SIAM J. COMPUT.
Vol. 40, No. 3, pp. 878–914

BEATING THE RANDOM ORDERING IS HARD: EVERY ORDERING CSP IS APPROXIMATION RESISTANT*

VENKATESAN GURUSWAMI†, JOHAN HÅSTAD‡, RAJSEKAR MANOKARAN§, PRASAD RAGHAVENDRA¶, AND MOSES CHARIKAR§

Abstract. We prove that, assuming the Unique Games conjecture (UGC), every problem in the class of ordering constraint satisfaction problems (OCSPs) where each constraint has constant arity is approximation resistant. In other words, we show that if ρ is the expected fraction of constraints satisfied by a random ordering, then obtaining a ρ' approximation for any $\rho' > \rho$ is UG-hard. For the simplest OCSP, the MAXIMUM ACYCLIC SUBGRAPH (MAS) problem, this implies that obtaining a ρ-approximation for any constant $\rho > 1/2$ is UG-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, it is UG-hard to find one with more than $(1/2 + \varepsilon)$ of its edges. 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 MAS problem has been well studied, and beating the random ordering for MAS has been a basic open problem. An OCSP of arity k is specified by a subset $\Pi \subseteq S_k$ of permutations on $\{1, 2, \ldots, k\}$. An instance of such an OCSP is a set V and a collection of constraints, each of which is an ordered k-tuple of V. The objective is to find a global linear ordering of V while maximizing the number of constraints ordered as in Π. A random ordering of V is expected to satisfy a $\rho = \frac{|\Pi|}{k!}$ fraction. We show that, for any fixed k, it is hard to obtain a ρ'-approximation for Π-OCSP for any $\rho' > \rho$. The result is in fact stronger: we show that for every $\Lambda \subseteq \Pi \subseteq S_k$, and an arbitrarily small ε, it is hard to distinguish instances where a $(1-\varepsilon)$ fraction of the constraints can be ordered according to Λ from instances where at most a $(\rho + \varepsilon)$ fraction can be ordered as in Π. A special case of our result is that the BETWEENNESS problem is hard to approximate beyond a factor 1/3. The results naturally generalize to OCSPs which assign a payoff to the different permutations. Finally, our results imply (unconditionally) that a simple semidefinite relaxation for MAS does not suffice to obtain a better approximation.

Key words. MAXIMUM ACYCLIC SUBGRAPH, feedback arc set, Unique Games conjecture, hardness of approximation, integrality gaps

AMS subject classification. 68Q17

DOI. 10.1137/090756144

1. Introduction.

We begin by discussing our results about the simplest ordering constraint satisfaction problem—MAXIMUM ACYCLIC SUBGRAPH (MAS)—that involves local ordering constraints on pairs of variables.

*Received by the editors April 16, 2009; accepted for publication (in revised form) January 17, 2011; published electronically June 23, 2011. Preliminary versions of this paper appeared in *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, 2008, pp. 573–582 and *Proceedings of the 24th IEEE Conference on Computational Complexity*, 2009, pp. 62–73.
http://www.siam.org/journals/sicomp/40-3/75614.html

†Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213 (guruswami@cmu.edu). Some of this work was done while this author was visiting the School of Mathematics, Institute for Advanced Study, Princeton, NJ. This author's research was supported in part by a Packard Fellowship, and by NSF grants CCF-0343672 and CCF-0963975.

‡School of Computer Science and Communication, KTH, S-100 44 Stockholm, Sweden (johanh@kth.se). This author's research was supported by ERC grant 226203.

§Department of Computer Science, Princeton University, Princeton, NJ 08540 (rajsekar@cs.princeton.edu, moses@cs.princeton.edu). The research of these authors was supported by NSF grants MSPA-MCS 0528414 and ITR 0205594.

¶College of Computing, Georgia Institute of Technology, Atlanta, GA 30332 (praghave@cc.gatech.edu). Some of this work was done while this author was visiting Princeton University and when at Microsoft Research New England, Cambridge, MA. This author's research was supported in part by NSF grant CCF-0343672.

1.1. MAS. 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 a fraction ε, 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., a fraction $(1 - \varepsilon)$, 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 MAS problem.

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 MAS. (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 MAS. The existence of such an algorithm has been a long-standing 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 (UGC). Formally, we show the following.

THEOREM 1.1. *Conditioned on the UGC, the following holds for every constant $\gamma > 0$. Given a weighted 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 a fraction $(1-\gamma)$ of the edges (in weight) directed forward (or, equivalently, G has an acyclic subgraph with at least a fraction $(1 - \gamma)$ of the weight).*
2. *For every ordering of the vertices of G, there are at most a fraction $(1/2+\gamma)$ of forward edges in weight (or, equivalently, every subgraph of G with more than a fraction $(1/2 + \gamma)$ of the weights 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 MIN FEEDBACK ARC SET (FAS) problem, where the objective is to minimize the number of back edges.

COROLLARY 1.2. *Conditioned on the UGC, for every $C > 0$, it is NP-hard to find a C-approximation to the* FAS *problem.*

Combining the unique game integrality gap instance of Khot and Vishnoi [26] along with the UG reduction, we obtain semidefinite programming (SDP) integrality gaps for the MAS problem. Our integrality gap instances also apply to a related SDP relaxation studied by Newman [33]. 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 [32].

1.2. General ordering constraints. Building on these techniques and the work of Raghavendra [35], we obtain tight UGC-based hardness results for the entire class of ordering constraint satisfaction problems (OCSPs).

An OCSP Λ of arity k is specified by a constraint payoff function $P : S_k \to [0, 1]$, where S_k is the set of permutations of $\{1, 2, \ldots, k\}$. An instance of such an OCSP consists of a set of variables V and a collection of constraint tuples, $\mathcal{T}$, each of which is an ordered k-tuple of V. The objective is to find a global ordering σ of V that maximizes the expected payoff $\mathbb{E}[P(\sigma_{|T})]$ for a random $T \in \mathcal{T}$, where $\sigma_{|T} \in S_k$ is the ordering of the k elements of T induced by the global ordering σ. This is just the natural extension of CSPs to the world of ordering problems. For generality, we allow payoff functions with range $[0, 1]$ instead of $\{0, 1\}$ which would correspond to True/False constraints. Without loss of generality, by reordering the inputs of any constraint, we may assume that the permutation σ which maximizes $P(\sigma)$ is the identity, id.

As with CSPs, we say that an OCSP of arity k and a payoff function P is *approximation resistant* if its approximation threshold equals

$$\frac{\mathbb{E}_{\alpha \in S_k}[P(\alpha)]}{P(\mathrm{id})},$$

which is the ratio that can be obtained by choosing a random ordering.

Note that in this language, MAS corresponds to the simplest OCSP: the arity 2 OCSP with a payoff function that gives value 1 to the identity permutation and 0 to its reverse.

Our main result is that *every* OCSP, of arity bounded by a fixed k, is approximation resistant. Specifically, for every such OCSP, outperforming the trivial approximation ratio achieved by random ordering is UG-hard.

THEOREM 1.3 (main). *Let k be a positive integer, and let Λ be an* OCSP *associated with a payoff function $P : S_k \to [0,1]$ on the set of k-permutations, S_k. Let $\Lambda_{\max} = \max_{\alpha \in S_k} P(\alpha)$ be the maximum payoff of P, and let $\Lambda_{random} = \mathbb{E}_{\alpha \in S_k} P(\alpha)$ be the average payoff of P (expected value achieved by a uniform random ordering).*

Then, for every $\varepsilon > 0$, the following hardness result holds. Given an instance of the OCSP specified by the payoff function P that admits an ordering with a payoff at least $\Lambda_{\max} - \varepsilon$, it is UG-hard to find an ordering of the instance that achieves a payoff of at least $\Lambda_{random} + \varepsilon$ with respect to the payoff function P.

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.

Indeed, our result holds in a more general setting where the OCSP could consist of a mixture of predicates—a formal statement appears in section 8 (Theorem 8.4).

1.3. Related work. MAS is a classic optimization problem, figuring in Karp's early list of NP-hard problems [22]; 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. MAS is also complete for the class of permutation optimization problems, $\mathsf{MAX\ SNP}[\pi]$, defined in [34], that can be approximated within a constant factor. It is shown in [32] that MAS 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 [27, 21] and reducible flow graphs [36]. Berger and Shor [5] 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 [32] 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 directed acyclic graph (DAG), leads to the FAS problem. This problem admits a factor $O(\log n \log \log n)$ approximation algorithm [37], where n is the number of vertices, based on bounding the integrality gap of the natural covering linear program for FAS; see also [11]. Using this algorithm, one can get an approximation ratio of $\frac{1}{2} + \Omega(1/(\log n \log \log n))$ for MAS.

Charikar, Makarychev, and Makarychev [7] gave a factor $(1/2 + \Omega(1/\log n))$-approximation algorithm for MAS. 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 specifically an instance showing tightness of its analysis from [7], is used as the combinatorial gadget for our hardness result for MAS.

Apart from MAS, another OCSP that has received some attention is the BETWEENNESS problem. BETWEENNESS is an OCSP where all the constraints are of the form "X appears between Y and Z" for variables X, Y, and Z. Chor and Sudan [9] gave a SDP-based factor $\frac{1}{2}$ approximation algorithm for BETWEENNESS on instances that were promised to be perfectly satisfiable; a simpler algorithm with the same guarantee was given by Makarychev [28]. Recently, Guruswami and Zhou [15] proved that the extension of MAS to higher arities, with constraints of the form $x_{i_1} < x_{i_2} < \cdots < x_{i_k}$, can be approximated within a factor greater than $1/k!$ on bounded-degree instances. They extend this to prove that all OCSPs of arity 3 (with arbitrary payoff functions) can be approximated beyond their random ordering threshold on bounded-degree instances.

1.3.1. Approximation resistance. Our main result is that every OCSP is approximation resistant under the UGC. In contrast, in the world of CSPs over fixed domains (such as Boolean CSPs), there are CSPs which are approximable beyond the random assignment threshold. There is by now a rich body of work on approximability of CSPs, though we are quite far from a complete classification of which CSPs are approximation resistant and which ones admit a nontrivial approximation algorithm that beats the trivial random assignment algorithm. But we now know fairly broad classes of CSPs which are approximation resistant, as well as those that are not. We mention some of these results below.

Håstad [17] proved many important CSPs to be approximation resistant, including Max 3SAT, Max 3LIN (whose predicate stipulates that the parity of 3 literals is 0), and in fact any binary 3CSP whose predicate is implied by the parity constraint $x \oplus y \oplus z = 0$, Max k-set splitting for $k \geqslant 4$, etc. Complementing Håstad's hardness result for 3CSPs, Zwick [38] 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 Boolean 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 [16].

It is known that every 2CSP, even over nonbinary domains, can be approximated better than the random assignment threshold [13, 10, 18]. The approximation threshold of 2CSPs (such as Max Cut) remained a fascinating mystery until recent progress based on the UGC tied it to the integrality gap of SDP relaxations [24, 2, 35]. In fact, under the UGC, Raghavendra showed the general result [35] that for every CSP, the approximation threshold equals the integrality gap of a natural SDP relaxation. Unfortunately, determining this integrality gap itself is often an extremely challenging task, so this does not immediately tell us which CSPs are approximation resistant (even assuming the UGC).

An elegant result of Austrin and Mossel [4] states that under the UGC any CSP whose satisfying assignments can support a pairwise independent distribution is approximation resistant. Using this, Austrin and Håstad [3] (see also [19]) showed that *most* k-ary predicates (a fraction approaching 1 for large k) are approximation resistant under the UGC.

Our main contribution in this work is to extend the above-mentioned result of Raghavendra [35] to OCSPs. Executing this plan requires several new ideas which we elaborate on in section 2. Roughly stated, we prove that for OCSPs, the existence of a certain kind of "weak" SDP integrality gap implies a corresponding UG-hardness. We are then able to construct instances whose integrality gap is close to the random ordering threshold. Together, these two results imply that all OCSPs are approximation resistant, assuming the UGC.

1.4. 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 and noise operators and restate the UGC. The groundwork for our reduction is laid in sections 4 and 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. Using this UG-hardness result we later, in section 12, establish present SDP integrality gaps for MAS.

Towards generalizing these hardness results, we begin with a formal definition of OCSPs and the natural semidefinite program for OCSPs in section 8. The construction of dictatorship tests for an OCSP starting from an object termed as a *multiscale gap instance* is presented in section 9. An important part of the soundness analysis is done in section 10 and is based on the ideas of [35]. Finally, in section 11, we exhibit the needed explicit construction of multiscale gap instances for every OCSP.

2. Proof overview. At the heart of all UG-hardness results lies a dictatorship testing result for an appropriate class of functions. As is standard we use $[m]$ to denote $\{1, \ldots, m\}$. 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 that, given a function $\mathcal{F} : [m]^R \to [m]$, makes a few queries to the values of $\mathcal{F}$ and distinguishes between whether $\mathcal{F}$ is a dictator or is *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 for which one is showing UG-hardness determines the nature of the dictatorship test needed for the purpose.

A dictatorship test (also referred to as long code test) serves as a gadget to be used in the reduction from UG. In UG, the input consists of a graph whose vertices are to be labeled, so as to satisfy the maximum number of constraints given on the edges. Given a UG instance Φ, a standard reduction technique is to introduce a dictatorship test gadget for each vertex in the instance Φ. We refer the reader to the work of Khot et al. [24] for an example of a long-code-based UG-hardness reduction.

Every ordering $\mathcal{O}$ of $[m]^R$ can be viewed as a function from $[m]^R$ to $\{1, 2, \ldots, m^R\}$. For the purpose of defining influence of orderings, we define m^{2R} functions $\mathcal{F}^{[s,t]} : [m]^R \to \{0, 1\}$ as follows:

$$\mathcal{F}^{[s,t]}(x) = \begin{cases} 1 & \text{if } s \leqslant \mathcal{O}(x) \leqslant t, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Given an ordering $\mathcal{O} : [m]^R \to \{1, \ldots, m^R\}$ of $[m]^R$, the ith coordinate of the input is said to be *influential* on $\mathcal{O}$ if it has a large influence ($> \tau$) on any of the functions $\mathcal{F}^{[s,t]}$. Here influence of a coordinate on a function $\mathcal{F}^{[s,t]}$ refers to the traditional notion of influence for real-valued functions on $[m]^R$. Roughly speaking, the influence of the ith coordinate is the expected variance of the output of the function $\mathcal{F}^{[s,t]}$ on fixing all but the ith coordinate randomly and varying the ith coordinate (see section 3). An ordering $\mathcal{O}$ is said to be τ-pseudorandom (*far* from a dictator) if it has no coordinate of influence at least τ.

For the sake of concreteness, let us consider the UG-hardness reduction to MAS. In this case, we introduce m^R vertices $\{(b,\mathbf{z}) \mid \mathbf{z} \in [m]^R\}$ for each vertex b of the UG instance Φ. Let $\mathcal{O}$ be an ordering of all the vertices of the resulting instance of MAS. Let $\mathcal{O}_b$ denote the induced ordering on the block of vertices $\{(b,\mathbf{z}) \mid \mathbf{z} \in [m]^R\}$ corresponding to a UG vertex b. The intent is to use $\mathcal{O}_b$ to decode a label for the UG vertex b.

Usually, in a long-code-based UG-hardness reduction, a small candidate set of labels decoded for a vertex b is given by the set of influential coordinates for the function corresponding to b. Hence, for the notion of influences for orderings to be useful, it is necessary that any given ordering $\mathcal{O}_b$ of $[m]^R$ 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.

2.1. MAS. Let us describe the proof strategy for the UG-hardness of MAS. 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 in $\mathcal{O}$ correctly.

Designing the appropriate dictatorship test for MAS amounts to the following: Construct a directed graph over the set of vertices $V = [m]^R$ (for some large constants m, R) such that the following hold:

- For a *dictator* ordering $\mathcal{O}$ of V, which is defined by using one of the coordinates of each vertex to give the ordering, $\mathsf{Val}(\mathcal{O}) \approx 1$.
- For any ordering $\mathcal{O}$ which is *far from a dictator*, $\mathsf{Val}(\mathcal{O}) \approx \frac{1}{2}$.

Recall that our definition of influential coordinates for orderings can be used to formalize the notion of being "far from dictator functions." Under this definition, 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 MAS problem. Furthermore, using the SDP gap instance for UG from the work of Khot and Vishnoi [26], the hardness reduction yields an integrality gap instance for a natural SDP relaxation (see subsection 3.2) of MAS.

Now we describe the design of the dictatorship test in greater detail. At the outset, the approach is similar to recent work on CSPs [35]. Fix a CSP Λ. Starting with an integrality gap instance $\Im$ for the natural semidefinite program for Λ, [35] constructs a dictatorship test $\mathsf{DICT}_\Im$. The $\mathsf{Completeness}$ of $\mathsf{DICT}_\Im$ is equal to the SDP value $\mathrm{sdp}(\Im)$, while the $\mathsf{Soundness}$ is close to the integral value $\mathrm{opt}(\Im)$.

Since the result of [35] applies to arbitrary CSPs, a natural direction would be to pose the MAS as a CSP. MAS is fairly similar to a CSP, with each vertex being a variable taking values in domain $[n]$ and each directed edge a constraint between two 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, if MAS was reducible to a 2CSP over a domain of fixed size, then we could obtain an approximation ratio better than a random assignment [18].

Towards using techniques from the CSP result, we define the following variant of MAS.

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

$$\text{val}_q(\mathcal{O}) = \mathop{\mathbf{Pr}}_{(u,v)\in E}\Big(\mathcal{O}(u) < \mathcal{O}(v)\Big) + \frac{1}{2}\mathop{\mathbf{Pr}}_{(u,v)\in E}\Big(\mathcal{O}(u) = \mathcal{O}(v)\Big).$$

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

The choice to give half credit for edges where the two endpoints are mapped to the same value is motivated by two similar reasons. The first reason is that the constraint is neither violated nor fulfilled, and the second is that the constraint is satisfied with probability $\frac{1}{2}$ if we choose a random, full ordering that respects the partial ordering defined by the given q-ordering.

On the one hand, the q-Order problem is a CSP over a fixed domain that is similar to MAS. However, to the best of our knowledge, for the q-Order problem, there are no known SDP gaps, which constitute the starting point for the results in [35]. For any fixed constant q, Charikar, Makarychev, and Makarychev [7] construct DAGs G (i.e., with the value of the best ordering equal to 1) such that the value of any q-ordering of G is close to $\frac{1}{2}$, say, at most $\frac{1}{2} + \eta$. We call such a graph an (η, q)-pseudorandom DAG. 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 an integrality gap example for the natural SDP relaxation of either the MAS problem or the q-Order problem.

As the graph G has only m vertices and an ordering of value ≈ 1, it has a good q-ordering for $q = m$. Viewing G as an instance of the m-Order CSP (corresponding to predicate $<$ and $=$), we obtain a directed graph, $\mathcal{G}$, on $[m]^R$. Loosely speaking, $\mathcal{G}$ is similar to a direct product of R copies of G, and hence the given good m-ordering of G ensures that the dictator m-orderings $\mathcal{O} : [m]^R \to [m]$ given by $\mathcal{O}(\mathbf{z}) = z_i$ for some $i \in [R]$ yield value ≈ 1 on $\mathcal{G}$. In other words, the dictator orderings have value ≈ 1 on $\mathcal{G}$, implying the completeness of the dictatorship test.

