma 4.3, the size of each set $S_{\tau/2}(\mathcal{O}_b)$ is at most $400/\epsilon\tau^3$. By double counting, we get that $|S_\tau(\mathcal{O}_a)|$ is at most $800/\epsilon\tau^4$. ☐

Theorem 7.5. *For all $\epsilon, \gamma > 0$, there exists constants $t, \tau > 0$ such that for any vertex $a \in \mathcal{A}$, if $\mathcal{O}_a$ is τ-pseudorandom then $\mathsf{Val}(\mathcal{O}_a) \leqslant \mathsf{Val}_t(G) + \gamma/4$.*

Proof. The proof outline is similar to that of Theorem 6.2. Let $\mathcal{O}_a^*$ denote the t-coarsening of $\mathcal{O}_a$. Then we can write,

$$\mathsf{Val}(\mathcal{O}_a) \leqslant \mathsf{Val}_t(\mathcal{O}_a^*) + \frac{1}{2}\Pr\left(\mathcal{O}_a^*(b, \pi_b(\tilde{\mathbf{z}}_u)) = \mathcal{O}_a^*(b', \pi_{b'}(\tilde{\mathbf{z}}_v))\right)$$

The t-coarsening $\mathcal{O}_a^*$ is obtained by dividing the order $\mathcal{O}_a$ in to t-blocks. Let $[p_1 + 1, p_2], [p_2 + 1, p_3], \ldots, [p_t + 1, p_{t+1}]$ denote the t blocks. For the sake of brevity, let us denote $\mathcal{F}_a^i = \mathcal{F}_a^{[p_i+1,p_{i+1}]}$ and $\mathcal{F}_b^i = \mathcal{F}_b^{[p_i+1,p_{i+1}]}$. In this notation, we can write:

$$\begin{aligned}
\Pr\left(\mathcal{O}_a^*(b, \pi_b(\tilde{\mathbf{z}}_u)) = \mathcal{O}_a^*(b', \pi_{b'}(\tilde{\mathbf{z}}_v))\right) &= \sum_{i \in [t]} \mathbf{E}_{e=(u,v)}\mathbf{E}_{b,b'}\mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v}\mathbf{E}_{\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v}\left[\mathcal{F}_b^i(\pi_b(\tilde{\mathbf{z}}_u)) \cdot \mathcal{F}_{b'}^i(\pi_{b'}(\tilde{\mathbf{z}}_v))\right] \\
&= \sum_{i \in [t]} \mathbf{E}_{e=(u,v)}\mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v}\mathbf{E}_{\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v}\left[\mathcal{F}_a^i(\tilde{\mathbf{z}}_u) \cdot \mathcal{F}_a^i(\tilde{\mathbf{z}}_v)\right] \\
&= \sum_{i \in [t]} \mathbf{E}_{e=(u,v)}\mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v}\left[T_{1-2\epsilon}\mathcal{F}_a^i(\mathbf{z}_u) \cdot T_{1-2\epsilon}\mathcal{F}_a^i(\mathbf{z}_v)\right]
\end{aligned}$$

As the ordering $\mathcal{O}_a$ is τ-pseudorandom, for every $k \in [R]$ and $i \in [t]$, $\mathrm{Inf}_k(T_{1-\epsilon}\mathcal{F}_a^i) \leqslant \tau$. Hence by Lemma 3.5, the above value is less than $O(t^{-\frac{\epsilon}{2-\epsilon}}) + o_\tau(1)$.

Now we shall bound the value of $\mathsf{Val}_t(\mathcal{O}_a^*)$. In terms of the functions $\mathcal{F}_b^i$, the expression for $\mathsf{Val}_t(\mathcal{O}_a^*)$ is as follows:

$$\begin{aligned}\mathsf{Val}_t(\mathcal{O}_a^*) &= \mathbf{E}_{e=(u,v)}\mathbf{E}_{b,b'}\mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v}\mathbf{E}_{\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v}\Big[\frac{1}{2}\sum_{i=j}\mathcal{F}_b^i(\pi_b(\tilde{\mathbf{z}}_u))\cdot\mathcal{F}_{b'}^j(\pi_{b'}(\tilde{\mathbf{z}}_v))+\sum_{i<j}\mathcal{F}_b^i(\pi_b(\tilde{\mathbf{z}}_u))\cdot\mathcal{F}_{b'}^j(\pi_{b'}(\tilde{\mathbf{z}}_v))\Big]\\ &= \mathbf{E}_{e=(u,v)}\mathbf{E}_{\mathbf{z}_u,\mathbf{z}_v}\mathbf{E}_{\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v}\Big[\frac{1}{2}\sum_{i=j}\mathcal{F}_a^i(\tilde{\mathbf{z}}_u)\cdot\mathcal{F}_a^j(\tilde{\mathbf{z}}_v)+\sum_{i<j}\mathcal{F}_a^i(\tilde{\mathbf{z}}_u)\cdot\mathcal{F}_a^j(\tilde{\mathbf{z}}_v)\Big]\end{aligned}$$