Now let us turn to the soundness analysis. Fix a τ-pseudorandom ordering $\mathcal{O}$. Obtain a q-ordering $\mathcal{O}^*$ by the following *coarsening* process: Divide the ordering $\mathcal{O}$ into q equal blocks, and map the vertices in the ith block to value i. The crucial observation relating $\mathcal{O}$ and $\mathcal{O}^*$, which relies on the fact that we have some noise in the construction, is as follows (proved in Lemma 6.3):

> *Coarsening observation.* For a τ-pseudorandom ordering $\mathcal{O}$, $\text{val}_q(\mathcal{O}^*) \approx \text{val}(\mathcal{O})$.

Note that $\text{val}(\mathcal{O}) - \text{val}_q(\mathcal{O}^*)$ is clearly bounded by the fraction of edges whose endpoints both fall in the same block during the coarsening. Using the Gaussian noise stability bounds of [30], we obtain a bound for the fraction of such edges, thereby proving the above observation. From the above observation, in order to prove $\text{val}(\mathcal{O}) \approx \frac{1}{2}$ for a τ-pseudorandom ordering $\mathcal{O}$, it is enough to bound $\text{val}_q(\mathcal{O}^*)$. Recall that the q-order problem is a CSP over a finite domain. Consequently, the soundness analysis of Raghavendra [35] can be used to show that $\text{val}_q(\mathcal{O}^*)$ is at most the value of the best q-ordering for the original graph 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 based on a certain gap instance for MAS. Using Gaussian noise stability bounds, we relate the value of a pseudorandom ordering to a related CSP and then apply techniques from [35]. Instantiating the gap instance with the (η, q)-pseudorandom DAG G finishes the proof.

2.2. OCSPs. The techniques developed in the case of MAS, along with ideas from [35], yield an approach to proving UG-hardness results for general OCSPs. In a general OCSP, a set of local ordering constraints such as "i is before j" or "i is between j and k" is given, and the goal is to find an ordering that satisfies the maximum number of constraints (see section 8 for a formal definition).

First, as in the case of MAS, for every OCSP Λ, it is possible to define a related CSP Λ_q over the domain $[q]$ for every positive integer q. Roughly speaking, the CSP Λ_q consists of the problem of finding the q-Order that achieves the maximum payoff. Given a q-Order $\mathcal{O}$ of an instance $\Im$ of Λ-OCSP, we use $\mathrm{val}_q(\mathcal{O})$ to denote its objective value (fraction of constraints satisfied). Further, let $\mathrm{opt}_q(\Im)$ denote the optimum value of a q-Order for the instance $\Im$.

In case of CSPs, the work of Raghavendra [35] established a black-box reduction from an integrality gap instance for a certain canonical SDP relaxation to a matching UG-hardness result. However, constructing integrality gap instances for OCSPs is in itself a challenging task. In this light, for every OCSP, we exhibit a black-box reduction to a UG-hardness result starting from what we refer to as a *multiscale gap* instance—a weaker object than an SDP integrality gap. Formally, a multiscale gap is defined as follows.

DEFINITION 2.3. *An instance $\Im$ of a Λ-OCSP is a (q, c, s)-multiscale gap instance if* $\mathrm{sdp}(\Im) \geqslant c$ *and* $\mathrm{opt}_q(\Im) \leqslant s$. *Here the SDP value refers to the optimum of a canonical SDP relaxation, described in section* 8.3.

It is not difficult to see that an integrality gap instance $\Im$ with $\mathrm{sdp}(\Im) = c$ and $\mathrm{opt}(\Im) = s$ (as opposed to $\mathrm{opt}_q(\Im) = s$) is a (q, c, s)-multiscale gap instance for all q (see Claim 8.6). Hence, a multiscale gap instance is formally easier to construct than an integrality gap instance. We give a reduction that obtains a UG-hardness result for an OCSP Λ starting with a multiscale gap instance for it. Specifically, we prove the following.

THEOREM 2.4. *If there exists a (q, c, s)-multiscale gap instance $\Im$ for an OCSP Λ, then, for every $\eta > 0$, it is UG-hard to distinguish Λ-OCSP instances with optimum at least $c - \eta$ from instances with optimum at most $s + \eta + O(q^{-\eta})$.*

To show Theorem 2.4, we give a black-box reduction that converts the instance $\Im$ with SDP solution $(\boldsymbol{V}, \boldsymbol{\mu})$ into a dictatorship test $\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}$ with completeness $c-\eta$ and soundness at most $s + \eta + O(q^{-\eta})$. Further, all the predicates checked by the dictatorship test $\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}$ belong to the family of predicates corresponding to the OCSP Λ.

Let m denote the number of variables in the instance $\Im$. The dictatorship test $\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}$ is constructed by viewing the instance $\Im$ as a CSP over a domain of size m. Specifically, $\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}$ is an instance of a Λ-OCSP over the set of variables indexed by $[m]^R$ for an integer R. The m-orderings of $[m]^R$ given by the dictator functions have an objective value close to the SDP value ($c-\eta$ in this case; the η loss is due to some noise added by the dictatorship test). To perform the soundness analysis, we appeal to the coarsening observation above. By using this observation, we can relate the value of an ordering $\mathcal{O}$ of $\Im$ to the value of the q-Order $\mathcal{O}_q$ obtained by coarsening $\mathcal{O}$. Finally, using a proof strategy along the lines of [35], we relate the value $\mathrm{val}_q(\mathcal{O}_q)$ of the q-Order $\mathcal{O}_q$ of $[m]^R$ to the optimum q-Order value $\mathrm{opt}_q(\Im)$ of the instance $\Im$.

Starting from the dictatorship test $\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}$, the UG-hardness result for OCSP Λ can be obtained exactly along the lines of MAS. Therefore, we omit the proof of the UG-hardness result from this presentation.

In section 11, we exhibit an explicit construction of multiscale gap instances for every OCSP, which, when plugged into Theorem 2.4, give our main result on the approximation resistance of all OCSPs under the UGC.

THEOREM 2.5. *For all positive integers q, k, for all $\eta > 0$, and for every OCSP Λ of arity k, there exists a $(q, \Lambda_{\max}, \Lambda_{random} + \eta)$-multiscale gap instance $\Im$ of Λ.*

The core of the above gap instance is our construction of a distribution D on $[m]^k$ with the following properties (here k, q are positive integers, $\eta > 0$ can be arbitrarily small, and m is a large enough integer):

- Completeness: $\mathbf{Pr}_{(x_1,x_2,\ldots,x_k)\in D}\big[x_i < x_2 < \cdots < x_k\big] = 1$.
- Soundness: For every permutation $\pi \in S_k$ and every q-ordering $\mathcal{O}_q$ of $[m]$, the probability over random linear extensions of $\mathcal{O}_q$ that a sample $(x_1, x_2, \ldots, x_k) \in D$ is ordered according to π is at most $\frac{1}{k!} + \eta$.

Theorems 2.4 and 2.5 together imply the main UG-hardness result for all OCSPs, and hence we obtain Theorem 1.3.

3. Preliminaries. For a positive integer q, Δ_q denotes the set of corners of the q-dimensional simplex, i.e., $\Delta_q = \{\mathbf{e}_i \mid i \in [q]\}$, where $\mathbf{e}_i$ is the unit vector in the ith dimension. Let $\blacktriangle_q$ denote the convex hull of the set Δ_q; in other words, $\blacktriangle_q$ is the q-dimensional simplex. More generally, for a set S, we use $\blacktriangle(S)$ to denote the set of probability distributions over the set S. For two sets A, B, let A^B denote the set of functions from B to A. For notational convenience, if $B = [n]$, then we write A^n instead of $A^{[n]}$. Let $o_\tau(1)$ denote a term that goes to 0 as $\tau \to 0$, while keeping all other parameters fixed.

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

Given an ordering $\mathcal{O}$ of the vertices of a directed graph G or more generally variables in an OCSP, we use $\mathrm{val}(\mathcal{O})$ to denote the fraction of constraints satisfied by $\mathcal{O}$. Furthermore, for a directed graph G, let $\mathrm{opt}(G)$ denote the largest value of $\mathrm{val}(\mathcal{O})$ for an ordering $\mathcal{O}$ of the vertices of the G. The quantities $\mathrm{val}_q(\mathcal{O})$ and $\mathrm{opt}_q(G)$ are defined analogously for q-Order $\mathcal{O}$ using Definition 2.2.

OBSERVATION 3.1. *For all directed graphs G and integers $q \leqslant q'$,* $\mathrm{opt}_q(G) \leqslant \mathrm{opt}_{q'}(G) \leqslant \mathrm{opt}(G)$.

While the first part of the inequality is trivial, let us elaborate on the latter half. Given a q'-ordering $\mathcal{O}^*$, construct a full ordering $\mathcal{O}$ by using a random permutation of the elements within each of the q' blocks, while retaining the natural order between the blocks. It is easy to check that the expected value of the ordering $\mathcal{O}$ is exactly equal to $\mathrm{val}_q(\mathcal{O}^*)$, thus proving the latter half of the inequality.

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 $L_2(\Omega)$ of real-valued functions over $[m]$ with the inner product

$$\langle f, g\rangle = \mathop{\mathbb{E}}_{x\in[m]}[f(x)g(x)].$$

For $\sigma \in \{0, 1, \ldots, m-1\}^R$, define

$$\chi_\sigma(\mathbf{z}) = \prod_{k\in[R]} \chi_{\sigma_k}(z^{(k)}).$$

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

$$||\mathcal{F}||_2^2 = \sum_{\sigma} \hat{\mathcal{F}}^2(\sigma).$$

For the sake of brevity, we denote $\langle m\rangle = \{0, 1, \ldots, m-1\}$. For $\sigma \in \langle m\rangle^R$, we define its "weight" $|\sigma|$ as

$$|\sigma| = \Big|\, \{i \in [R] \mid \sigma_i \neq 0\} \,\Big|.$$

Definition 3.2. *For a function $\mathcal{F} : \Omega^R \to \mathbb{R}$, define*

$$\mathrm{Inf}_k(\mathcal{F}) = \mathop{\mathbb{E}}_{\mathbf{z}}[\mathop{\mathbf{Var}}_{z^{(k)}}[\mathcal{F}]] = \sum_{\sigma:\sigma_k \neq 0} \hat{\mathcal{F}}^2(\sigma).$$

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

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

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.

It is useful for us that indicator functions of small support that have no influential coordinates are not very stable under the noise operator T_ρ.

Lemma 3.4. *For every $\varepsilon > 0$, there exists a $\mu_0 > 0$ such that for all $\mu < \mu_0$ the following holds: Let $\mathcal{F} : [m]^R \to [0, 1]$ be any function with $\mathbb{E}[\mathcal{F}] = \mu$, and let*

$$\mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{F}) \leqslant \tau$$

for all $k \in \{1, 2, \ldots, R\}$. Then,

$$||T_{1-2\varepsilon}\mathcal{F}||_2^2 \leqslant \mu^{1+\varepsilon/2} + o_\tau(1).$$

Proof. The lemma essentially follows from the Majority is Stablest theorem (see Theorem 4.4 in [31]). We have

$$\begin{aligned} ||T_{1-2\varepsilon}\mathcal{F}||_2^2 = \sum_{\sigma \in \langle m\rangle^R} (1-2\varepsilon)^{2|\sigma|} \hat{\mathcal{F}}^2(\sigma) &\leqslant \sum_{\sigma \in \langle m\rangle^R} (1-\varepsilon)^{|\sigma|} \hat{\mathcal{F}}(\sigma)(1-\varepsilon)^{2|\sigma|} \hat{\mathcal{F}}(\sigma) \\ &\leqslant \mathbb{E}[(T_{1-\varepsilon}\mathcal{F})(\mathbf{x}) T_{1-\varepsilon}(T_{1-\varepsilon}\mathcal{F})(\mathbf{x})]. \end{aligned}$$

Since the influences of $T_{1-\varepsilon}\mathcal{F}$ are low, we can apply Theorem 4.4 from [31] to bound the last expression by noise stability in Gaussian space $\Gamma_{1-\varepsilon}(\mu)$:

$$\mathbb{E}[(T_{1-\varepsilon}\mathcal{F}) T_{1-\varepsilon}(T_{1-\varepsilon}\mathcal{F})] \leqslant \Gamma_{1-\varepsilon}(\mu) + o_\tau(1).$$

By Theorem B.5 from [31], $\Gamma_{1-\varepsilon}(\mu)$ is bounded by $\mu^{1+\varepsilon/2}$ for μ small enough compared to ε, establishing the desired bound. ∎

We have the following immediate consequence of Lemma 3.4.

Lemma 3.5. *Let $\mathcal{F}, \mathcal{G} : [m]^R \to [0, 1]$ be any two functions satisfying the assumption of Lemma 3.4, and let $\mathbf{x}, \mathbf{y}$ be random vectors in $[m]^R$ whose marginal distributions are uniform over $[m]^R$ but are arbitrarily correlated. Then,*

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

Proof. The quantity in question is upper bounded by $||T_{1-2\varepsilon}\mathcal{F}||_2||T_{1-2\varepsilon}\mathcal{G}||_2$ by the Cauchy–Schwarz inequality. The result now follows from the previous lemma. ∎

The following lemma is useful in bounding the number of influential coordinates of a function.

LEMMA 3.6 (sum of influences lemma). *Given a function* $\mathcal{F} : [m]^R \to [0,1]$*, if* $\mathcal{H} = T_{1-\varepsilon}\mathcal{F}$*, then*

$$\sum_{k=1}^{R} \mathrm{Inf}_k(\mathcal{H}) \leqslant \frac{1}{2e\ln 1/(1-\varepsilon)} \leqslant \frac{1}{\varepsilon}.$$

Proof. Let $\mathcal{F}(\mathbf{x}) = \sum_\sigma \hat{\mathcal{F}}(\sigma)\chi_\sigma(\mathbf{x})$ denote the multilinear expansion of $\mathcal{F}$. The function $\mathcal{H}$ is given by $\mathcal{H}(\mathbf{x}) = \sum_\sigma (1-\varepsilon)^{|\sigma|}\hat{\mathcal{F}}(\sigma)\chi_\sigma(x)$. Hence we get

$$\begin{aligned}\sum_{i=1}^{R} \mathrm{Inf}_i(\mathcal{H}) &= \sum_{i=1}^{R}\sum_{\sigma,\sigma_i\neq 0}(1-\varepsilon)^{2|\sigma|}\hat{\mathcal{F}}^2(\sigma) = \sum_\sigma (1-\varepsilon)^{2|\sigma|}|\sigma|\hat{\mathcal{F}}^2(\sigma)\\ &\leqslant \max_{\sigma\in\langle m\rangle^R}\left((1-\varepsilon)^{2|\sigma|}|\sigma|\right)\cdot\sum_\sigma \hat{\mathcal{F}}(\sigma)^2 \leqslant \max_\sigma (1-\varepsilon)^{2|\sigma|}|\sigma|.\end{aligned}$$

The function $h(x) = x(1-\varepsilon)^{2x}$ achieves a maximum at $x = -1/2\ln(1-\varepsilon)$. Substituting, we get $\sum_{i=1}^{R}\mathrm{Inf}_i(\mathcal{H}) \leqslant \frac{1}{2e\ln 1/(1-\varepsilon)} \leqslant \frac{1}{\varepsilon}$. ∎

3.2. Semidefinite program. We use the following natural SDP relaxation of the MAS problem. Given a directed graph $G = (V,E)$ with $|V| = n$, the program has n variables $\{\boldsymbol{b}_{u,i} \mid i \in [n]\}$ for each vertex $u \in V$ and a unit vector $\mathbf{I}$ representing the constant 1. In the intended solution, we have $\boldsymbol{b}_{u,i} = \mathbf{I}$ and $\boldsymbol{b}_{u,j} = 0$ for all $j \neq i$ if u is assigned the ith location in the ordering.

MAS-SDP Relaxation

(MAS − SDP)

$$\text{maximize} \quad \mathop{\mathbb{E}}_{e=(u,v)\sim E}\left[\sum_{\substack{i<j\\ i,j\in[n]}}\langle \boldsymbol{b}_{u,i},\boldsymbol{b}_{v,j}\rangle + \frac{1}{2}\sum_{i\in[n]}\langle \boldsymbol{b}_{u,i},\boldsymbol{b}_{v,i}\rangle\right]$$

(2)

subject to $\langle \boldsymbol{b}_{u,i},\boldsymbol{b}_{u,j}\rangle = 0 \quad \forall\, u\in V, i,j\in[n], i\neq j,$

(3) $\langle \boldsymbol{b}_{u,i},\boldsymbol{b}_{v,j}\rangle \geqslant 0 \quad \forall\, u,v\in V, i,j\in[n],$

(4) $\sum_{i\in[n]}\|\boldsymbol{b}_{u,i}\|_2^2 = 1 \quad \forall u\in V,$

(5) $\langle \boldsymbol{b}_{u,i},\mathbf{I}\rangle = \|\boldsymbol{b}_{u,i}\|_2^2 \quad \forall u\in V, i\in[n],$

(6) $\|\mathbf{I}\|_2^2 = 1.$

The above semidefinite program has the same set of constraints as the relaxations for Max Dicut [12], Linear Equations Mod p [1], and UG [23, 8].

The program can also be written succinctly in terms of distributions over local integral assignments. Specifically, define a set of probability distributions $\boldsymbol{\mu} = \{\mu_e \mid e\in E\}$ over $[n]^2$, one for each edge. The probability distribution μ_e is to be thought of as a distribution over local assignments to the vertices of the edge e.

LC Relaxation for MAS

(7)

$$\text{maximize} \quad \mathop{\mathbb{E}}_{e=(u,v)\sim E}\left[\mathop{\mathbf{Pr}}_{(x_u,x_v)\in\mu_e}\left\{x_u < x_v\right\} + \frac{1}{2}\mathop{\mathbf{Pr}}_{(x_u,x_v)\in\mu_e}\left\{x_u = x_v\right\}\right]$$

$$\text{subject to} \quad \langle \boldsymbol{b}_{u,i}, \boldsymbol{b}_{v,j}\rangle = \mathop{\mathbf{Pr}}_{(x_u,x_v)\in\mu_e}\left\{x_u = i, x_v = j\right\} \quad (e=(u,v)\in E,\ \ i,j\in[n]),$$

$$\mu_e \in \blacktriangle([n]^2) \qquad \forall e \in E.$$

3.3. UGC. Let us give a formal definition of the constraint satisfaction problem that underlies this famous conjecture.

DEFINITION 3.7. *An instance of* UG *represented as* $\Phi = (\mathcal{A}_\Phi \cup \mathcal{B}_\Phi, E, \Pi, [R])$ *consists of a bipartite graph over node sets* $\mathcal{A}_\Phi, \mathcal{B}_\Phi$ *with the edges* E *between them. Also part of the instance is a set of labels* $[R] = \{1, \ldots, R\}$ *and a set of bijections* $\pi_{a\to b} : [R] \to [R]$ *for each edge* $e = (a, b) \in E$*, where* $a \in \mathcal{A}_\Phi$ *and* $b \in \mathcal{B}_\Phi$*. (We will sometimes also denote the bijection* $\pi_{a\to b}$ *for an edge* $e = (a, b)$ *by* π_e*.)*

An assignment $A : \mathcal{A}_\Phi \cup \mathcal{B}_\Phi \to [R]$ *of labels to vertices is said to satisfy an edge* $e = (a, b)$ *if* $\pi_{a\to b}(A(a)) = A(b)$*. The objective is to find an assignment* A *of labels that satisfies the maximum number of edges.*

For the sake of convenience, we use the following version of the UGC, which was shown to be equivalent to the original conjecture [25].

CONJECTURE 3.8 (UGC). *For every* $\delta > 0$*, the following problem is* NP*-hard for a sufficiently large choice of* R*: Given a bipartite* UG *instance* $\Phi = (\mathcal{A}_\Phi \cup \mathcal{B}_\Phi, E, \Pi = \{\pi_{a\to b} : [R] \to [R] \mid e = (a, b) \in E\}, [R])$ *with number of labels* R*, distinguish between the following two cases:*

- $(1-\delta)$-strongly satisfiable instances: *There exists an assignment* A *of labels such that a fraction* $(1-\delta)$ *of vertices* $w \in \mathcal{A}_\Phi$ *are strongly satisfied; i.e.,* all *the edges* (w, v) *are satisfied.*
- Instances that are not δ-satisfiable: *No assignment satisfies more than a* δ*-fraction of the edges* E*.*

4. Orderings and their influences. In this section, we develop the notions of influences for orderings and prove some basic results about them.

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

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

For an ordering $\mathcal{O}$ of points in $[m]^R$, we have functions $\mathcal{F}^{[s,t]} : [m]^R \to \{0,1\}$ for integers s, t defined by (1). For the sake of brevity, we write $\mathcal{F}^i$ for $\mathcal{F}^{[i,i]}$, and $\boldsymbol{\mathcal{F}} = (\mathcal{F}^1, \ldots, \mathcal{F}^q)$.

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

$$\mathsf{L}_\tau(\mathcal{O}) = \{k \mid \mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{F}^{[s,t]}) \geqslant \tau \text{ for some } s,t \in \mathbb{Z}\}.$$

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

It is not difficult to see that we can bound the number of influential coordinates.

LEMMA 4.3 (few influential coordinates). *For any ordering* $\mathcal{O}$ *of* $[m]^R$*, we have* $|\mathsf{L}_\tau(\mathcal{O})| \leqslant \frac{400}{\varepsilon\tau^3}$.

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

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

Hence, using $a^2 \leqslant 2(b^2 + (a-b)^2)$, we get

$$\mathrm{Inf}_k(f) = \sum_{\sigma:\sigma_k \neq 0} \hat{f}^2(\sigma) \leqslant 2\left[\sum_{\sigma_k \neq 0} \hat{g}^2(\sigma) + \sum_{\sigma_k \neq 0}\left(\hat{f}(\sigma) - \hat{g}(\sigma)\right)^2\right] \leqslant 2 \cdot \mathrm{Inf}_k(g) + \frac{\tau}{2}.$$

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 $N = \{\mathcal{F}^{[s,t]}\}$ of size at most $100/\tau^2$ such that for every $\mathcal{F}^{[s',t']}$ there is a $\mathcal{F}^{[s,t]} \in N$ such that $\max |s-s'|, |t-t'| < \frac{\tau m^R}{8}$. Further, by Lemma 3.6, each function $T_{1-\varepsilon}\mathcal{F}^{[s,t]}$ has at most $\frac{4}{\varepsilon\tau}$ coordinates with influence more than $\tau/4$. Hence, $|\mathsf{L}_\tau(\mathcal{O})| \leqslant \frac{400}{\varepsilon\tau^3}$. ∎

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

Proof. Since the functions $\{\mathcal{F}^{[\cdot,\cdot]}\}$ with respect to the ordering $\mathcal{O}^*$ are a subset of the same functions with respect to $\mathcal{O}$, we have $S_\tau(\mathcal{O}^*) \subseteq S_\tau(\mathcal{O})$. ∎

5. Gap instances for MAS. In this section, we construct DAGs with no good q-ordering. These graphs are crucial in designing the dictatorship test in section 6. Actually, in section 11, we construct such instances for ordering constraints of higher arity, which in particular proves the existence of the needed graphs. In particular, Lemma 5.3 is a special case of Theorem 11.1 when the arity k equals 2. However, for self-contained treatment of the MAS result, we present the specialized construction for graphs separately in this section. Even though it is of little importance for our applications, we note that the constants obtained in this section are superior to those of the general construction.

DEFINITION 5.1. *For $\eta > 0$ and a positive integer q, an (η, q)-pseudorandom DAG is a weighted directed graph $G = (V, E)$ with the following properties:*

$$\mathrm{opt}(G) = 1 \quad \textit{and} \quad \mathrm{opt}_q(G) \leqslant \frac{1}{2} + \eta.$$

Clearly, if $\mathrm{opt}(G) = 1$, then the value of the LC relaxation for MAS (from section 3.2) on G is also at least 1. Thus, a pseudorandom DAG as above gives a "weak" integrality gap, where the optimum for q-orderings is small. Specifically, an (η, q)-pseudorandom DAG is also a $(q, 1, 1/2 + \eta)$-multiscale gap instance for MAS, in the sense of Definition 2.3. The formal claim, along with certain smoothness properties of the SDP solution, is made at the end of this section in Corollary 5.4. We now turn to the construction of (η, q)-pseudorandom DAGs.

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

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

We need the following theorem from [7] relating the cut norm of a directed graph G to $\mathrm{opt}(G)$.

THEOREM 5.2 (Theorem 3.1 in [7]). *If a directed graph G on n vertices has an acyclic subgraph with at least a fraction $(\frac{1}{2} + \delta)$ of the edges, then $||G||_C \geqslant \Omega\left(\frac{\delta}{\log n}\right)$.*

The following lemma constructs (η, q)-pseudorandom DAGs from graphs that are the "tight cases" of the above theorem.

LEMMA 5.3. *Given $\eta > 0$ and a positive integer q, for every sufficiently large n, there exists a directed graph $G = (V, E)$ on n vertices such that* $\text{opt}(G) = 1$, $\text{opt}_q(G) \leqslant \frac{1}{2} + \eta$.

Proof. Charikar, Makarychev, and Makarychev (section 4 in [7]) 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 $w_{ij} = \sum_{k=1}^{n} \sin \frac{\pi(j-i)k}{n+1}$. It is easy to verify that for every $0 < q < n$, $\sum_{k=1}^{n} \sin\left(\frac{\pi q k}{n+1}\right) \geqslant 0$. Thus, $w_{ij} \geqslant 0$ whenever $i < j$, implying that the graph is acyclic (in other words, $\text{opt}(G) = 1$).

We bound $\text{opt}_q(G)$ as follows. Let $\text{opt}_q(G) = \frac{1}{2} + \delta$, and let $\mathcal{O} : V \to [q]$ be the optimal q-ordering. Construct a (multi)graph H on q 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 $\Omega\left(\frac{\delta}{\log q}\right)$. Moreover, since $\mathcal{O}$ is a partition of V, the cut norm of G is at least the cut norm of H. Thus, $\Omega\left(\frac{\delta}{\log q}\right) \leqslant ||H||_C \leqslant ||G||_C \leqslant O(1/\log n)$. This gives $\delta \leqslant O\left(\frac{\log q}{\log n}\right)$, implying that $\text{opt}_q(G) \leqslant \frac{1}{2} + O\left(\frac{\log q}{\log n}\right)$. Choosing n sufficiently large (specifically $n \geqslant q^{\Omega(1/\eta)}$) gives the required result. □

We now have the following corollary to Lemma 5.3, which shows how to obtain a "smooth" SDP gap instance from the (η, q)-pseudorandom DAG.

COROLLARY 5.4. *For every $\eta > 0$ and positive integer q, there exists a $(q, 1-\eta, 1/2+\eta)$-multiscale gap instance with a corresponding SDP solution $\boldsymbol{V} = \{\boldsymbol{b}_{u,i} \mid u \in V, i \in [|V|]\}$ and $\boldsymbol{\mu} = \{\mu_e \mid e \in E\}$ of objective value $1 - \eta$ which further satisfies*

$$\|\boldsymbol{b}_{u,i}\|_2^2 = 1/|V| \qquad \forall u \in V, i \in [|V|]. \tag{8}$$

Proof. Let $G = (V, E)$ be the graph obtained by taking $b = \lceil 1/\eta \rceil$ disjoint copies of the graph guaranteed by Lemma 5.3, and let $m = |V|$. Note that the graph still satisfies the required properties: $\text{opt}(G) = 1$, $\text{opt}_q(G) \leqslant \frac{1}{2} + \eta$. The ordering, $\mathcal{O}$, that satisfies every edge of G is obtained by taking the good ordering inside any copy and letting each copy have contiguous places in the ordering. Let D denote the distribution over labelings obtained by shifting $\mathcal{O}$ by a random offset cyclically. For every $u \in V$, $i \in [m]$, $\mathbf{Pr}[D(u) = i] = 1/m$. Further, every directed edge is satisfied with probability at least $1 - 1/b \geqslant 1 - \eta$. Being a distribution over integral labelings, D gives rise to a set of vectors satisfying the constraints in (8). The graph G along with these vectors form the claimed multiscale gap instance. □

6. Dictatorship test for MAS.

Let $G = (V, E)$ be a $(q, 1-\eta, 1/2+\eta)$-multiscale gap instance on m vertices, where m is divisible by q, with corresponding SDP solution $(\boldsymbol{V}, \boldsymbol{\mu})$ as guaranteed by Corollary 5.4. Using the graph G and the SDP solution, we construct a dictatorship test $\mathsf{DICT}_G^{\varepsilon}$ on $[m]^R$ as follows.

$\mathsf{DICT}_G^{\varepsilon}$ Test:

- Pick an edge $e = (u, v) \in E$ at random from G.
- Sample $\mathbf{z}_e = \{\mathbf{z}_u, \mathbf{z}_v\}$ from the product distribution μ_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 μ_e given by $\mu_e(i, j) = \langle \boldsymbol{b}_{u,i}, \boldsymbol{b}_{v,j} \rangle$.
- 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 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 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 $\mathcal{O}(\tilde{\mathbf{z}}_u) < \mathcal{O}(\tilde{\mathbf{z}}_v)$).

Note that since the test takes a form of a directed edge, $\mathsf{DICT}_G^\varepsilon$ can be viewed as a weighted MAS instance where the weight of a particular directed edge $\tilde{\mathbf{z}}_u \to \tilde{\mathbf{z}}_v$ is the probability the above test outputs it. Let us first establish that the test indeed accepts dictator orderings with high probability.

LEMMA 6.1.

$$\mathsf{Completeness}(\mathsf{DICT}_G^\varepsilon) \geqslant 1 - \eta - 4\varepsilon.$$

Proof. A dictator m-ordering $\mathcal{O}$ is given by $\mathcal{O}(\mathbf{z}) = z^{(j)}$. With probability $(1-2\varepsilon)^2$, $\tilde{z}_u^{(j)} = z_u^{(j)}$ and $\tilde{z}_u^{(v)} = z_v^{(j)}$. As the value of the ordering of G is at least $1-\eta$, the lemma follows. ∎

THEOREM 6.2 (soundness analysis). *For every $\varepsilon > 0$, there exist sufficiently large m, q such that for any τ-pseudorandom ordering $\mathcal{O}$ of $[m]^R$,*

$$\mathrm{val}(\mathcal{O}) \leqslant \mathrm{opt}_q(G) + O(q^{-\frac{\varepsilon}{2}}) + o_\tau(1).$$

Let $\mathcal{F}^{[s,t]} : [m]^R \to \{0,1\}$ denote the functions associated with the q-ordering $\mathcal{O}^*$, and remember that we write $\mathcal{F}^i$ for $\mathcal{F}^{[i,i]}$. The result follows from Lemmas 6.3 and 6.4.

LEMMA 6.3. *For every $\varepsilon > 0$, there exist sufficiently large m, q such that for any τ-pseudorandom ordering $\mathcal{O}$ of $[m]^R$,*

$$\mathrm{val}(\mathcal{O}) \leqslant \mathrm{val}_q(\mathcal{O}^*) + O(q^{-\frac{\varepsilon}{2}}) + o_\tau(1),$$

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

Proof. The loss in $\mathrm{val}(\mathcal{O})$ due to coarsening is because some edges $e = (\mathbf{z}, \mathbf{z}')$ which are oriented correctly in $\mathcal{O}$ fall into the same block during coarsening, i.e., $\mathcal{O}^*(z) = \mathcal{O}^*(z')$. Thus we can write

$$\mathrm{val}(\mathcal{O}) \leqslant \mathrm{val}_q(\mathcal{O}^*) + \frac{1}{2}\,\mathbf{Pr}\left(\mathcal{O}^*(\tilde{\mathbf{z}}_u) = \mathcal{O}^*(\tilde{\mathbf{z}}_v)\right),$$

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

As $\mathcal{O}^*$ is a q-coarsening of $\mathcal{O}$, for each value $i \in [q]$, there is exactly a fraction $\frac{1}{q}$ of $\mathbf{z}$ for which $\mathcal{O}^*(\mathbf{z}) = i$. Hence, for each $i \in [q]$, $\mathbb{E}_{\mathbf{z}}[\mathcal{F}^i(\mathbf{z}) = \frac{1}{q}]$. Further, since the ordering $\mathcal{O}^*$ is τ-pseudorandom, for every $k \in [R]$ and $i \in [q]$, $\mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{F}^i) \leqslant \tau$. From Corollary 5.4 we know that $\mathbf{z}_u$ and $\mathbf{z}_v$ individually are uniformly distributed, and hence using Lemma 3.5, for sufficiently large q, the above probability is bounded by $q \cdot q^{-1-\frac{\varepsilon}{2}} + q \cdot o_\tau(1) = O(q^{-\frac{\varepsilon}{2}}) + o_\tau(1)$. ∎

We proceed with the other essential lemma to prove Theorem 6.2.

LEMMA 6.4. *For every choice of m, q, ε and any τ-pseudorandom q-ordering $\mathcal{O}^*$ of $[m]^R$,*

$$\mathrm{val}_q(\mathcal{O}^*) \leqslant \mathrm{opt}_q(G) + o_\tau(1).$$

In section 10 we give a proof of the more general Lemma 9.4, and to avoid duplication of arguments we here give only a sketch of the main ideas behind the proof of Lemma 6.4.

The q-ordering problem is a CSP over a finite domain and is thus amenable to techniques of [35]. Specifically, consider the payoff function $P : [q]^2 \to [0, 1]$ defined by $P(i, j) = 1$ for $i < j$, $P(i, j) = 0$ for $i > j$, and $P(i, j) = \frac{1}{2}$ otherwise.

First, we can extend the domain of the payoff $[q]^2$ to $\blacktriangle_q^2$ using the following multilinear extension:

$$P(\mathbf{x}, \mathbf{y}) = \frac{1}{2} \sum_{i=j} x^{(i)} y^{(j)} + \sum_{i<j} x^{(i)} y^{(j)}$$

for all $\mathbf{x} = (x^{(1)}, \ldots, x^{(q)})$, $\mathbf{y} = (y^{(1)}, \ldots, y^{(q)}) \in \blacktriangle_q$.

Let $\mathcal{F}^{[s,t]} : [m]^R \to \{0, 1\}$ denote the functions associated with a τ-pseudorandom q-ordering $\mathcal{O}^*$, and recall that we write $\mathcal{F}^i$ for $\mathcal{F}^{[i,i]}$, and $\boldsymbol{\mathcal{F}} = (\mathcal{F}^1, \ldots, \mathcal{F}^q)$. Arithmetizing $\mathrm{val}_q(\mathcal{O}^*)$ in terms of functions $\mathcal{F}^i$ we get

$$\begin{aligned} \mathrm{val}_q(\mathcal{O}^*) &= \mathop{\mathbb{E}}_{e} \mathop{\mathbb{E}}_{\mathbf{z}_u, \mathbf{z}_v} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v} \left[\frac{1}{2} \sum_{i=j} \mathcal{F}^i(\tilde{\mathbf{z}}_u) \cdot \mathcal{F}^j(\tilde{\mathbf{z}}_v) + \sum_{i<j} \mathcal{F}^i(\tilde{\mathbf{z}}_u) \cdot \mathcal{F}^j(\tilde{\mathbf{z}}_v) \right] \\ &= \mathop{\mathbb{E}}_{e} \mathop{\mathbb{E}}_{\mathbf{z}_u, \mathbf{z}_v} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v} \Big[P(\boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_u), \boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_v)) \Big], \end{aligned}$$

where the expectation is over the edge $e = (u, v)$, $\mathbf{z}_u$, $\mathbf{z}_v$, $\tilde{\mathbf{z}}_u$, and $\tilde{\mathbf{z}}_v$. If we denote $\boldsymbol{\mathcal{H}} = T_{1-\varepsilon}\boldsymbol{\mathcal{F}}$, then, using the multilinearity of P to transfer the expectation inside the application of P, we can rewrite the preceding expression as

$$\mathrm{val}_q(\mathcal{O}^*) = \mathop{\mathbb{E}}_{e} \mathop{\mathbb{E}}_{\mathbf{z}_u, \mathbf{z}_v} \Big[P(\boldsymbol{\mathcal{H}}(\mathbf{z}_u), \boldsymbol{\mathcal{H}}(\mathbf{z}_v)) \Big].$$

Being functions on a product space $[m]^R$, $\boldsymbol{\mathcal{F}}, \boldsymbol{\mathcal{H}}$ can be expressed as vectors of multilinear polynomials in variables $x_{i,j}$, $i \in [m]$, $j \in [R]$, where $x_{i,j}$ is the indicator variable for the event that the jth input takes the value i. Let $\boldsymbol{F}$ and $\boldsymbol{H}$ denote the vector of multilinear polynomials associated with the functions $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{H}}$, respectively.

Let $\{\boldsymbol{b}_{u,i} \mid u \in V, i \in [m]\}$ denote the SDP solution associated with the $(q, 1-\eta, 1/2+\eta)$-multiscale gap instance G. We exhibit a randomized rounding $\mathsf{Round}_{\boldsymbol{\mathcal{F}}}$ of this SDP solution into a q-ordering for the graph G. If $\mathsf{Round}_{\boldsymbol{\mathcal{F}}}(G)$ denotes the expected value of the ordering returned by the rounding scheme, then we show that $\mathsf{Round}_{\boldsymbol{\mathcal{F}}}(G) \approx \mathrm{val}_q(\mathcal{O}^*)$. Clearly, the expected value of the q-ordering returned by the rounding scheme has value at most $\mathrm{opt}_q(G)$. Hence we get

$$\mathrm{val}_q(\mathcal{O}^*) \leqslant \mathsf{Round}_{\boldsymbol{\mathcal{F}}}(G) + o_\tau(1) \leqslant \mathrm{opt}_q(G) + o_\tau(1).$$

The rounding scheme $\mathsf{Round}_{\boldsymbol{\mathcal{F}}}$ proceeds as follows: Pick R random Gaussian vectors, and project the SDP solution along these directions. For each vertex $v \in V$, the values of the projections of corresponding vectors $\{\boldsymbol{b}_{v,i} \mid i \in [m]\}$ are fed as inputs to the multilinear polynomial $\boldsymbol{H}$ to obtain a vector $\mathbf{p}_v$ in $\mathbb{R}^q$. The $\mathbf{p}_v$ is rounded to a point $\mathbf{p}_v^*$ on the q-dimensional simplex $\blacktriangle_q$ using a fairly natural procedure. Finally, the vertex v is assigned a label $\ell \in [q]$ by independently sampling from the distribution $\mathbf{p}_v^*$.

The vector of multilinear polynomials $\boldsymbol{H}$ has no input coordinates with influence greater than τ, since the ordering $\mathcal{O}^*$ is τ-pseudorandom. Furthermore, since $\boldsymbol{H} = T_{1-\varepsilon}\boldsymbol{F}$, the polynomial $\boldsymbol{H}$ is close to a low-degree polynomial.