Again, since the ordering $\mathcal{O}_a$ is τ-pseudorandom, for every $k \in [R]$ and $i \in [t]$, $\mathrm{Inf}_k(T_{1-\epsilon}\mathcal{F}_a^i) \leqslant \tau$. Hence by Claim 6.5, the above value is bounded by $\mathsf{Val}_t(G) + o_\tau(1)$. From the above inequalities, we get $\mathsf{Val}(\mathcal{O}_a) \leqslant \mathsf{Val}_t(G) + O(t^{-\frac{\epsilon}{2-\epsilon}}) + o_\tau(1)$, which finishes the proof. □

7.2.2 Defining a Labelling

Define the labelling Λ for the unique games instance Γ as follows: For each $a \in \mathcal{A}$, $\Lambda(a)$ is a uniformly random element from $S_\tau(\mathcal{O}_a)$ if it is non-empty, and a random label otherwise. Similarly for each $b \in \mathcal{B}$, assign $\Lambda(b)$ to be a random element of $S_{\tau/2}(\mathcal{O}_b)$ if it is nonempty, else an arbitrary label.

If $\mathsf{Val}(\mathcal{O})$ is greater than $\frac{1}{2} + \gamma$, then

$$\mathsf{Val}(\mathcal{O}) = \mathbf{E}_{a\in\mathcal{A}}[\mathsf{Val}(\mathcal{O}_a)] \geq \frac{1}{2} + \gamma$$

For at least $\gamma/2$ fraction of vertices $a \in \mathcal{A}$, we have $\mathsf{Val}(\mathcal{O}_a) \geq \frac{1}{2}+\gamma/2$. Let us refer to these vertices a as *good* vertices. From Theorem 7.5, for every *good* vertex the order $\mathcal{O}_a$ is not τ-pseudorandom. In other words, for every *good* vertex a, the set $S_\tau(\mathcal{O}_a)$ is non-empty. Further by Lemma 7.3 for every label $l \in S_\tau(\mathcal{O}_a)$, for at least $\tau/2$ fraction of the neighbours $b \in N(a)$, $\pi_b(l)$ belongs to $S_{\tau/2}(\mathcal{O}_b)$. For every such b, the edge (a, b) is satisfied with probability at least $1/|S_\tau(\mathcal{O}_a)| \times 1/|S_{\tau/2}(\mathcal{O}_b)|$. By Lemma 4.3 and Lemma 7.4, this probability is at least $\epsilon\tau^4/800 \times \epsilon\tau^3/3200$. Summarizing the argument, the expected fraction of edges satisfied by the labelling Λ is at least $\gamma\epsilon^2\tau^8/10240000$. By a small enough choice of δ, this yields the required result.

8 SDP Integrality Gap

In this section, we construct integrality gaps for the MAS-SDP relaxation using the unique games hardness reduction. Specifically we show,

Theorem 8.1. *For any $\gamma > 0$, there exists a directed graph G such that the value of semi definite program* (MAS-SDP) *is at least $1 - \gamma$, while $\mathsf{Val}(G) \leqslant \frac{1}{2} + \gamma$.*

The proof uses a bipartite variant of the Khot-Vishnoi [10] Unique Games integrality gap instance as in [19, 12]. Specifically, the following is a direct consequence of [10].

Theorem 8.2. *[10] For every $\delta > 0$, there exists a UG instance, $\Gamma = (\mathcal{A}\cup\mathcal{B}, E, \Pi = \{\pi_e : [R] \to [R] \mid e \in E\}, [R])$ and vectors $\{\mathbf{V}_b^k\}$ for every $b \in \mathcal{B}$, $k \in [R]$ such that the following conditions hold :*

- *No assignment satisfies more than δ fraction of constraints in Π.*
- *For all $b, b' \in \mathcal{B}, k, l \in [R]$, $\mathbf{V}_b^k \cdot \mathbf{V}_{b'}^l \geqslant 0$ and $\mathbf{V}_b^k \cdot \mathbf{V}_b^l = 0$.*
- *For all $b, b' \in \mathcal{B}, k, l \in [R]$, $\mathbf{V}_b^k \cdot \sum_{l \in [R]} \mathbf{V}_{b'}^l = |\mathbf{V}_b^k|^2$ and $\sum_{k \in [R]} |\mathbf{V}_b^k|^2 = 1$*
- *The SDP value is at least $1 - \delta$: $E_{a \in \mathcal{A}, b, b' \in \mathcal{B}} \Big[\sum_{k \in [R]} \mathbf{V}_b^{\pi(k)} \cdot \mathbf{V}_{b'}^{\pi'(k)} \Big] \geqslant 1 - \delta$*

Let G be a (η, t)-multiscale gap instance with m vertices. Apply Theorem 8.2, with a sufficiently small δ to obtain a UGC instance Γ and SDP vectors $\{\mathbf{V}_b^k | b \in \mathcal{B}, k \in [R]\} \cup \{\mathbf{I}\}$. Consider the instance $\mathcal{G}$ constructed by running the UG hardness reduction in Section 7 on the UG instance Γ. The set of vertices of $\mathcal{G}$ is given by $\mathcal{B} \times [m]^R$. Set $M = |\mathcal{B}| \times m^R$ and $N = |\mathcal{B}|$. Further, fix an arbitrary ordering $\{b_1, \ldots, b_N\}$ of the vertices in $\mathcal{B}$.