Roughly speaking, the invariance principle of Mossel [30] asserts that low-degree and low-influence polynomials cannot distinguish between two distributions over inputs with matching moments up to order two. More precisely, the distribution of the output of the multilinear polynomial $\boldsymbol{H}$ depends only on the first two moments of the distribution of inputs. Note that the distribution used in the dictatorship test is inspired by the vectors $\{\boldsymbol{b}_{u,i}\}$. This leads to closeness in the distribution of $\boldsymbol{H}$ when applied to the Gaussians used in Round and $\boldsymbol{H}$ applied to evaluate the payoff of a pseudorandom ordering $\mathcal{O}^\star$. This, in turn, implies that $\mathsf{Round}_{\boldsymbol{\mathcal{F}}}(G) \approx \mathrm{val}_q(\mathcal{O}^*)$, completing the outline of the proof of Lemma 6.4.

Lemma 6.4 asserts that the value of the q-ordering is bounded by $\mathrm{opt}_q(G)+o_\tau(1)$ for all τ-pseudorandom functions $\boldsymbol{\mathcal{F}} = (\mathcal{F}^1, \dots, \mathcal{F}^q)$ that correspond to a q-ordering. Specifically, for each $\mathbf{z} \in [m]^R$, $\boldsymbol{\mathcal{F}}(\mathbf{z})$ is a corner of the simplex; $\boldsymbol{\mathcal{F}}(\mathbf{z}) \in \Delta_q$.

For the UG-hardness reduction, we need the above lemma to hold for the more general class of functions that take values in $\blacktriangle_q$—the q-dimensional simplex—and indeed we need the following stronger claim.

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

$$\mathbb{E}\left[\frac{1}{2}\sum_{i=j}\mathcal{F}^i(\tilde{\mathbf{z}}_u)\mathcal{F}^j(\tilde{\mathbf{z}}_u) + \sum_{i<j}\mathcal{F}^i(\tilde{\mathbf{z}}_u)\mathcal{F}^j(\tilde{\mathbf{z}}_u)\right] \leqslant \mathrm{opt}_q(G) + o_\tau(1),$$

where the expectation is over the edge $e = (u, v)$, $\mathbf{z}_u$, $\mathbf{z}_v$, $\tilde{\mathbf{z}}_u$, *and* $\tilde{\mathbf{z}}_v$.

We give the proof of the above claim in the more general setting (see Lemma 10.5) of OCSPs in section 10.

7. Hardness reduction for MAS. In this section we describe how to turn the dictator test of the previous section into a UG-hardness result for MAS. This is a quite standard procedure, and hence we do not repeat the argument for the case of general k-ary ordering constraints.

Let $G = (V, E)$ be a $(q, 1-\eta, 1/2+\eta)$-multiscale gap instance, and let $\boldsymbol{V} = \{\boldsymbol{b}_{v,i}\}$ and $\boldsymbol{\mu} = \{\mu_e \mid e \in E\}$ be the corresponding SDP solution as guaranteed by Corollary 5.4. Let $m = |V|$.

Let $\Phi = (\mathcal{A}_\Phi \cup \mathcal{B}_\Phi, E, \Pi = \{\pi_e : [R] \to [R] \mid e \in E\}, [R])$ be a bipartite UG instance. Towards constructing a MAS instance $\Psi = (\mathcal{V}, \mathcal{E})$ from Φ, we introduce a long code for each vertex in $\mathcal{B}_\Phi$. Specifically, the set of vertices $\mathcal{V}$ of the directed graph Ψ is indexed by $\mathcal{B}_\Phi \times [m]^R$.

Hardness Reduction:
Input: UG instance $\Phi = (\mathcal{A}_\Phi \cup \mathcal{B}_\Phi, E, \Pi = \{\pi_e : [R] \to [R] \mid e \in E\}, [R])$.
Output: Directed graph $\Psi = (\mathcal{V}, \mathcal{E})$ with set of vertices $\mathcal{V} = \mathcal{B}_\Phi \times [m]^R$ and edges $\mathcal{E}$ given by the following verifier:

- Pick a random vertex $a \in \mathcal{A}_\Phi$. Choose two neighbors $b, b' \in \mathcal{B}_\Phi$ of a independently at random. Let $\pi = \pi_{a\to b}$ and $\pi' = \pi_{a\to b'}$ denote the permutations on the edges (a, b) and (a, b'), respectively.
- Pick an edge $e = (u, v) \in E$ at random from G.
- Sample $\mathbf{z}_e = \{\mathbf{z}_u, \mathbf{z}_v\}$ from the product distribution μ_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 μ_e given by $\mu_e(i, j) = \langle \boldsymbol{b}_{u,i}, \boldsymbol{b}_{v,j}\rangle$.
- 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 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 with the remaining probability, $\tilde{z}_u^{(k)}$ is a new sample from Ω.
- Introduce a directed edge $(b, \pi(\tilde{\mathbf{z}}_u)) \to (b', \pi'(\tilde{\mathbf{z}}_v))$, where for a vector $\mathbf{z} = (z^{(1)}, z^{(2)}, \dots, z^{(R)}) \in [m]^R$ and a permutation σ of $[R]$, $\sigma(\mathbf{z}) \in [m]^R$ is defined by $\sigma(\mathbf{z})^{(i)} = z^{(\sigma^{-1}(i))}$.

THEOREM 7.1. *For every $\eta > 0$, there exists a choice of parameters ε, q, δ such that the following hold:*

- Completeness: *If Φ is a $(1-\delta)$-strongly satisfiable instance of* UG, *then there is an ordering $\mathcal{O}$ for the graph Ψ with value at least $(1-5\eta)$, i.e.,* $\mathrm{val}(\Psi) \geqslant 1 - 5\eta$.
- Soundness: *If Φ is not δ-satisfiable, then no ordering to Ψ has value more than $\frac{1}{2} + 4\eta$, i.e.,* $\mathrm{val}(\Psi) \leqslant \frac{1}{2} + 4\eta$.

In the rest of the section, we present the proof of the above theorem. To begin with, we fix the parameters of the reduction.

Parameters. Fix $\varepsilon = \eta/100$. Let τ, q be the constants obtained from Theorem 7.5. Finally, let us choose $\delta = \min\{\eta/4, \eta\varepsilon^2\tau^8/10^9\}$.

7.1. Completeness. In order to show that $\mathrm{val}(\Psi) \geqslant 1 - 5\eta$, we instead show that $\mathrm{val}_m(\Psi) \geqslant 1 - 5\eta$, which, by Observation 3.1, implies the required result.

By assumption, there exist labelings to the UG instance Φ such that for a fraction $(1-\delta)$ of the vertices $a \in \mathcal{A}_\Phi$ all the edges (a,b) are satisfied. Let $\mathcal{A} : \mathcal{B}_\Phi \cup \mathcal{A}_\Phi \to [R]$ denote one such labeling. Define an m-ordering of Ψ as follows:

$$\mathcal{O}(b, \mathbf{z}) = z^{(\mathcal{A}(b))} \qquad\qquad \forall b \in \mathcal{B}_\Phi, \mathbf{z} \in [m]^R.$$

Clearly the mapping $\mathcal{O} : \mathcal{V} \to [m]$ defines an m-ordering of the vertices $\mathcal{V} = \mathcal{B}_\Phi \times [m]^R$. To determine $\mathrm{val}_m(\mathcal{O})$, 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 whether the edge is satisfied. Arithmetizing this probability, we can write

$$\mathrm{val}_m(\mathcal{O}) = \frac{1}{2}\,\mathbf{Pr}\left(\mathcal{O}(b, \pi(\tilde{\mathbf{z}}_u)) = \mathcal{O}(b', \pi'(\tilde{\mathbf{z}}_v))\right) + \mathbf{Pr}\left(\mathcal{O}(b, \pi(\tilde{\mathbf{z}}_u)) < \mathcal{O}(b', \pi'(\tilde{\mathbf{z}}_v))\right).$$

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

$$\mathrm{val}_m(\mathcal{O}) \geqslant (1-\delta) \cdot \left(\frac{1}{2}\,\mathbf{Pr}\left(\tilde{z}_u^{(l)} = \tilde{z}_v^{(l)}\right) + \mathbf{Pr}\left(\tilde{z}_u^{(l)} < \tilde{z}_v^{(l)}\right)\right).$$

With probability at least $(1-2\varepsilon)^2$, we have $\tilde{z}_u^{(l)} = z_u^{(l)}$ and $\tilde{z}_v^{(l)} = z_v^{(l)}$. Further, note that each coordinate $z_u^{(l)}, z_v^{(l)}$ is generated according to the local distribution μ_e for the edge $e = (u,v)$. Substituting in the expression for $\mathrm{val}_m(\mathcal{O})$, we get

$$\mathrm{val}_m(\mathcal{O}) \geqslant (1-\delta)(1-2\varepsilon)^2 \mathop{\mathbb{E}}_{e=(u,v)}\left[\mathop{\mathbf{Pr}}_{(x_u,x_v)\in\mu_e}\left\{x_u < x_v\right\} + \frac{1}{2}\mathop{\mathbf{Pr}}_{(x_u,x_v)\in\mu_e}\left\{x_u = x_v\right\}\right].$$

Recall that the SDP solution $(\boldsymbol{V}, \boldsymbol{\mu})$ has an objective value at least $(1-\eta)$. Thus for a small enough choice of δ and ε, we have $\mathrm{val}_m(\mathcal{O}) \geqslant 1 - 5\eta$.

7.2. Soundness. Let $\mathcal{O}$ be an ordering of Ψ with $\mathrm{val}(\mathcal{O}) \geqslant \frac{1}{2} + 4\eta$. Using the ordering, we will obtain a labeling $\mathcal{A}$ for the UG instance Φ. Towards this, we build machinery to deal with multiple long codes. For $b \in \mathcal{B}_\Phi$, 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 $a \in \mathcal{A}_\Phi$, 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 fix a vertex $a \in \mathcal{A}_\Phi$ and analyze the long codes corresponding to all neighbors of a. For ease of notation, for a neighbor $b \in N(a)$, we use π_b to denote the permutation $\pi_{a\to b}$ along the edge (a, b). Let $\mathcal{F}_b^{[s,t]}$ denote the functions associated with the ordering $\mathcal{O}_b$. Define functions $\mathcal{F}_a^{[s,t]} : [m]^R \to \mathbb{R}$ as follows:

$$\mathcal{F}_a^{[s,t]}(\mathbf{z}) = \underset{b\in N(a)}{\mathbf{Pr}} \Big(\mathcal{O}_a(b, \pi_b(\mathbf{z})) \in [s, t]\Big) = \underset{b\in N(a)}{\mathbb{E}} [\mathcal{F}_b^{[s,t]}(\pi_b(\mathbf{z}))].$$

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

$$\mathsf{L}_\tau(\mathcal{O}_a) = \big\{k \;\big|\; \mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{F}_a^{[s,t]}) \geqslant \tau \text{ for some } s, t \in \mathbb{Z}\big\}.$$

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

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

Proof. As the coordinate k is influential on $\mathcal{O}_a$, there exist s, t such that $\mathrm{Inf}_k(\mathcal{F}_a^{[s,t]}) \geqslant \tau$. Recall that $\mathcal{F}_a^{[s,t]}(\mathbf{z}) = \mathbb{E}_{b\in N(a)}[\mathcal{F}_b^{[s,t]}(\pi_b(\mathbf{z}))]$. Using the convexity of Inf, this implies $\mathbb{E}_{b\in N(a)}[\mathrm{Inf}_{\pi_b(k)}(\mathcal{F}_b^{[s,t]})] \geqslant \tau$. All the influences $\mathrm{Inf}_{\pi_b(k)}(\mathcal{F}_b^{[s,t]})$ are bounded by 1, since each of the functions $\mathcal{F}_b^{[s,t]}$ takes values in the range $[0, 1]$. Therefore, for at least a fraction $\tau/2$ of vertices $b \in N(a)$, we have $\mathrm{Inf}_{\pi_b(k)}(\mathcal{F}_b^{[s,t]}) \geqslant \tau/2$. This concludes the proof. ∎

Lemma 7.4. *For any vertex* $a \in \mathcal{A}_\Phi$*,* $|\mathsf{L}_\tau(\mathcal{O}_a)| \leqslant \frac{800}{\varepsilon\tau^4}$*.*

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

It is in fact not difficult to get a better bound than obtained in Lemma 7.4 by applying an extension of Lemma 4.3 directly to the function $\mathcal{F}_a$. Note that the lemma does not apply directly, as $\mathcal{O}_a$ is not an ordering but a set of orderings. This extension is not difficult, but the improvement in parameters is not a real concern to us.

Theorem 7.5. *For all* $\varepsilon, \eta > 0$*, there exist constants* $q, \tau > 0$ *such that for any vertex* $a \in \mathcal{A}_\Phi$*, if* $\mathcal{O}_a$ *is* τ*-pseudorandom, then* $\mathrm{val}(\mathcal{O}_a) \leqslant \mathrm{opt}_q(G) + \eta/4$*.*

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

$$\mathrm{val}(\mathcal{O}_a) \leqslant \mathrm{val}_q(\mathcal{O}_a^*) + \frac{1}{2}\,\mathbf{Pr}\Big(\mathcal{O}_a^*(b, \pi_b(\tilde{\mathbf{z}}_u)) = \mathcal{O}_a^*(b', \pi_{b'}(\tilde{\mathbf{z}}_v))\Big).$$

The q-coarsening $\mathcal{O}_a^*$ is obtained by dividing the order $\mathcal{O}_a$ into q-blocks. Let $[1 = p_1 + 1, p_2], [p_2 + 1, p_3], \ldots, [p_q + 1, p_{q+1} = m]$ denote the q 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

$$\begin{aligned}
&\mathbf{Pr}\Big(\mathcal{O}_a^*(b, \pi_b(\tilde{\mathbf{z}}_u)) = \mathcal{O}_a^*(b', \pi_{b'}(\tilde{\mathbf{z}}_v))\Big)\\
&\quad= \sum_{i\in[q]} \underset{e=(u,v)}{\mathbb{E}}\ \underset{b,b'}{\mathbb{E}}\ \underset{\mathbf{z}_u,\mathbf{z}_v,\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v}{\mathbb{E}} \Big[\mathcal{F}_b^i(\pi_b(\tilde{\mathbf{z}}_u)) \cdot \mathcal{F}_{b'}^i(\pi_{b'}(\tilde{\mathbf{z}}_v))\Big]\\
&\quad= \sum_{i\in[q]} \underset{e=(u,v)}{\mathbb{E}}\ \underset{\mathbf{z}_u,\mathbf{z}_v}{\mathbb{E}}\ \underset{\tilde{\mathbf{z}}_u,\tilde{\mathbf{z}}_v}{\mathbb{E}} \Big[\mathcal{F}_a^i(\tilde{\mathbf{z}}_u) \cdot \mathcal{F}_a^i(\tilde{\mathbf{z}}_v)\Big]\\
&\quad= \sum_{i\in[q]} \underset{e=(u,v)}{\mathbb{E}}\ \underset{\mathbf{z}_u,\mathbf{z}_v}{\mathbb{E}} \Big[T_{1-2\varepsilon}\mathcal{F}_a^i(\mathbf{z}_u) \cdot T_{1-2\varepsilon}\mathcal{F}_a^i(\mathbf{z}_v)\Big].
\end{aligned}$$

As the ordering $\mathcal{O}_a$ is τ-pseudorandom, for every $k \in [R]$ and $i \in [q]$, $\mathrm{Inf}_k(T_{1-\varepsilon}\mathcal{F}_a^i) \leqslant \tau$. Hence by the fact that $\mathbf{z}_u$ and $\mathbf{z}_v$ are uniformly distributed and using Lemma 3.5, the above value is less than $O(q^{-\frac{\varepsilon}{2}}) + o_\tau(1)$.

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

$$\begin{aligned}\mathrm{val}_q(\mathcal{O}_a^*) &= \mathbb{E}\left[\frac{1}{2}\sum_{i=j}\mathcal{F}_b^i(\pi_b(\tilde{\mathbf{z}}_u))\cdot\mathcal{F}_{b'}^j(\pi_{b'}(\tilde{\mathbf{z}}_v)) + \sum_{i<j}\mathcal{F}_b^i(\pi_b(\tilde{\mathbf{z}}_u))\cdot\mathcal{F}_{b'}^j(\pi_{b'}(\tilde{\mathbf{z}}_v))\right]\\ &= \mathbb{E}\left[\frac{1}{2}\sum_{i=j}\mathcal{F}_a^i(\tilde{\mathbf{z}}_u)\cdot\mathcal{F}_a^j(\tilde{\mathbf{z}}_v) + \sum_{i<j}\mathcal{F}_a^i(\tilde{\mathbf{z}}_u)\cdot\mathcal{F}_a^j(\tilde{\mathbf{z}}_v)\right].\end{aligned}$$

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

7.2.2. Defining a labeling. Define the labeling $\mathcal{A}$ for the UG instance Φ as follows: For each $a \in \mathcal{A}_\Phi$, $\mathcal{A}(a)$ is a uniformly random element from $\mathsf{L}_\tau(\mathcal{O}_a)$ if it is nonempty and a random label otherwise. Similarly, for each $b \in \mathcal{B}_\Phi$, assign $\mathcal{A}(b)$ to be a random element of $\mathsf{L}_{\tau/2}(\mathcal{O}_b)$ if it is nonempty and an arbitrary label otherwise.

If $\mathrm{val}(\mathcal{O}) = \mathbb{E}_{a\in\mathcal{A}_\Phi}[\mathrm{val}(\mathcal{O}_a)] \geqslant \frac{1}{2} + 4\eta$, then for at least a fraction 2η of vertices $a \in \mathcal{A}_\Phi$, we have $\mathrm{val}(\mathcal{O}_a) \geqslant \frac{1}{2} + 2\eta$. 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 $\mathsf{L}_\tau(\mathcal{O}_a)$ is nonempty. Further, by Lemma 7.3, for every label $l \in \mathsf{L}_\tau(\mathcal{O}_a)$, for at least a fraction $\tau/2$ of the neighbors, $b \in N(a)$, $\pi_b(l)$ belongs to $\mathsf{L}_{\tau/2}(\mathcal{O}_b)$. For every such b, the edge (a, b) is satisfied with probability at least $1/|\mathsf{L}_\tau(\mathcal{O}_a)| \times 1/|\mathsf{L}_{\tau/2}(\mathcal{O}_b)|$. By Lemmas 4.3 and 7.4, this probability is at least $\varepsilon\tau^4/800 \times \varepsilon\tau^3/3200$. Summarizing the argument, the expected fraction of edges satisfied by the labeling $\mathcal{A}$ is at least $\eta\varepsilon^2\tau^8/10240000$. By a small enough choice of δ (the soundness of the original UG instance), this yields the required result and completes the proof of Theorem 7.1.

8. OCSPs. In this section, we outline the ideas of the proof of Theorem 1.3. To this end, we begin by formally defining a class of ordering constraint satisfaction problems.

8.1. Formal definitions.

DEFINITION 8.1. *An ordering constraint satisfaction problem (OCSP) Λ is specified by a probability distribution over the family of* payoff functions $P : S_k \to [0, 1]$ *on the set S_k of permutations on k elements. The integer k is referred to as the arity of the OCSP Λ.*

An example of an OCSP would be all instances that contain 75% of constraints of the form "i before j" and 25% of constraints of the form "i between j and k." Hence the definition fixes not only the set of predicates but also the proportion of each predicate that appears in an instance.

Let us use the notation $P \sim \Lambda$ to denote a payoff sampled from the distribution Λ. Notice that every payoff $P \sim \Lambda$ is assumed to be on the set of permutations of exactly k elements. However, there is no loss of generality, since for every $q \leqslant k$, a payoff on set Π_q of permutations on q elements can be expressed as a payoff on S_k by including dummy variables.

Let $\Pi_{k\to\mathbb{N}}$ denote the set of one-to-one maps from $[k] \to \mathbb{N}$. The domain of a payoff function P can be extended naturally from the set of permutations S_k to $\Pi_{k\to\mathbb{N}}$. In particular, an injective map $f \in \Pi_{k\to\mathbb{N}}$, along with the standard ordering on the range $\mathbb{N}$, induces a permutation π_f on $[k]$. To extend the payoff, just define $P(f) = P(\pi_f)$ for all $f \in \Pi_{k\to\mathbb{N}}$.

DEFINITION 8.2 (Λ-OCSP). *An instance $\Im$ of OCSP Λ is given by $\Im = (\mathcal{V}, \mathcal{P})$, where the following hold:*

- *$\mathcal{V} = \{y_1, \ldots, y_m\}$ is the set of variables that need to be ordered. Thus an ordering $\mathcal{O}$ is a one-to-one map from $\mathcal{V}$ to natural numbers $\mathbb{N}$.*
- *$\mathcal{P}$ is a probability distribution over constraints/payoffs applied to subsets of at most k variables from $\mathcal{V}$. More precisely, a sample $P \sim \mathcal{P}$ would be a payoff function from Λ, applied on a sequence of variables $y_S = (y_{s_1}, \ldots, y_{s_k})$. If $\mathcal{O}_{|S}$ denotes the injective map from $y_S \to \mathbb{N}$ obtained by restricting $\mathcal{O}$ to y_S, then the payoff returned is $P(\mathcal{O}_{|S})$.*

  Moreover, the type of payoff $P \sim \mathcal{P}$ sampled from $\mathcal{P}$ is identical to the distribution associated with the OCSP Λ.

For a payoff $P \in \mathcal{P}$, we define $\mathcal{V}(P) \subseteq \mathcal{V}$ to denote the set of variables on which P is applied. The objective is to find an ordering $\mathcal{O}$ of the variables that maximizes the total weighted payoff/expected payoff, i.e.,

$$\mathop{\mathbb{E}}_{P\sim\mathcal{P}}\left[P(\mathcal{O}_{|P})\right].$$

Here $\mathcal{O}_{|P}$ denotes the ordering $\mathcal{O}$ restricted to the variables in $\mathcal{V}(P)$. We define the value $\text{opt}(\mathcal{P})$ *as*

$$\text{opt}(\Im) \stackrel{\text{def}}{=} \max_{\mathcal{O}:\Pi_{\mathcal{V}\to\mathbb{N}}} \mathop{\mathbb{E}}_{P\sim\mathcal{P}} P(\mathcal{O}_{|P}).$$

Observe that if the payoff functions P are predicates, then maximizing the payoff amounts to maximizing the number of constraints satisfied. We use the notions "payoff function" and "constraint" interchangeably. As noted earlier, by reordering the inputs, we can assume that $P(\sigma)$ is maximized when σ is the identity, id.

DEFINITION 8.3. *Given an OCSP Λ, let*

$$\Lambda_{\max} = \mathop{\mathbb{E}}_{P\sim\Lambda}\left[P(\text{id})\right], \qquad \Lambda_{random} = \mathop{\mathbb{E}}_{P\sim\Lambda}\mathop{\mathbb{E}}_{\sigma\in S_k}\left[P(\sigma)\right].$$

With these definitions, we can state the following general UG-hardness for OCSPs.

THEOREM 8.4 (general UG-hardness). *For every $\eta > 0$ and every OCSP of bounded arity k, the following holds: Given an instance of the OCSP Λ that admits an ordering with payoff at least $\Lambda_{\max} - \eta$, it is UG-hard to find an ordering of the instance that achieves a payoff of at least $\Lambda_{random} + \eta$.*

Notice that Theorem 1.3 corresponds to the special case where the probability distribution Λ consists of a single payoff function. For the sake of exposition, we present the proof of Theorem 1.3 here. The proof of the more general Theorem 8.4 is essentially the same.

8.2. Relation to CSPs. An ordering $\mathcal{O}$ can be thought of as an assignment of values from $[m]$ to each variable y_i such that $y_i \neq y_j$ for all $i \neq j$. By suitably extending the payoff functions $P \in \Lambda$, it is possible to eliminate the "one-to-one" condition ($y_i \neq y_j$ whenever $i \neq j$). More precisely, we extend the domain of payoff functions $P \in \Lambda$ from $\Pi_{k\to[m]}$ to $\mathbb{N}^{[k]}$—the set of all maps from $[k]$ to $\mathbb{N}$.

Given an arbitrary function $f : [k] \to \mathbb{N}$, define a probability distribution $\mathcal{D}_f$ on the set of permutations S_k by the following random procedure:

1. For each $j \in \mathbb{N}$ with $f^{-1}(j) \neq \phi$, pick a uniform random permutation π_j of elements in $f^{-1}(j)$.
2. Concatenate the permutations π_j in the natural ordering on $j \in \mathbb{N}$ to obtain the permutation $\pi \in S_k$. For a payoff $P \in \Lambda$, define

$$P(f) = \mathop{\mathbb{E}}_{\pi \sim \mathcal{D}_f} [P(\pi)]. \tag{9}$$

With this extension of payoff functions, the following lemma shows that optimizing over all orderings is equivalent to optimizing over all assignments of values in $[m]$ to variables $\{y_1, \ldots, y_m\}$.

LEMMA 8.5. *For an instance $\Im = (\mathcal{V}, \mathcal{P})$ of a Λ-OCSP with $|\mathcal{V}| = m$, we have*

$$\max_{\mathcal{O} \in \Pi_{\mathcal{V} \to \mathbb{N}}} \mathop{\mathbb{E}}_{P \in \mathcal{P}} P(\mathcal{O}_{|P}) = \max_{f \in [m]^{\mathcal{V}}} \mathop{\mathbb{E}}_{P \in \mathcal{P}} P(f_{|P}).$$

Here $[m]^{\mathcal{V}}$ denotes the set of all functions from $\mathcal{V}$ to $[m]$.

Proof. For every injective map $\mathcal{O} : \mathcal{V} \to \mathbb{N}$, there is an injective map $\mathcal{O}' : \mathcal{V} \to [m]$ corresponding to the permutation induced by $\mathcal{O}$. Clearly, the objective value of $\mathcal{O}$ is the same as $\mathcal{O}'$. Since $\mathcal{O}' \in [m]^{\mathcal{V}}$, we have

$$\max_{\mathcal{O} \in \Pi_{\mathcal{V} \to \mathbb{N}}} \mathop{\mathbb{E}}_{P \in \mathcal{P}} P(\mathcal{O}_{|P}) \leqslant \max_{f \in [m]^{\mathcal{V}}} \mathop{\mathbb{E}}_{P \in \mathcal{P}} P(f_{|P}).$$

As the payoff of an arbitrary function $f : \mathcal{V} \to [m]$ is defined as an expectation of the payoff of permutations, the reverse inequality follows, finishing the proof. ∎

By virtue of Lemma 8.5, the Λ-OCSP instance $\Im = (\mathcal{V}, \mathcal{P})$ is transformed into a CSP over variables $\mathcal{V}$, albeit over a domain $[m]$ whose size is not fixed. Specifically, the problem of finding an optimal ordering $\mathcal{O}$ for the Λ-OCSP instance can be reformulated as computing

$$\operatorname{opt}(\Im) = \max_{y \in [m]^{\mathcal{V}}} \mathop{\mathbb{E}}_{P \in \mathcal{P}} \Big[P(y_{\mathcal{V}(P)}) \Big]. \tag{10}$$

Here P refers to the extended payoff function as defined in (9). For the sake of convenience, we use y_P as shorthand for $y_{\mathcal{V}(P)}$.

Taking the analogy with CSPs a step further, one can define a CSP Λ_q for every positive integer $q > 0$. Given an instance $\Im = (\mathcal{V}, \mathcal{P})$ of Λ-OCSP, the corresponding Λ_q problem is to find a q-ordering that maximizes the expected payoff. Formally, the goal of the Λ_q-CSP instance $\Im$ is to compute an assignment $y \in [q]^m$ that maximizes the following:

$$\operatorname{opt}_q(\Im) = \max_{y \in [q]^m} \mathop{\mathbb{E}}_{P \in \mathcal{P}} \Big[P(y_P) \Big]. \tag{11}$$

The following claim is an easy consequence of the above definitions.

CLAIM 8.6. *For every Λ-OCSP instance $\Im = (\mathcal{V}, \mathcal{P})$ and integers $q \leqslant q'$,*

$$\operatorname{opt}_q(\Im) \leqslant \operatorname{opt}_{q'}(\Im) \leqslant \operatorname{opt}(\Im).$$

Further, if $|\mathcal{V}| = m$, then $\operatorname{opt}_m(\Im) = \operatorname{opt}(\Im)$.

8.3. SDP relaxation. Inspired by the interpretation of a Λ-OCSP as a CSP over a large domain, one can formulate a generic semidefinite program along the lines of [35]. The details of the generic semidefinite program are described here.

Given a Λ-OCSP instance $\Im = (\mathcal{V}, \mathcal{P})$, the goal is to find a collection of vectors $\{\boldsymbol{b}_{i,a}\}_{i\in\mathcal{V},a\in[m]}$ in a sufficiently high-dimensional space and a collection $\{\mu_P\}_{P\in\text{supp}(\mathcal{P})}$ of distributions over local assignments. For each payoff $P \in \mathcal{P}$, the distribution μ_P is a distribution over $[m]^{\mathcal{V}(P)}$ corresponding to assignments for the variables $\mathcal{V}(P)$. We write $\mathbf{Pr}_{x\in\mu_P}\{E\}$ to denote the probability of an event E under the distribution μ_P.

LC Relaxation

(LC)

$$\text{maximize} \quad \mathop{\mathbb{E}}_{P\sim\mathcal{P}} \mathop{\mathbb{E}}_{x\sim\mu_P} P(x)$$

(12)

$$\text{subject to} \quad \langle \boldsymbol{b}_{s,i}, \boldsymbol{b}_{s',j}\rangle = \mathop{\mathbf{Pr}}_{x\sim\mu_P}\Big\{x_s = i, x_{s'} = j\Big\} \quad (P \in \text{supp}(\mathcal{P}),\ \ s, s' \in \mathcal{V}(P),\ \ i, j \in [m]),$$

$$\mu_P \in \blacktriangle([m]^{\mathcal{V}(P)}) \qquad \forall P \in \text{supp}(\mathcal{P}).$$

We claim that the above optimization problem can be solved in polynomial time. To show this claim, let us introduce additional real-valued variables $\mu_{P,x}$ for $P \in \text{supp}(\mathcal{P})$ and $x \in [m]^{V(P)}$. We add the constraints $\mu_{P,x} \geqslant 0$ and $\sum_{x\in[m]^{V(P)}} \mu_{P,x} = 1$. We can now make the following substitutions to eliminate the distributions μ_P:

$$\mathop{\mathbb{E}}_{x\sim\mu_P} P(x) = \sum_{x\in[m]^{V(P)}} P(x)\mu_{P,x}, \qquad \mathop{\mathbf{Pr}}_{x\sim\mu_P}\Big\{x_i = a\Big\} = \sum_{\substack{x\in[m]^{V(P)}\\ x_i=a}} \mu_{P,x},$$

$$\mathop{\mathbf{Pr}}_{x\sim\mu_P}\Big\{x_i = a, x_j = b\Big\} = \sum_{\substack{x\in[m]^{V(P)}\\ x_i=a, x_j=b}} \mu_{P,x}.$$

After substituting the distributions μ_P by the scalar variables $\mu_{P,x}$, it is clear that an optimal solution to the relaxation of $\mathcal{P}$ can be computed in time $\text{poly}(m^k, |\text{supp}(\mathcal{P})|)$ using standard results about semidefinite programming.

The LC relaxation succinctly encodes several constraints. In the following claim, we present some of the additional properties that a feasible solution to LC can be assumed to satisfy.

CLAIM 8.7. *Given a feasible solution $\{\boldsymbol{b}_{s,i} \mid s \in \mathcal{V},\ i \in [m]\}$, $\boldsymbol{\mu} = \{\mu_e \mid e \in E\}$ to the LC relaxation, the vectors can be transformed to another SDP solution $\{\boldsymbol{b}^*_{s,i}\}$ with the same objective value such that for some unit vector $\mathbf{I}$ the following hold:*

$$\langle \boldsymbol{b}^*_{s,i}, \boldsymbol{b}^*_{s,j}\rangle = 0 \qquad \forall\ i, j \in [m],\ i \neq j,$$

$$\sum_{i\in[m]} \langle \boldsymbol{b}^*_{s,i}, \boldsymbol{b}^*_{s,i}\rangle = 1,$$

$$\sum_{i\in[m]} \boldsymbol{b}^*_{s,i} = \mathbf{I} \qquad \forall s \in \mathcal{V},$$

$$\langle \boldsymbol{b}^*_{s,i}, \mathbf{I}\rangle = \|\boldsymbol{b}^*_{s,i}\|_2^2 \qquad \forall s \in \mathcal{V}, i \in [m],$$

$$\|\mathbf{I}\|_2^2 = 1.$$

We do not formally verify this claim, but any reader that doubts the claim can include these conditions, as they are of the correct form, into LC. In any case from now on we assume that the conditions of Claim 8.7 are fulfilled.

Note that an integrality gap instance to the above relaxation would be a Λ-OCSP instance, $\Im$, such that $\mathrm{sdp}(\Im)$ is "large" while $\mathrm{opt}(\Im)$ is "small." A *multiscale gap* instance, on the other hand, has much weaker properties—requiring only $\mathrm{opt}_q(\Im)$ to be small—thus making it easier to construct. Recall Definition 2.3 of multiscale gap instances: An instance $\Im$ of a Λ-OCSP is a (q, c, s)-multiscale gap instance if $\mathrm{sdp}(\Im) \geqslant c$ and $\mathrm{opt}_q(\Im) \leqslant s$.

8.3.1. Smoothing gap instances. Let us start with a definition.

DEFINITION 8.8. *For $\alpha > 0$, a (q, c, s)-multiscale gap instance $\Im = (\mathcal{V}, \mathcal{P})$ over m variables is said to be α-smooth if for every $P \in \mathcal{P}$ and $x \in [m]^k$, $\mu_{P,x} \geqslant \alpha$.*

Here we outline a transformation on a multiscale gap instance $\Im^*$ to another multiscale gap instance $\Im$ with certain special properties including α-smoothness. In particular, the lemma implies that the smoothness parameter of the resulting solutions is $\alpha = \frac{\eta}{10m^k}$.

LEMMA 8.9. *For all $\eta > 0$, the following holds: given a (q, c, s)-multiscale gap instance $\Im^* = (\mathcal{V}^*, \mathcal{P}^*)$ of a Λ-OCSP, for large enough m, there exists a $(q, c - \eta/5, s + \eta/5)$-multiscale gap instance $\Im = (\mathcal{V}, \mathcal{P})$ on m variables, an SDP solution $\{\boldsymbol{b}_{s,i}\}_{s\in\mathcal{V}, i\in[m]}, \{\mu_P\}_{P\in\mathrm{supp}(\mathcal{P})}$, and a vector $\mathbf{I}$ satisfying*

$$\langle \boldsymbol{b}_{v,i}, \boldsymbol{b}_{v,i}\rangle = \frac{1}{m} \qquad \forall v \in \mathcal{V}, i \in [m], \tag{13}$$

$$\mu_{P,x} \geqslant \frac{\eta}{10m^k} \qquad \forall P \in \mathcal{P}, x \in [m]^k, \tag{14}$$

and

$$\mathop{\mathbb{E}}_{P\sim\mathcal{P}} \mathop{\mathbb{E}}_{x\sim\mu_P} P(x) \geqslant c - \frac{\eta}{5}, \qquad \mathrm{opt}_q(\Im) \leqslant s + \frac{\eta}{5}.$$

Note that although $\mathbf{I}$ does not appear in the claim explicitly it does so implicitly by our assumption that the conditions of Claim 8.7 are valid.

Proof. Intuitively, the SDP solution corresponding to instance $\Im$ assigns each of the variables in $\mathcal{V}$ to each of the locations in $[m]$ with equal probability. The instance $\Im$ is constructed by taking many copies of $\Im^*$ and joining them side by side such that cyclic shifts of orderings obtain around the same payoff.

More formally, let $L = \lceil \frac{20}{\eta} \rceil$ and set $\mathcal{V} = \mathcal{V}^* \times [L]$. The distribution $\mathcal{P}$ is simply the product of the distribution $\mathcal{P}^*$ and the uniform distribution over $[L]$. That is, for every $p = (y_1, y_2, \ldots, y_k)$ in the support of $\mathcal{P}^*$ and for every $l \in [L]$, $\mathbf{Pr}_{\mathcal{P}}((y_1, l), (y_2, l), \ldots, (y_k, l)) = \mathbf{Pr}_{\mathcal{P}^*}(p)/L$.

Let $\mathcal{O}$ be an optimal ordering for $\Im$. Let $m = |\mathcal{V}| = L|\mathcal{V}^*|$. For every $i \in [m]$, define ordering $\mathcal{O}^*_{(i)} : \mathcal{V} \to [m]$ to be $\mathcal{O}^*(v, k) = i + k|\mathcal{V}| + \mathcal{O}(v)$ (addition modulo m). Since, except for at most one copy of $\mathcal{P}^*$, every other constraint is ordered as in $\mathcal{O}$, the payoff of $\mathcal{O}^*_{(i)}$ is at least $c - \eta/20$.

Further, since the q-ordering value of $\mathcal{P}$ is simply the average of the q-ordering values of the individual pieces, $\mathrm{val}_q(\mathcal{P}) \leqslant s$.

To construct the vectors, we consider the distribution over assignments obtained by taking, with probability $1 - \eta/10$, one of $\mathcal{O}^*_{(i)}$ with equal probability and taking a completely random assignment with probability $\eta/10$. It is easy to see that the probability that $y \in \mathcal{V}$ is assigned to $a \in [m]$ is exactly $1/m$. Further, since we take a completely random assignment with probability $\eta/10$, for any constraints $p \in \mathcal{P}$ and $x \in [m]^k$, the distribution assigns x to p with probability at least $\frac{\eta}{10m^k}$. The payoff obtained by this distribution is at least $(1-\eta/10)(c-\eta/20) \geqslant c-\eta/5$. The distribution over assignments naturally gives vectors satisfying the required constraints. ∎

9. Dictatorship test for OCSP. In this section, we construct a dictatorship test for an OCSP Λ starting with a multiscale gap instance $\Im$ for the problem. Formally, let $\Im^* = (\mathcal{V}^*, \mathcal{P}^*)$ be a (q, c, s)-multiscale gap instance with $|\mathcal{V}| = m$. Let $\Im = (\mathcal{V}, \mathcal{P})$ denote the $(q, c - \frac{\eta}{5}, s + \frac{\eta}{5})$-multiscale gap instance, which is $\alpha = \eta/10m^k$-smooth, obtained from Lemma 8.9. Let $(\boldsymbol{V}, \boldsymbol{\mu})$ denote the SDP solution associated with the instance $\Im$. Define a dictatorship test $\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}$ on orderings $\mathcal{O}$ of $[m]^R$ as follows.