The program MAS-SDP on the instance $\mathcal{G}$ contains M vectors $\{\mathbf{W}_i^{(b_j, \mathbf{z})} | i \in [M]\}$ for each vertex $(b_j, \mathbf{z}) \in \mathcal{B} \times [q]^R$ and a special vector $\mathbf{I}$ denoting the constant 1. Define a solution to MAS-SDP as follows: Set the vector $\mathbf{I}$ to be the corresponding vector in the instance Γ. For each $(b_j, \mathbf{z}) \in \mathcal{B} \times [m]^R$ and $i \in [R]$ define

$$
\begin{aligned}
\mathbf{W}_{N(i-1)+j}^{(b_j, \mathbf{z})} &= \sum_{z_k = i} \mathbf{V}_{b_j}^k & & \forall i \in [R], (b_j, \mathbf{z}) \in \mathcal{B} \\
\mathbf{W}_l^{(b_j, \mathbf{z})} &= 0 & & \text{for any other choice of } l \in [M], (b_j, \mathbf{z}) \in \mathcal{B}
\end{aligned}
$$

It is easy to check that the vectors $\{\mathbf{W}_i^{(b_j, \mathbf{z})}\}$ satisfy the constraints of MAS-SDP and have an SDP value close to 1. On the other hand, the soundness analysis in Section 7 implies that the integral optimum for $\mathcal{G}$ is at most $\frac{1}{2} + \gamma$. The details of the proof will appear in the full version of the paper.

9 Acknowledgments

We would like to thank Alantha Newman for pointing out the similarity between Maximum Acyclic Subgraph and CSPs, and suggesting the problem. We would like to thank Sanjeev Arora for many fruitful discussions.

References

[1] *46th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2005), 23-25 October 2005, Pittsburgh, PA, USA, Proceedings*. IEEE Computer Society, 2005.

[2] Bonnie Berger and Peter W. Shor. Tight bounds for the Maximum Acyclic Subgraph problem. *J. Algorithms*, 25(1):1–18, 1997.

[3] Moses Charikar, Konstantin Makarychev, and Yury Makarychev. On the advantage over random for maximum acyclic subgraph. In *FOCS*, pages 625–633. IEEE Computer Society, 2007.

[4] Guy Even, Joseph Naor, Baruch Schieber, and Madhu Sudan. Approximating minimum feedback sets and multicuts in directed graphs. *Algorithmica*, 20(2):151–174, 1998.

[5] Johan Håstad. Every 2-csp allows nontrivial approximation. In Harold N. Gabow and Ronald Fagin, editors, *STOC*, pages 740–746. ACM, 2005.

[6] A. Karazanov. On the minimal number of arcs of a digraph meeting all its directed cutsets. *Graph Theory Newsletters*, 8, 1979.

[7] Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of Computer Computations. New York: Plenum*, pages 85–103, 1972.

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

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

[10] Subhash Khot and Nisheeth K. Vishnoi. The unique games conjecture, integrality gap for cut problems and embeddability of negative type metrics into l_1. In *FOCS* [1], pages 53–62.

[11] Cláudio Lucchesi and Daniel H. Younger. A minimax theorem for directed graphs. *Journal London Math. Society*, 17:369–374, 1978.

[12] Rajsekar Manokaran, Joseph (Seffi) Naor, Prasad Raghavendra, and Roy Schwartz. Sdp gaps and ugc hardness for multiway cut, 0-extension and metric labelling. *In STOC '08: Proceedings of the 40th ACM Symposium on Theory of Computing*, 2008.

[13] Elchanan Mossel. Gaussian bounds for noise correlation of functions (arxiv: math/0703683) [http://front.math.ucdavis.edu/0703.5683]. To appear in FOCS, 2008.

[14] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. In *FOCS* [1], pages 21–30.

[15] Alantha Newman. Approximating the maximum acyclic subgraph. Master's thesis, MIT, June 2000.

[16] Alantha Newman. *Algorithms for String and Graph Layout.* PhD thesis, MIT, August 2004.

[17] Alantha Newman. Cuts and orderings: On semidefinite relaxations for the linear ordering problem. In Klaus Jansen, Sanjeev Khanna, José D. P. Rolim, and Dana Ron, editors, *APPROX-RANDOM*, volume 3122 of *Lecture Notes in Computer Science*, pages 195–206. Springer, 2004.

[18] Christos H. Papadimitriou and Mihalis Yannakakis. Optimization, approximation, and complexity classes. *J. Comput. Syst. Sci.*, 43(3):425–440, 1991.

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

[20] V. Ramachandran. Finding a minimum feedback arc set in reducible flow graphs. *J. Algorithms*, 9:299–313, 1988.

[21] Paul Seymour. Packing directed circuits fractionally. *Combinatorica*, 15(2):281–288, 1995.