$\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}$ Test
Let $\Im = (\mathcal{V}, \mathcal{P})$ denote a $(q, c - \frac{\eta}{5}, s + \frac{\eta}{5})$-multiscale gap instance for OCSP Λ, which is $\alpha = \eta/10m^k$-smooth. Let $(\boldsymbol{V}, \boldsymbol{\mu})$ denote the SDP solution associated with the instance $\Im$.

- Sample a payoff P from the distribution $\mathcal{P}$. Let $\mathcal{V}(P) = S = \{s_1, s_2, \ldots, s_k\}$.
- Sample $\mathbf{z}_S = \{\mathbf{z}_{s_1}, \ldots, \mathbf{z}_{s_k}\}$ from the product distribution μ_P^R; i.e., for each $1 \leqslant j \leqslant R$, $z_S^{(j)} = \{z_{s_1}^{(j)}, \ldots, z_{s_k}^{(j)}\}$ is sampled using the local distribution μ_P on $[m]^{\mathcal{V}(P)}$.
- For each $s_i \in S$ and each $1 \leqslant j \leqslant R$, sample $\tilde{z}^j_{s_i}$ as follows: With probability $(1-\varepsilon)$, $\tilde{z}^{(j)}_{s_i} = z^{(j)}_{s_i}$, and with the remaining probability, $\tilde{z}^{(j)}_{s_i}$ is a uniform random element from $[m]$.
- Query the ordering values $\mathcal{O}(\tilde{\mathbf{z}}_{s_1}), \ldots, \mathcal{O}(\tilde{\mathbf{z}}_{s_k})$.
- Return the payoff: $P\big(\mathcal{O}(\tilde{\mathbf{z}}_{s_1}), \ldots, \mathcal{O}(\tilde{\mathbf{z}}_{s_k})\big)$.

9.1. Completeness analysis. It is fairly simple to check that the completeness of the dictatorship test $\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}$ is close to the SDP value of $\Im$. Specifically, we now state the following lemma.

LEMMA 9.1.

$$\mathsf{Completeness}(\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}) \geqslant \mathrm{val}(\boldsymbol{V}, \boldsymbol{\mu}) - \varepsilon k = c - \frac{\eta}{5} - \varepsilon k.$$

Proof. A dictator m-ordering $\mathcal{O}$ is given by $\mathcal{O}(\mathbf{z}) = z^{(j)}$. The expected payoff returned by the verifier $\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}$ on $\mathcal{O}$ is given by

$$\mathop{\mathbb{E}}_{P \in \mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z_S}} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_S} \Big[P\Big(\mathcal{O}\big(\tilde{\mathbf{z}}_{s_1}\big), \ldots, \mathcal{O}\big(\tilde{\mathbf{z}}_{s_k}\big)\Big)\Big] = \mathop{\mathbb{E}}_{P \in \mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z_S}} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_S} \Big[P_S\Big(\tilde{z}^{(j)}_{s_1}, \ldots, \tilde{z}^{(j)}_{s_k}\Big)\Big].$$

With probability $(1-\varepsilon)^k$, $\tilde{z}^{(j)}_{s_i} = z^{(j)}_{s_i}$ for each $s_i \in S$. Hence a lower bound for the expected payoff is given by

$$\mathop{\mathbb{E}}_{P \in \mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z_S}} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_S} \Big[P\Big(\mathcal{O}\big(\tilde{z}_{s_1}\big), \ldots, \mathcal{O}\big(\tilde{z}_{s_q}\big)\Big)\Big] \geqslant (1-\varepsilon)^k \mathop{\mathbb{E}}_{P \in \mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z_S}} \Big[P\Big(z^{(j)}_{s_1}, \ldots, z^{(j)}_{s_q}\Big)\Big].$$

The jth coordinates $\mathbf{z}^{(j)}_S = \{z^{(j)}_{s_1}, \ldots, z^{(j)}_{s_q}\}$ are generated from the local probability distribution μ_P. Thus we get

$$\mathop{\mathbb{E}}_{P \in \mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z_S}} \Big[P\Big(z^{(j)}_{s_1}, \ldots, z^{(j)}_{s_q}\Big)\Big] = \mathop{\mathbb{E}}_{P \in \mathcal{P}} \mathop{\mathbb{E}}_{x \in \mu_P} \Big[P(x) \Big] = \mathrm{val}(\boldsymbol{V}, \boldsymbol{\mu}). \tag{15}$$

The expected payoff is at least $(1-\varepsilon)^k \cdot \mathrm{val}(\boldsymbol{V}, \boldsymbol{\mu}) \geqslant \mathrm{val}(\boldsymbol{V}, \boldsymbol{\mu}) - \varepsilon k$. ◻

9.2. Soundness of dictatorship test. Let us state our soundness claim.

THEOREM 9.2 (soundness analysis). *For every $\varepsilon > 0$, for any τ-pseudorandom ordering $\mathcal{O}$ of $[m]^R$,*

$$\mathrm{val}(\mathcal{O}) \leqslant \mathrm{opt}_q(\Im) + O(q^{-\frac{\varepsilon}{2}}) + o_\tau(1).$$

This theorem is an immediate consequence of Lemmas 9.3 and 9.4, and let us turn to the first of these statements.

LEMMA 9.3. *For every $\varepsilon > 0$, for any τ-pseudorandom ordering $\mathcal{O}$ of $[m]^R$,*

$$\mathrm{val}(\mathcal{O}) \leqslant \mathrm{val}_q(\mathcal{O}^*) + \binom{k}{2} q^{-\frac{\varepsilon}{2}} + o_\tau(1),$$

where $\mathcal{O}^*$ is the q-coarsening of $\mathcal{O}$ and k denotes the arity of the OCSP Λ.

Proof. Let $\mathcal{F}^{[s,t]} : [m]^R \to \{0,1\}$ denote the functions associated with the q-ordering $\mathcal{O}^*$.

Note that the loss due to coarsening arises because for some payoffs P the k variables in $\mathcal{V}(P)$ do not fall into distinct bins during coarsening. Let us upper bound the probability that some two of the variables queried, $\tilde{\mathbf{z}}_{s_i}$ and $\tilde{\mathbf{z}}_{s_j}$, fall into the same block during coarsening, i.e., $\mathcal{O}^*(\tilde{\mathbf{z}}_{s_i}) = \mathcal{O}^*(\tilde{\mathbf{z}}_{s_j})$. Observe that

$$\begin{aligned}\mathbf{Pr}\left(\mathcal{O}^*(\tilde{\mathbf{z}}_{s_i}) = \mathcal{O}^*(\tilde{\mathbf{z}}_{s_j})\right) &= \sum_{i\in[q]} \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z}_{s_i},\mathbf{z}_{s_j}} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_{s_i},\tilde{\mathbf{z}}_{s_j}} \left[\mathcal{F}^i(\tilde{\mathbf{z}}_{s_i}) \cdot \mathcal{F}^i(\tilde{\mathbf{z}}_{s_j})\right] \\ &= \sum_{i\in[q]} \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z}_{s_i},\mathbf{z}_{s_j}} \left[T_{1-2\varepsilon}\mathcal{F}^i(\mathbf{z}_{s_i}) \cdot T_{1-2\varepsilon}\mathcal{F}^i(\mathbf{z}_{s_j})\right].\end{aligned}$$

As $\mathcal{O}$ is a q-coarsening of $\mathcal{O}$, for each value $i \in [q]$, there is exactly a fraction $\frac{1}{q}$ of $\mathbf{z}$ for which $\mathcal{O}^*(\mathbf{z}) = i$. Hence, for each $i \in [q]$, $\mathbb{E}_{\mathbf{z}}[\mathcal{F}^i(\mathbf{z}) = \frac{1}{q}]$. Further, since the ordering $\mathcal{O}^*$ is τ-pseudorandom, for every $j \in [R]$ and $i \in [q]$, $\mathrm{Inf}_j(T_{1-\varepsilon}\mathcal{F}^i) \leqslant \tau$. Hence, using Lemma 3.5 and the fact that $\mathbf{z}_{s_i}$ and $\mathbf{z}_{s_j}$ are uniformly distributed, for sufficiently large q, the above probability is bounded by $q \cdot q^{-1-\frac{\varepsilon}{2}} + q \cdot o_\tau(1)$. By a simple union bound, the probability that two of the queried values fall in the same bin is at most $\binom{k}{2}\big(q \cdot q^{-1-\frac{\varepsilon}{2}} + q \cdot o_\tau(1)\big)$. As all the payoffs are bounded by 1 in absolute value, we can write

$$\begin{aligned}\mathrm{val}(\mathcal{O}) &\leqslant \mathrm{val}_q(\mathcal{O}^*) + \mathbf{Pr}\left(\exists i, j \in [k] \text{ such that } \mathcal{O}^*(\tilde{\mathbf{z}}_{s_i}) = \mathcal{O}^*(\tilde{\mathbf{z}}_{s_j})\right) \\ &\leqslant \mathrm{val}_q(\mathcal{O}^*) + \binom{k}{2} q^{-\frac{\varepsilon}{2}} + o_\tau(1). \qquad \square\end{aligned}$$

We now state the second lemma needed to prove Theorem 9.2. The proof of this lemma is postponed to section 10.

LEMMA 9.4. *For every choice of m, q, ε and any τ-pseudorandom q-ordering $\mathcal{O}^*$ of $[m]^R$, $\mathrm{val}_q(\mathcal{O}^*) \leqslant \mathrm{opt}_q(\Im) + o_\tau(1)$.*

10. Soundness analysis for q-orderings.

In this section, we give the proof of Lemmas 9.4 and 6.4. As Lemma 6.4 is a special case of Lemma 9.4, we restrict ourselves to the proof of Lemma 9.4, which completes the soundness analysis for the dictatorship test for arbitrary OCSPs. The proof of Lemma 9.4 closely resembles the soundness analysis of dictatorship tests for the case of generalized CSPs in [35]. However, in [35], the dictatorship test is analyzed for functions with domain $[q]^R$ and range $\blacktriangle_q$. In our application, we are interested in functions whose domain is $[m]^R$ while the output is in $\blacktriangle_q$ for some q. Hence Lemma 9.4 is not a formal consequence of the lemmas in [35]. We start with some preliminaries and tools.

10.1. Invariance principle.

The following invariance principle is an immediate consequence of Theorem 3.6 in the work of Isaksson and Mossel [20].

THEOREM 10.1 (invariance principle [20]). *Let Ω be a finite probability space with the least nonzero probability of an atom at least $\alpha \leqslant 1/2$. Let $\mathcal{L} = \{\ell_1, \ell_1, \ldots, \ell_m\}$ be an ensemble of random variables over Ω. Let $\mathcal{G} = \{g_1, \ldots, g_m\}$ be an ensemble of Gaussian random variables satisfying the following conditions:*

$$\mathbb{E}[\ell_i] = \mathbb{E}[g_i], \qquad \mathbb{E}[\ell_i^2] = \mathbb{E}[g_i^2], \qquad \mathbb{E}[\ell_i \ell_j] = \mathbb{E}[g_i g_j] \qquad \forall i, j \in [m].$$

Let $K = \log(1/\alpha)$. Let $\mathbf{F} = (F_1, \ldots, F_d)$ denote a vector-valued multilinear polynomial, and let $H_i = (T_{1-\varepsilon} F_i)$ and $\boldsymbol{H} = (H_1, \ldots, H_d)$. Further assume that $\mathrm{Inf}_i(\boldsymbol{H}) \leqslant \tau$ and $\mathbf{Var}[H_i] \leqslant 1$ for all i.

If $\Psi : \mathbb{R}^d \to \mathbb{R}$ is a Lipschitz-continuous function with Lipschitz constant C_0 (with respect to the L_2-norm), then

$$\left| \mathbb{E}\left[\Psi(\boldsymbol{H}(\mathcal{L}^R))\right] - \mathbb{E}\left[\Psi(\boldsymbol{H}(\mathcal{G}^R))\right] \right| \leqslant C_d \cdot C_0 \cdot \tau^{\varepsilon/18K} = o_\tau(1)$$

for some constant C_d depending on d.

10.2. Payoff functions. For the sake of the proof, we extend the payoff functions P corresponding to the CSP Λ_q to a multilinear polynomial on $\blacktriangle_q^k$. Specifically, the payoff functions $P \in \Lambda_q$ are defined over the set $[q]^k$, where k is the arity of Λ. Given a payoff function $P : [q]^k \to [0, 1]$, we define a function $P' : \mathbb{R}^{qk} \to \mathbb{R}$ follows:

– Define the function P' on $\blacktriangle_q^k$ as a multilinear polynomial:

$$P'(\mathbf{x}_1, \ldots, \mathbf{x}_k) = \sum_{\beta \in [q]^k} P(\beta) \prod_{i=1}^{k} x_{(i,\beta_i)} \qquad \forall \{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \in \blacktriangle_q^k.$$

Note that when the inputs belong to Δ_q^k the sum contains only one nonzero element, and we have the following simpler definition:

– The function P' on Δ_q^k equals

$$P'(\mathbf{e}_{\beta_1}, \ldots, \mathbf{e}_{\beta_k}) = P(\beta) \qquad \forall \beta \in [q]^k.$$

Abusing notation, we use $P \in \Lambda_q$ to denote both the payoff function over $[q]^k$ and the corresponding multilinear function (the P' defined above) over $\blacktriangle_q^k$. The domain of the input to P will hopefully be clear from the context.

10.3. Local and global distributions. Now, we describe two ensembles of random variables, namely the local integral ensembles $\mathcal{L}_P$ for each payoff P, and a global Gaussian ensemble $\mathcal{G}$.

DEFINITION 10.2. *For every payoff $P \in \mathcal{P}$ of size at most k, the local distribution μ_P is a distribution over $[m]^{\mathcal{V}(P)}$. In other words, the distribution μ_P is a distribution over assignments to the CSP variables in set $\mathcal{V}(P)$. The corresponding local integral ensemble is a set of random variables $\mathcal{L}_P = \{\ell_{s_1}, \ldots, \ell_{s_k}\}$ each taking values in Δ_m.*

DEFINITION 10.3. *The global ensemble $\mathcal{G} = \{\boldsymbol{g}_s \mid s \in \mathcal{V},\ j \in [m]\}$ is generated by setting $\boldsymbol{g}_s = \{g_{s,1}, \ldots, g_{s,m}\}$, where*

$$g_{s,j} = \langle \mathbf{I}, \boldsymbol{b}_{s,j} \rangle + \langle \boldsymbol{b}_{s,j} - \langle \mathbf{I}, \boldsymbol{b}_{s,j} \rangle \mathbf{I}, \zeta \rangle$$

and ζ is a normal Gaussian random vector of appropriate dimension.

It is easy to see that the local and global integral ensembles have matching moments up to degree two. Let $\ell_{s,j}$ denote the jth component of ℓ_s.

LEMMA 10.4. *For any set $P \in \mathcal{P}$, the global ensemble $\mathcal{G}$ matches the following moments of the local integral ensemble $\mathcal{L}_P$:*

$$\mathbb{E}[g_{s,j}] = \mathbb{E}[\ell_{s,j}] = \langle \mathbf{I}, \boldsymbol{b}_{s,j}\rangle, \qquad \mathbb{E}[g_{s,j}^2] = \mathbb{E}[\ell_{s,j}^2] = \langle \mathbf{I}, \boldsymbol{b}_{s,j}\rangle,$$
$$\mathbb{E}[g_{s,j}g_{s',j'}] = \mathbb{E}[\ell_{s,j}\ell_{s',j'}] = \langle \boldsymbol{b}_{s,j}, \boldsymbol{b}_{s',j'}\rangle \qquad \forall j, j', s, s'.$$

Proof. The statement of the expectations of the expressions involving the ℓ-variables is easy to check. For the expressions involving the g-variables, we need the fact that

$$E[\langle \boldsymbol{b}, \zeta\rangle\langle \boldsymbol{b}', \zeta\rangle] = \langle \boldsymbol{b}, \boldsymbol{b}'\rangle$$

and $E[\langle \boldsymbol{b}, \zeta\rangle] = 0$ for any vectors $\boldsymbol{b}$ and $\boldsymbol{b}'$. The quantity that requires some calculation to compute is $\mathbb{E}[g_{s,j}g_{s',j'}]$, which equals

$$\langle \mathbf{I}, \boldsymbol{b}_{s,j}\rangle\langle \mathbf{I}, \boldsymbol{b}_{s',j'}\rangle + \langle \boldsymbol{b}_{s,j} - \langle \mathbf{I}, \boldsymbol{b}_{s,j}\rangle \mathbf{I}, \boldsymbol{b}_{s',j'} - \langle \mathbf{I}, \boldsymbol{b}_{s',j'}\rangle \mathbf{I}\rangle = \langle \boldsymbol{b}_{s,j}, \boldsymbol{b}_{s',j'}\rangle. \qquad \square$$

Note that local distributions μ_P for different payoffs $P \in \mathcal{P}$ do not fit together to form a global distribution, and in fact when applying Theorem 10.1 we use this theorem on each term in the payoff function locally. From this we can conclude that the value obtained by the global ensemble on each local condition gives about the same expected contribution to the objective function as the local distribution specific for that constraint.

10.4. Putting it all together. Finally, we now show the following lemma, which forms the core of the soundness argument in Lemma 9.4 and is a generalization of Claim 6.5.

LEMMA 10.5. *For a function $\boldsymbol{\mathcal{F}} : [m]^R \to \blacktriangle_q$ satisfying $\mathrm{Inf}_j(T_{1-\varepsilon}\boldsymbol{\mathcal{F}}) \leqslant \tau$ for all $j \in [R]$,*

$$\mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z}_S} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_S} \Big[P\Big(\boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_{s_1}), \ldots, \boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_{s_k})\Big)\Big] \leqslant \mathrm{opt}_q(\Im) + o_\tau(1).$$

Before proving this lemma let us establish Lemma 9.4.

Proof of Lemma 9.4. Let $\mathcal{F}^{[s,t]} : [m]^R \to \{0,1\}$ denote the functions associated with the q-ordering $\mathcal{O}^*$. The expected payoff returned by the verifier in the dictatorship test $\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}$ is given by

$$\mathrm{val}_q(\mathcal{O}^*) = \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z}_S} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_S} \Big[P\Big(\boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_{s_1}), \ldots, \boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_{s_k})\Big)\Big].$$

Further, since the ordering $\mathcal{O}^*$ is τ-pseudorandom, for every $j \in [R]$, we have $\mathrm{Inf}_j(T_{1-\varepsilon}\mathcal{F}^i) \leqslant \tau$, and thus Lemma 10.5 concludes the proof. $\square$

Let us turn to establishing Lemma 10.5.

Proof of Lemma 10.5. Let us denote $\boldsymbol{\mathcal{H}} = T_{1-\varepsilon}\boldsymbol{\mathcal{F}}$. Let $\mathbf{F}(\mathbf{x}), \boldsymbol{H}(\mathbf{x})$ denote the multilinear polynomials corresponding to functions $\boldsymbol{\mathcal{F}}, \boldsymbol{\mathcal{H}}$, respectively, where the variables $x_{i,j}$ for $i \in [R]$ and $j \in [m]$ can be thought of as indicator variables if the jth input equals i. Let us denote

$$\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}(\boldsymbol{\mathcal{F}}) = \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z}_S} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_S} \Big[P\Big(\boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_{s_1}), \ldots, \boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_{s_k})\Big)\Big].$$

Each vector $\mathbf{z}_{s_i}$ is independently perturbed to obtain $\tilde{\mathbf{z}}_{s_i}$. The payoff functions P are multilinear when restricted to the domain $\blacktriangle_q$. Consequently, we can write

$$\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}(\boldsymbol{\mathcal{F}}) = \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z}_S} \Big[P\Big(\mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_{s_1}} [\boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_{s_1})|\mathbf{z}_{s_1}], \dots, \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_{s_1}} [\boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_{s_q}|\mathbf{z}_{s_k}]) \Big) \Big]$$
$$= \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z}_S} \Big[P\Big(\boldsymbol{\mathcal{H}}(\mathbf{z}_{s_1}), \dots, \boldsymbol{\mathcal{H}}(\mathbf{z}_{s_k}) \Big) \Big].$$

The last equality is due to the fact that $\mathbb{E}_{\tilde{\mathbf{z}}_{s_i}}[\boldsymbol{\mathcal{F}}(\tilde{\mathbf{z}}_{s_i})|\mathbf{z}_{s_i}] = T_{1-\varepsilon}\boldsymbol{\mathcal{F}}(\mathbf{z}_{s_i}) = \boldsymbol{\mathcal{H}}(z_{s_i})$. For each $s \in S$, the coordinates of $\mathbf{z}_s$ are generated by the distribution μ_P. Therefore the above expectation can be written in terms of the polynomial $\boldsymbol{H}$ applied to an instance of the integral ensemble $\mathcal{L}_P$. Specifically, we can write

(16)
$$\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}(\boldsymbol{\mathcal{F}}) = \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathbf{z}_S} \Big[P\Big(\boldsymbol{\mathcal{H}}(\mathbf{z}_{s_1}), \dots, \boldsymbol{\mathcal{H}}(\mathbf{z}_{s_k}) \Big) \Big] = \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathcal{L}_P^R} \Big[P\Big(\boldsymbol{H}(\ell^R_{s_1}), \dots, \boldsymbol{H}(\ell^R_{s_k}) \Big) \Big].$$

The following procedure $\mathsf{Round}_{\boldsymbol{\mathcal{F}}}$ returns an ordering for the original Λ-OCSP instance $\Im$.

$\mathsf{Round}_{\boldsymbol{\mathcal{F}}}$ Scheme

Input. A Λ-OCSP instance $\Im = (\mathcal{V}, \mathcal{P})$ with m variables and an SDP solution $\{\boldsymbol{b}_{v,i}\}, \{\mu_P\}$.

Truncation Function. Let $f_\blacktriangle : \mathbb{R}^q \to \blacktriangle_q$ be a Lipschitz-continuous function such that for all $\mathbf{x} \in \blacktriangle_q$, $f_\blacktriangle(\mathbf{x}) = \mathbf{x}$. Clearly, a function $f_\blacktriangle$ of this nature can be constructed with a Lipschitz constant C_q depending on q.

Scheme. Sample R vectors $\zeta^{(1)}, \dots, \zeta^{(R)}$ with each coordinate being an independent and identically distributed normal random variable.

For each $v \in \mathcal{V}$ do

- For all $1 \leqslant j \leqslant R$ and $i \in [m]$, compute the projection $g^{(j)}_{v,i}$ of the vector $\boldsymbol{b}_{v,i}$ as follows:

$$g^{(j)}_{v,i} = \langle \mathbf{I}, \boldsymbol{b}_{v,i} \rangle + \Big[\langle (\boldsymbol{b}_{v,i} - \langle \mathbf{I}, \boldsymbol{b}_{v,i} \rangle \mathbf{I}), \zeta^{(j)} \rangle \Big].$$

- Evaluate the function $\boldsymbol{\mathcal{H}} = T_{1-\varepsilon}\boldsymbol{\mathcal{F}}$ with $g^{(j)}_{v,i}$ as inputs. In other words, compute $\mathbf{p}_v = (p_{v,1}, \dots, p_{v,q})$ as follows:

$$\mathbf{p}_v = \boldsymbol{H}(\mathbf{g}_\mathbf{v}).$$

- Round $\mathbf{p}_v$ to $\mathbf{p}^*_v \in \blacktriangle_q$ by using the Lipschitz-continuous truncation function $f_\blacktriangle : \mathbb{R}^q \to \blacktriangle_q$:

$$\mathbf{p}^*_v = f_\blacktriangle(\mathbf{p}_v).$$

- Assign the Λ-OCSP variable $v \in \mathcal{V}$ the value $j \in [q]$ with probability $p^*_{v,j}$.

Let $\mathsf{Round}_{\boldsymbol{\mathcal{F}}}(\boldsymbol{V}, \boldsymbol{\mu})$ denote the expected payoff of the ordering returned by the rounding scheme $\mathsf{Round}_{\boldsymbol{\mathcal{F}}}$ on the SDP solution $(\boldsymbol{V}, \boldsymbol{\mu})$ for the Λ-OCSP instance $\Im$. By definition, we have

(17)
$$\mathsf{Round}_{\boldsymbol{\mathcal{F}}}(\boldsymbol{V}, \boldsymbol{\mu}) \leqslant \mathrm{opt}_q(\Im).$$

In the remainder of the proof, we show the following inequality:

$$\mathsf{Round}_{\boldsymbol{\mathcal{F}}}(\boldsymbol{V}, \boldsymbol{\mu}) \geqslant \mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}(\boldsymbol{\mathcal{F}}) - o_\tau(1).$$

Along with (17), this would imply that $\mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}(\boldsymbol{\mathcal{F}})$ is less than $\mathrm{opt}_q(\Im)+o_\tau(1)$, thus showing the required claim. To this end, we arithmetize the value of $\mathsf{Round}_{\boldsymbol{\mathcal{F}}}(\boldsymbol{V},\boldsymbol{\mu})$. Notice that the $\boldsymbol{g}_i$ are nothing but samples of the global ensemble $\mathcal{G}$ associated with $\Im$. By definition, the expected payoff is given by

$$\mathsf{Round}_{\boldsymbol{\mathcal{F}}}(\boldsymbol{V},\boldsymbol{\mu}) = \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathcal{G}_P^R} \left[P\Big(f_{\blacktriangle}(\boldsymbol{H}(\mathbf{g}^R_{s_1})), \ldots, f_{\blacktriangle}(\boldsymbol{H}(\mathbf{g}^R_{s_k}))\Big)\right]. \tag{18}$$

We show that the quantities in (16) and (18) are roughly equal. Fix a payoff $P\in\mathcal{P}$. Let $\Psi_P:\mathbb{R}^{qk}\to\mathbb{R}$ be a Lipschitz-continuous function defined as follows:

$$\Psi_P(\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_k) = P\Big(f_{\blacktriangle}(\mathbf{p}_1),\ldots,f_{\blacktriangle}(\mathbf{p}_k)\Big) \qquad \forall \mathbf{p}_1,\ldots,\mathbf{p}_k\in\blacktriangle_q.$$

Applying the invariance principle (Theorem 10.1) with the ensembles $\mathcal{L}_P$, $\mathcal{G}_P$, Lipschitz-continuous functional Ψ, and the vector of kq multilinear polynomials given by $(\boldsymbol{H},\boldsymbol{H},\ldots,\boldsymbol{H})$ where $\boldsymbol{H}=(H_1,\ldots,H_q)$, we get the required result:

$$\begin{aligned}
\mathsf{Round}_{\boldsymbol{\mathcal{F}}}(\boldsymbol{V},\boldsymbol{\mu}) &= \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathcal{G}_P^R}\left[\Psi_P\Big(\boldsymbol{H}(\mathbf{g}^R_{s_1}),\ldots,\boldsymbol{H}(\mathbf{g}^R_{s_k})\Big)\right] \\
&\geqslant \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathcal{L}_P^R}\left[\Psi_P\Big(\boldsymbol{H}(\ell^R_{s_1}),\ldots,\boldsymbol{H}(\ell^R_{s_k})\Big)\right] - o_\tau(1) \\
&\quad (\because \text{invariance principle (Theorem 10.1)}) \\
&= \mathop{\mathbb{E}}_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\mathcal{L}_P^R}\left[P\Big(\boldsymbol{H}(\ell^R_{s_1}),\ldots,\boldsymbol{H}(\ell^R_{s_k})\Big)\right] - o_\tau(1) \\
&\quad (\because \Psi_P(\mathbf{p}_1,\ldots,\mathbf{p}_k) = P(\mathbf{p}_1,\ldots,\mathbf{p}_k) \text{ if } \forall i,\ \mathbf{p}_i\in\blacktriangle_q) \\
&= \mathsf{DICT}^{\varepsilon}_{\boldsymbol{V},\boldsymbol{\mu}}(\boldsymbol{\mathcal{F}}) - o_\tau(1) \qquad (\because (16)). \qquad \square
\end{aligned}$$

11. Constructing multiscale gap instances for general OCSP. In this section, we prove Theorem 11.1, which is the last piece to complete the proof of Theorem 2.5. We remind the reader that the step moving from the dictator test to the UG-hardness is completely analogous to the transition done in section 7 for MAS and is not presented in this paper.

THEOREM 11.1. *For every $\eta>0$ and positive integers q,k, there is an $m=m(k,q,\eta)$ and a distribution, D, over k-tuples of $[m]$ such that the following hold:*

- *The support of D is contained in the set of strictly increasing k-tuples of $[m]$.*
- *For any $f:[m]\to[q]$, let $\mathcal{D}_f$ denote the distribution over permutations of $[m]$ obtained by extending f as in section 8.2. For any $\sigma=(\sigma_1,\sigma_2,\ldots,\sigma_k)\in S_k$,*

$$\left| \mathop{\mathbf{Pr}}_{y\in\mathcal{D}_f;(d_1,d_2,\ldots,d_k)\in D}[y(d_{\sigma_1})<y(d_{\sigma_2})<\cdots<y(d_{\sigma_k})] - \tfrac{1}{k!}\right| \leqslant \eta.$$

Note that the $k=2$ case of Theorem 11.1 is the content of Lemma 5.3. Before delving into the proof of Theorem 11.1, let us see why it implies a proof of Theorem 2.5.

Proof of Theorem 2.5. Let P be the payoff associated with OCSP Λ, and remember that it is maximized by the identity permutation. Let $m=m(k,q,\eta/k!)$, let D be the distribution as promised by Theorem 11.1, and let $\Im=([m],D)$ be an OCSP instance with payoff P. Almost by definition, the value obtained by the trivial ordering of $[m]$ is the maximal possible $P(\mathrm{id})$. Certainly, the SDP value can be only higher.

Now, viewing $\Im$ as an instance of the OCSP Λ, define f to be an optimal q-ordering of $\Im$. Then, $\mathrm{opt}_q(\Im)$ can be bounded as follows:

$$\mathrm{opt}_q(\Im) = \mathrm{val}(f) = E_D E_{\pi \in \mathcal{D}_f}[P(\pi)] \leqslant \sum_{\sigma} \frac{P(\sigma)}{k!} + \eta \leqslant \Lambda_{random} + \eta. \qquad \square$$

Let us now turn to the proof of Theorem 11.1.

We set $m = k^s$ for some integer s to be chosen depending on q and η, and we think of $[m]$ as the s-tuples of $[k]$, ordered by the lexicographic ordering of the tuples. For an integer r, $0 \leqslant r \leqslant s$, a k-adic interval of order r is an interval of $[m]$ specified by an element $\alpha \in [k]^r$ and denotes the subset of $[k]^s$ whose first r coordinates match those in α. It is easy to check that for every r and α, such a set is, in fact, an interval of $[m]$ (due to the lexicographic ordering) and is of length k^{s-r}. For $r = 0$, it is the entire set $[m]$, and for $r = s$, such a set consists of a single element. A random k-adic interval of order r is a k-adic interval of order r where α is chosen uniformly at random from $[k]^r$.

Every k-adic interval, I, of order r strictly smaller than s naturally contains k disjoint k-adic intervals of order $r + 1$, denoted by $I_1, I_2, \ldots, I_k$ in the order they appear in I. A random k-adic subinterval, J, is obtained by picking one of these k subintervals uniformly at random. Let us define our distribution.

Definition 11.2. *The distribution D_s is a distribution over k-tuples from $[m]$ for $m = k^s$ defined as follows:*

1. *Pick a random r uniformly in $0 \leqslant r \leqslant s - 1$.*
2. *Pick a random k-adic interval I of order r from $[k]^s$.*
3. *Pick x_j uniformly at random from the k-adic subinterval I_j of I, for $j = 1, 2, \ldots, k$.*
4. *Output $(x_1, x_2, \ldots, x_k)$.*

The first claim of Theorem 11.1 follows immediately from the definition, since the elements chosen are always in increasing order in the lexicographic ordering of $[m]$.

In the rest of this section, we prove that, for the distribution D_s, no function $f : [m] \to [q]$ obtains more than negligible advantage over random with respect to any permutation π (for large enough s).

Fix a particular function $f : [m] \to [q]$. For $p \in [q]$ and an interval I, let $\mu_p(I)$ denote the fraction of I mapped to p by f. The following lemma is the heart of our analysis.

Lemma 11.3. *For a random k-adic interval I chosen as in Definition* 11.2 *and a random k-adic subinterval, J, of I, we have*

$$\sum_{p=1}^{q} \mathop{\mathbb{E}}_{I,J}\Big[|\mu_p(I) - \mu_p(J)|\Big] \leqslant \sqrt{\frac{q}{s}}.$$

Proof. Let $\beta_{r,p}$ be $\mathbb{E}[\mu_p(I)^2]$ when I is a random k-adic interval of order r. Note first that as $E[X]^2 \leqslant E[X^2]$ and

$$\mu_p(I) = E_J[\mu_p(J)],$$

where the expectation is over a random k-adic subinterval J of I, we have that $\beta_{r+1,p} \geqslant \beta_{r,p}$. Now, for any $p \in \{1, 2, \ldots, q\}$, and a random k-adic interval I of order r and a random k-adic subinterval J of I,

$$\mathop{\mathbb{E}}_{I,J}\Big[|\mu_p(I)-\mu_p(J)|\Big] \leqslant \Big(\mathop{\mathbb{E}}_{I,J}\big[(\mu_p(J)-\mu_p(I))^2\big]\Big)^{1/2} = \left(\mathop{\mathbb{E}}_{I}\left[\frac{1}{k}\sum_{j=1}^{k}(\mu_p(I_j)-\mu_p(I))^2\right]\right)^{1/2}$$
$$= \left(\mathop{\mathbb{E}}_{I}\left[\frac{1}{k}\sum_{j=1}^{k}\mu_p(I_j)^2-\mu_p(I)^2\right]\right)^{1/2} = (\beta_{r+1,p}-\beta_{r,p})^{1/2}.$$

Thus, averaging over the choice of random $r \in \{0,1,\dots,s-1\}$ and summing over all values p in the range $[q]$, we have

$$\sum_{p=1}^{q}\mathop{\mathbb{E}}_{I,J}\Big[|\mu_p(I)-\mu_p(J)|\Big] \leqslant \frac{1}{s}\sum_{p=1}^{q}\sum_{r=0}^{s-1}(\beta_{r+1,p}-\beta_{r,p})^{1/2}$$
$$\leqslant \frac{1}{s}\sum_{p=1}^{q}\left(\sum_{r=0}^{s-1}1\right)^{1/2}\left(\sum_{r=0}^{s-1}(\beta_{r+1,p}-\beta_{r,p})\right)^{1/2}$$
$$\leqslant \frac{1}{\sqrt{s}}\sum_{p=1}^{q}\beta_{s,p}^{1/2} \leqslant \sqrt{\frac{q}{s}}\left(\sum_{p=1}^{q-1}\beta_{s,p}\right)^{1/2} = \sqrt{\frac{q}{s}}. \qquad \square$$

Next we have the following lemma.

LEMMA 11.4. *Given nonnegative numbers $a_i^{(j)}$, $b_i^{(j)}$, $i \in [q]$, $j \in [k]$, such that for every j, $\sum_i a_i^{(j)} = \sum_i b_i^{(j)} = 1$, we have*

$$\sum_{\sigma\in[q]^k}\left|\prod_{j=1}^{k}a_{\sigma(j)}^{(j)}-\prod_{j=1}^{k}b_{\sigma(j)}^{(j)}\right| \leqslant \sum_{j=1}^{k}\sum_{i=1}^{q}\left|\,a_i^{(j)}-b_i^{(j)}\,\right|,$$

where $\sigma(j)$ denotes the jth element of the k-tuple σ.

Proof. The proof follows by an induction over k. The two sides of the expression are equal for $k=1$. For $k>1$,

$$\sum_{\sigma\in[q]^k}\left|\prod_{j=1}^{k}a_{\sigma(j)}^{(j)}-\prod_{j=1}^{k}b_{\sigma(j)}^{(j)}\right|$$
$$\leqslant \sum_{\sigma\in[q]^k}\left(\left|a_{\sigma(1)}^{(1)}\prod_{j=2}^{k}a_{\sigma(j)}^{(j)}-a_{\sigma(1)}^{(1)}\prod_{j=2}^{k}b_{\sigma(j)}^{(j)}\right|+\left|a_{\sigma(1)}^{(1)}\prod_{j=2}^{k}b_{\sigma(j)}^{(j)}-b_{\sigma(1)}^{(1)}\prod_{j=2}^{k}b_{\sigma(j)}^{(j)}\right|\right)$$
$$=\left(\sum_{i\in[q]}a_i^{(1)}\right)\cdot\sum_{\sigma\in[q]^{k-1}}\left|\prod_{j=1}^{k-1}a_{\sigma(j)}^{(j+1)}-\prod_{j=1}^{k-1}b_{\sigma(j)}^{(j+1)}\right|+\left(\prod_{j=2}^{k}\left(\sum_{i=1}^{q}b_i^{(j)}\right)\right)\sum_{i=1}^{q}\left|a_i^{(1)}-b_i^{(1)}\right|$$
$$=\sum_{\sigma\in[q]^{k-1}}\left|\prod_{j=1}^{k-1}a_{\sigma(j)}^{(j+1)}-\prod_{j=1}^{k-1}b_{\sigma(j)}^{(j+1)}\right|+\sum_{i=1}^{q}\left|a_i^{(1)}-b_i^{(1)}\right|$$
$$\leqslant \sum_{j=2}^{k}\sum_{i=1}^{q}\left|a_i^{(j)}-b_i^{(j)}\right|+\sum_{i=1}^{q}\left|a_i^{(1)}-b_i^{(1)}\right| \quad \text{(by induction hypothesis)}$$
$$=\sum_{j=1}^{k}\sum_{i=1}^{q}\left|a_i^{(j)}-b_i^{(j)}\right|. \qquad \square$$

We can now finish the proof of Theorem 11.1 using the above two lemmas. For any $\sigma \in [q]^k$, let $\Pr(\sigma)$ denote the probability of the event $f(x_j) = \sigma(j)$ when $x = (x_1, x_2, \ldots, x_k)$ is chosen according to the distribution D_s. For an interval I, let $\Pr(\sigma, I)$ denote the above probability conditioned on D_s choosing I in the second claim. Since x_j is chosen uniformly from the k-adic subinterval I_j of I, we have $\Pr(\sigma, I) = \prod_{j=1}^{k} \mu_{\sigma(j)}(I_j)$. Now

$$\begin{aligned}
\sum_{\sigma} \left| \Pr(\sigma) - \mathop{\mathbb{E}}_{I} \left[\prod_j \mu_{\sigma(j)}(I) \right] \right| &= \sum_{\sigma} \left| \mathop{\mathbb{E}}_{I} [\Pr(\sigma, I)] - \mathop{\mathbb{E}}_{I} \left[\prod_j \mu_{\sigma_j}(I) \right] \right| \\
&\leqslant \mathop{\mathbb{E}}_{I} \left[\sum_{\sigma} \left| \Pr(\sigma, I) - \prod_j \mu_{\sigma(j)}(I) \right| \right] \\
&= \mathop{\mathbb{E}}_{I} \left[\sum_{\sigma} \left| \prod_j \mu_{\sigma(j)}(I_j) - \prod_j \mu_{\sigma(j)}(I) \right| \right] \\
&\leqslant \mathop{\mathbb{E}}_{I} \left[\sum_{j=1}^{k} \sum_{p \in [q]} \left| \mu_p(I_j) - \mu_p(I) \right| \right] \qquad \text{(by Lemma 11.4)} \\
&\leqslant k\sqrt{q/s} \qquad \text{(by Lemma 11.3)}.
\end{aligned}$$

For any permutation $\pi \in S_k$, the value of the q-ordering f with respect to π is $\mathrm{val}_t^{\pi}(D_s, f) = \sum_{\sigma} \Pr(\sigma)\mathrm{Payoff}^{\pi}(\sigma)$. Since Payoff takes values in $[0, 1]$, from the above argument,

$$\left| \mathrm{val}_t^{\pi}(D, f) - \sum_{\sigma} \mathrm{Payoff}^{\pi}(\sigma) \mathop{\mathbb{E}}_{I} \left[\Pi_j \mu_{\sigma_j}(I) \right] \right| \leqslant k\sqrt{q/s}.$$

Further, since the value of the second factor in terms of the sum depends on which values appear in σ and not their order, the sum is independent of the permutation π, and we get

$$\left| \mathrm{val}_t^{\pi}(D, f) - \tfrac{1}{k!} \right| \leqslant k\sqrt{q/s}.$$

Choosing s greater than $k^2 q/\eta^2$, we immediately obtain the statement of Theorem 11.1.

Remark 11.5. In Lemma 5.3, we constructed an (η, q)-pseudorandom DAG, which corresponds to the $k = 2$ case of Theorem 11.1, with $n \leqslant q^{O(1/\eta)}$ vertices. The above construction gives an instance of size $k^s = k^{O(k^2 q/\eta^2)}$ which is $\exp(O(q/\eta^2))$ for constant k. This is somewhat worse than the size of the DAG construction from section 5. But the above construction works for all k, and since we treat q, η as constants, the exact dependence of the size on these parameters does not matter for our applications.

12. SDP integrality gap. In this section, we construct integrality gaps for the MAS-SDP relaxation using the hardness reduction from UG. Specifically, we show the following result.

THEOREM 12.1. *For any $\gamma > 0$, there exists a directed graph G such that the value of semidefinite program* (MAS − SDP) *is at least $1 - \gamma$, while* $\mathrm{opt}(G) \leqslant \frac{1}{2} + \gamma$.

The proof uses a bipartite variant of the Khot–Vishnoi UG integrality gap instance [26] as in [35, 29]. Specifically, the following is a direct consequence of [26].

The integrality gap instance $\Phi = (\mathcal{A}_\Phi \cup \mathcal{B}_\Phi, E, \Pi = \{\pi_e : [R] \to [R] \mid e \in E\}, [R])$ presented in [26] is not bipartite. To obtain a bipartite UG instance Φ', duplicate the vertices by setting $\mathcal{A}_\Phi = \{(b, 0) \mid b \in V\}$ and $\mathcal{B}_\Phi = \{(b, 1) \mid b \in V\}$. Further, for each edge $(a, b) \in E$, introduce two edges $((a, 0), (b, 1))$ and $((a, 1), (b, 0))$ in Φ'. The SDP solution for the bipartite instance Φ' is obtained by assigning the vector corresponding to $b \in V$ to both vertices $(b, 0)$ and $(b, 1)$. Except for these minor modifications, the following theorem is a direct consequence of [26].

THEOREM 12.2 (see [26]). *For every $\delta > 0$, there exist a UG instance, $\Phi = (\mathcal{A}_\Phi \cup \mathcal{B}_\Phi, E, \Pi = \{\pi_e : [R] \to [R] \mid e \in E\}, [R])$, and vectors $\{\boldsymbol{v}_{b,\ell}\}$ for every $b \in \mathcal{B}_\Phi$, $\ell \in [R]$, and a unit vector $\mathbf{I}$ such that the following conditions hold:*

- *No assignment satisfies more than a fraction δ of constraints in Π.*
- *For all $b, b' \in \mathcal{B}_\Phi$, $\ell, \ell' \in [R]$, $\langle \boldsymbol{v}_{b,\ell}, \boldsymbol{v}_{b',\ell'} \rangle \geqslant 0$, and, if $\ell \neq \ell'$, $\langle \boldsymbol{v}_{b,\ell}, \boldsymbol{v}_{b,\ell'} \rangle = 0$.*
- *For all $b \in \mathcal{B}_\Phi$, $\sum_{\ell \in [R]} \boldsymbol{v}_{b,\ell} = \mathbf{I}$ and $\langle \boldsymbol{v}_{b,\ell}, \mathbf{I} \rangle = \|\boldsymbol{v}_{b,\ell}\|_2^2$.*
- *The SDP value is at least $1-\delta$: $\mathbb{E}_{a \in \mathcal{A}_\Phi, b, b' \in \mathcal{B}_\Phi} \left[\sum_{\ell \in [R]} \langle \boldsymbol{v}_{b,\pi_{a\to b}(\ell)}, \boldsymbol{v}_{b',\pi'_{a\to b'}(\ell)} \rangle \right] \geqslant 1 - \delta$.*

Proof of Theorem 12.1. Let G be an (η, t)-multiscale gap instance with m vertices. Apply Theorem 12.2, with a sufficiently small δ, to obtain a UGC instance Φ and SDP vectors $\{\boldsymbol{v}_{b,\ell} \mid b \in \mathcal{B}_\Phi,\ \ell \in [R]\}$. Consider the instance Ψ constructed by running the UG-hardness reduction in section 7 on the UG instance Φ. The set of vertices of Ψ is given by $\mathcal{B}_\Phi \times [m]^R$. Set $M = |\mathcal{B}_\Phi| \times m^R$ and $N = |\mathcal{B}_\Phi|$.

The program $(\mathsf{MAS} - \mathsf{SDP})$ on the instance Ψ contains M vectors $\{\mathbf{W}_i^{(b,\mathbf{z})} \mid i \in [M]\}$ for each vertex $(b, \mathbf{z}) \in \mathcal{B}_\Phi \times [m]^R$.

Define a solution to $(\mathsf{MAS} - \mathsf{SDP})$ as follows: Set the vector $\mathbf{I}$ to be the corresponding vector in the instance Φ. For each vertex $(b, \mathbf{z})$ of the graph Ψ, define SDP vectors $\{\mathbf{W}_i^{(b,\mathbf{z})} \mid i \in [M]\}$ as follows:

$$\mathbf{W}_i^{(b,\mathbf{z})} = \begin{cases} \sum_{z_\ell = i} \boldsymbol{v}_{b,\ell} & \forall i \in [m], (b, \mathbf{z}) \in \mathcal{B}_\Phi \times [m]^R, \\ 0 & \forall i \notin [m]. \end{cases}$$

Now we check that the SDP vectors $\{\mathbf{W}_i^{(b,\mathbf{z})}\}$ satisfy conditions (2)–(6) of the MAS-SDP relaxation.

- (Constraint (3)) Since the vectors $\{\boldsymbol{v}_{b,\ell}\}$ have a nonnegative dot product, the vectors $\{\mathbf{W}_i^{(b,\mathbf{z})}\}$ have nonnegative inner products too.
- (Constraint (2)) For a fixed b and $\mathbf{z}$, the vectors $\{\mathbf{W}_i^{(b,\mathbf{z})}\}$ are constructed by *partitioning* the vectors $\{\boldsymbol{v}_{b,\ell}\}$ and assigning the vector sum over the partitions. Hence, for any i, j, the vectors $\mathbf{W}_i^{(b,\mathbf{z})}$ and $\mathbf{W}_j^{(b,\mathbf{z})}$ sum over the disjoint set of ℓ. Thus,

$$\langle \mathbf{W}_i^{(b,\mathbf{z})}, \mathbf{W}_j^{(b,\mathbf{z})} \rangle = \left\langle \sum_{z_\ell = i} \boldsymbol{v}_{b,\ell}, \sum_{z_{\ell'} = j} \boldsymbol{v}_{b,\ell'} \right\rangle = 0.$$

- (Constraint (4)) For every vertex $(b, \mathbf{z})$ we have

$$\sum_{i,j \in [M]} \langle \mathbf{W}_i^{(b,\mathbf{z})}, \mathbf{W}_j^{(b,\mathbf{z})} \rangle = \sum_{\ell,\ell' \in [R]} \langle \boldsymbol{v}_{b,\ell}, \boldsymbol{v}_{b,\ell'} \rangle = \sum_{\ell \in [R]} \langle \boldsymbol{v}_{b,\ell}, \boldsymbol{v}_{b,\ell} \rangle = 1.$$

– (Constraint (5)) For $i \notin m$, we have $\mathbf{W}_i^{(b,\mathbf{z})} = 0$, thereby trivially satisfying constraint (5). For $i \in [m]$, we can write

$$\langle \mathbf{I}, \mathbf{W}_i^{(b,\mathbf{z})} \rangle = \sum_{z_\ell = i} \langle \mathbf{I}, \boldsymbol{v}_{b,\ell} \rangle = \sum_{z_\ell = i} \|\boldsymbol{v}_{b,\ell}\|_2^2.$$

Due to orthogonality of the vectors $\{\boldsymbol{v}_{b,i}\}$ for every vertex $b \in \mathcal{B}_\Phi$, we get

$$\langle \mathbf{W}_i^{(b,\mathbf{z})}, \mathbf{W}_i^{(b,\mathbf{z})} \rangle = \left\langle \sum_{z_\ell = i} \boldsymbol{v}_{b,\ell}, \sum_{z_\ell = i} \boldsymbol{v}_{b,\ell} \right\rangle = \sum_{z_\ell = i} \|\boldsymbol{v}_{b,\ell}\|_2^2 = \langle \mathbf{I}, \mathbf{W}_i^{(b,\mathbf{z})} \rangle.$$

– (Constraint (6)) This is satisfied by the choice of $\mathbf{I}$.

To prove that the SDP value is close to 1, we first fix a particular choice of $a \in \mathcal{A}_\Phi$, $b, b' \in \mathcal{B}_\Phi$. Set $\pi = \pi_{a \to b}$, $\pi' = \pi_{a \to b'}$. The SDP value of edges from $(b, *)$ to $(b', *)$ is

$$\begin{aligned} \mathop{\mathbb{E}}_{e \in G} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v} \sum_{i<j} \langle \mathbf{W}_i^{(b,\pi(\tilde{\mathbf{z}}_u))}, \mathbf{W}_j^{(b',\pi'(\tilde{\mathbf{z}}_v))} \rangle &= \mathop{\mathbb{E}}_{e \in G} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v} \sum_{i<j} \left\langle \left(\sum_{\tilde{\mathbf{z}}_u^\ell = i} \boldsymbol{v}_{b,\ell} \right), \left(\sum_{\tilde{\mathbf{z}}_v^{\ell'} = j} \boldsymbol{v}_{b',\ell'} \right) \right\rangle \\ &\geqslant \sum_\ell (\langle \boldsymbol{v}_{b,\pi(\ell)}, \boldsymbol{v}_{b',\pi'(\ell)} \rangle) \mathop{\mathbb{E}}_{e \in G} \mathop{\mathbf{Pr}}_{\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v} [\tilde{\mathbf{z}}_u^\ell < \tilde{\mathbf{z}}_v^\ell]. \end{aligned}$$

With probability at least $(1 - 2\varepsilon)^2$, $\tilde{\mathbf{z}}_u = \mathbf{z}_u$, $\tilde{\mathbf{z}}_v = \mathbf{z}_v$. Further, since the coordinates of $\mathbf{z}_u, \mathbf{z}_v$ are generated from the multiscale gap instance, G, $\mathbf{Pr}[\mathbf{z}_u^\ell < \mathbf{z}_v^\ell] \geqslant 1 - \eta$. Hence,

$$\mathop{\mathbb{E}}_{e \in G} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v} \sum_{i<j} \langle \mathbf{W}_i^{(b,\pi(\tilde{\mathbf{z}}_u))}, \mathbf{W}_j^{(b',\pi'(\tilde{\mathbf{z}}_v))} \rangle \geqslant (1 - 2\varepsilon)^2 (1 - \eta) \sum_\ell \langle \boldsymbol{v}_{b,\pi(\ell)}, \boldsymbol{v}_{b',\pi'(\ell)} \rangle.$$

Thus, the expected payoff over the whole instance is

$$\begin{aligned} &\mathop{\mathbb{E}}_{a,b,b'} \mathop{\mathbb{E}}_{e \in G} \mathop{\mathbb{E}}_{\tilde{\mathbf{z}}_u, \tilde{\mathbf{z}}_v} \sum_{i<j} \langle \mathbf{W}_i^{(b,\pi(\tilde{\mathbf{z}}_u))}, \mathbf{W}_j^{(b',\pi'(\tilde{\mathbf{z}}_v))} \rangle \\ &\quad \geqslant (1 - 2\varepsilon)^2 (1 - \eta) \mathop{\mathbb{E}}_{a \in \mathcal{A}_\Phi, b, b' \in \mathcal{B}_\Phi} \sum_\ell \langle \boldsymbol{v}_{b,\pi(\ell)}, \boldsymbol{v}_{b',\pi'(\ell)} \rangle \\ &\quad \geqslant (1 - 2\varepsilon)^2 (1 - \eta)(1 - \delta). \end{aligned}$$

Hence, for a sufficiently small choice of parameters ε, η, and δ, the SDP value for Ψ is greater than $1 - \gamma$. On the other hand, the soundness analysis in section 7 (Theorem 7.1) implies that the integral optimum for Ψ is at most $\frac{1}{2} + \gamma$ with a sufficiently small choice of ε, η, and δ. ▯

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

REFERENCES

[1] G. Andersson, L. Engebretsen, and J. Håstad, *A new way of using semidefinite programming with applications to linear equations mod p*, J. Algorithms, 39 (2001), pp. 162–204.

[2] P. Austrin, *Towards sharp inapproximability for any 2-CSP*, SIAM J. Comput., 39 (2010), pp. 2430–2463.

[3] P. Austrin and J. Håstad, *Randomly supported independence and resistance*, SIAM J. Comput., 40 (2011), pp. 1–27.

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

[5] B. Berger and P. W. Shor, *Tight bounds for the maximum acyclic subgraph problem*, J. Algorithms, 25 (1997), pp. 1–18.

[6] M. Charikar, V. Guruswami, and R. Manokaran, *Every permutation CSP of arity* 3 *is approximation resistant*, in Proceedings of the 24th Annual IEEE Conference on Computational Complexity, 2009, pp. 62–73.

[7] M. Charikar, K. Makarychev, and Y. Makarychev, *On the advantage over random for maximum acyclic subgraph*, in Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science, 2007, pp. 625–633.

[8] M. Charikar, Y. Makarychev, and K. Makarychev, *Near-optimal algorithms for unique games*, in Proceedings of the 38th Annual ACM Symposium on Theory of Computing, 2006, pp. 205–214.

[9] B. Chor and M. Sudan, *A geometric approach to betweenness*, SIAM J. Discrete Math., 11 (1998), pp. 511–523.

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

[11] G. Even, J. Naor, B. Schieber, and M. Sudan, *Approximating minimum feedback sets and multicuts in directed graphs*, Algorithmica, 20 (1998), pp. 151–174.

[12] U. Feige and M. X. Goemans, *Approximating the value of two proper proof systems, with applications to MAX-*2*SAT and MAX-DICUT*, in Proceedings of the Third Israel Symposium on Theory of Computing and Systems, 1995, pp. 182–189.

[13] M. X. Goemans and D. P. Williamson, *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming*, J. Assoc. Comput. Mach., 42 (1995), pp. 1115–1145.

[14] V. Guruswami, R. Manokaran, and P. Raghavendra, *Beating the random ordering is hard: Inapproximability of maximum acyclic subgraph*, in Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science, 2008, pp. 573–582.

[15] V. Guruswami and Y. Zhou, *Approximating Bounded Occurrence Ordering CSPs*, manuscript, 2010.

[16] G. Hast, *Beating a Random Assignment*, Ph.D. thesis, Royal Institute of Technology, Stockholm, Sweden, 2005.

[17] J. Håstad, *Some optimal inapproximability results*, J. ACM, 48 (2001), pp. 798–859.

[18] J. Håstad, *Every* 2*-CSP allows nontrivial approximation*, Comput. Complexity, 17 (2008), pp. 549–566.

[19] J. Håstad, *On the approximation resistance of a random predicate*, Comput. Complexity, 18 (2009), pp. 413–434.

[20] M. Isaksson and E. Mossel, *Maximally stable Gaussian partitions with discrete applications*, Israel J. Math., to appear.

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

[22] R. M. Karp, *Reducibility among combinatorial problems*, in Complexity of Computer Computations, Plenum, New York, 1972, pp. 85–103.

[23] S. Khot, *On the power of unique* 2*-prover* 1*-round games*, in Proceedings of the 34th Annual ACM Symposium on Theory of Computing, 2002, pp. 767–775.

[24] 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 (2007), pp. 319–357.

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

[26] S. Khot and N. K. Vishnoi, *The Unique Games conjecture, integrality gap for cut problems and embeddability of negative type metrics into* l_1, in Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science, 2005, pp. 53–62.

[27] C. Lucchesi and D. H. Younger, *A minimax theorem for directed graphs*, J. London Math. Soc. (2), 17 (1978), pp. 369–374.

[28] Y. Makarychev, *Simple Linear Time Approximation Algorithm for Betweenness*, Microsoft Research Technical report MSR-TR-2009-74, 2009.

[29] R. Manokaran, J. S. Naor, P. Raghavendra, and R. Schwartz, *SDP gaps and UGC hardness for multiway cut,* 0*-extension and metric labelling*, in Proceedings of the 40th ACM Symposium on Theory of Computing, 2008, pp. 11–20.

[30] E. Mossel, *Gaussian bounds for noise correlation of functions and tight analysis of long codes*,

in Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science, 2008, pp. 156–165.

[31] E. Mossel, R. O'Donnell, and K. Oleszkiewicz, *Noise stability of functions with low influences: Invariance and optimality*, Ann. of Math. (2), 171 (2010), pp. 295–341.

[32] A. Newman, *Approximating the Maximum Acyclic Subgraph*, Master's thesis, MIT, Cambridge, MA, 2000.

[33] A. Newman, *Cuts and orderings: On semidefinite relaxations for the linear ordering problem*, in Proceedings of APPROX-RANDOM, 2004, pp. 195–206.

[34] C. H. Papadimitriou and M. Yannakakis, *Optimization, approximation, and complexity classes*, J. Comput. System Sci., 43 (1991), pp. 425–440.

[35] P. Raghavendra, *Optimal algorithms and inapproximability results for every CSP?*, in Proceedings of the 40th ACM Symposium on Theory of Computing, 2008, pp. 245–254.

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

[37] P. Seymour, *Packing directed circuits fractionally*, Combinatorica, 15 (1995), pp. 281–288.

[38] U. Zwick, *Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint*, in Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, 1998, pp. 201–210.