THEORY OF COMPUTING, Volume 11 (10), 2015, pp. 257–283
www.theoryofcomputing.org

SPECIAL ISSUE: APPROX-RANDOM 2013

On the NP-Hardness of Approximating Ordering-Constraint Satisfaction Problems

Per Austrin*   Rajsekar Manokaran†   Cenny Wenner‡

Received October 2, 2013; Revised November 9, 2014; Published June 18, 2015

Abstract: We show improved NP-hardness of approximating *Ordering-Constraint Satisfaction Problems* (OCSPs). For the two most well-studied OCSPs, MAXIMUM ACYCLIC SUBGRAPH and MAXIMUM BETWEENNESS, we prove NP-hard approximation factors of $14/15+\varepsilon$ and $1/2+\varepsilon$, respectively. When it is hard to approximate an OCSP by a constant better than taking a uniform random ordering, then the OCSP is said to be *approximation resistant*. We show that the MAXIMUM *Non*-BETWEENNESS PROBLEM is approximation resistant and that there are width-m approximation-resistant OCSPs accepting only a fraction $1/(m/2)!$ of assignments. These results provide the first examples of approximation-resistant OCSPs subject only to $\mathsf{P} \neq \mathsf{NP}$.

ACM Classification: F.2.2

AMS Classification: 68Q17, 68W25

Key words and phrases: acyclic subgraph, APPROX, approximation, approximation resistance, betweenness, constraint satisfaction, CSPs, feedback arc set, hypercontractivity, NP-completeness, orderings, PCP, probabilistically checkable proofs

---

An extended abstract of this paper appeared in the 15th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX 2013) [4].

*Supported by Swedish Research Council Grant 621-2012-4546.
†Supported by ERC Advanced Grant 226203.
‡Supported by ERC Advanced Grant 226203.

 DOI: 10.4086/toc.2015.v011a010

1 Introduction

We study the NP-hardness of approximating a rich class of optimization problems known as *Ordering-Constraint Satisfaction Problems* (OCSPs). An instance of an OCSP is given by a set $\mathcal{X}$ of variables and a set $\mathcal{C}$ of *local ordering constraints* where each constraint is specified by a set of permitted permutations on a subset of the variables. The objective is to find an ordering of $\mathcal{X}$ maximizing the fraction of constraints satisfied by the induced local permutations.

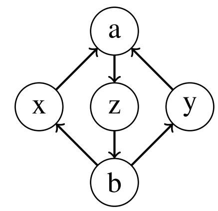

Figure 1: An MAS instance with value $5/6$.

A simple example of an OCSP is the problem MAXIMUM ACYCLIC SUBGRAPH (MAS) where one is given a directed graph $G = (V, A)$ with the task of finding an acyclic subgraph of G containing as many arcs as possible. Phrased as an OCSP, V is the set of variables and each arc $u \to v$ is a constraint "$u \prec v$" dictating that u should precede v. The maximum fraction of constraints simultaneously satisfiable by an ordering of V is then precisely the fraction of edges in a largest acyclic subgraph. Since each constraint involves exactly two variables, MAS is an OCSP of *width* two. Another example of an OCSP is the MAXIMUM BETWEENNESS (MAX BTW) problem [12]. In this width-three OCSP, a constraint on a triplet of variables (x, y, z) is satisfied by the local orderings $x \prec z \prec y$ and $y \prec z \prec x$ – in other words, z has to be between x and y, giving rise to the name of the problem.

Determining the optimal value of an MAS instance is NP-hard and one turns to approximations. An algorithm is called a c-approximation if, when applied to an instance $\mathcal{I}$, the algorithm is guaranteed to produce an ordering $\mathcal{O}$ satisfying a fraction of constraints within a factor c of the optimum, i. e., $\mathsf{val}(\mathcal{O};\mathcal{I}) \geq c \cdot \mathsf{val}(\mathcal{I})$. As in the case of classical constraint satisfaction, every OCSP admits a naive approximation algorithm which picks an ordering of $\mathcal{X}$ uniformly at random without even looking at the constraints. For example, this algorithm yields an $1/2$-approximation in expectation for MAS.

Surprisingly, there is evidence that this mindless procedure achieves the best polynomial-time approximation constant: assuming the Unique Games Conjecture (UGC) [17], MAS is hard to approximate within $1/2 + \varepsilon$ for every $\varepsilon > 0$ [14, 13]. An OCSP is called *approximation resistant* if it exhibits this behavior, i. e., if it is NP-hard to approximate within a constant better than the random-ordering algorithm. In fact, the results of Guruswami et al. [13] are more general and contrasting with classical constraint satisfaction: assuming the UGC, they prove that *every* OCSP of bounded width is approximation resistant.

In many cases—such as for VERTEX COVER, MAX CUT, and as we just mentioned, for all OCSPs—the UGC implies optimal NP-hard inapproximability constants which are not known without the conjecture. For instance, the problems MAS and MAX BTW have, until the present work, only been known to be NP-hard to approximate within $65/66 + \varepsilon$ [19] and $47/48 + \varepsilon$ [9], resp., far from matching the respective random-assignment thresholds of $1/2$ and $1/3$. In fact, while the UGC implies that all OCSPs are approximation resistant, there were no results proving NP-hard approximation resistance of an OCSP prior to this work. In contrast, there is a significant body of work on NP-hard approximation resistance of classical Constraint Satisfaction Problems (CSPs) [16, 21, 11, 6]. Furthermore, the UGC is still very much open and recent algorithmic advances have given rise to subexponential algorithms (as the completeness-error ε goes to 0) for Unique Games [1, 5] calling the conjecture into question. Several recent works have also been aimed at bypassing the UGC for natural problems by providing comparable results without assuming the conjecture [15, 6].

Problem	Approx. factor	UG-inapprox.	NP-inapprox.	This work
MAS	$1/2+\Omega(1/\log n)$ [8]	$1/2+\varepsilon$ [14]	$65/66+\varepsilon$ [19]	$14/15+\varepsilon$
MAX BTW	$1/3$	$1/3+\varepsilon$ [7]	$47/48+\varepsilon$ [9]	$1/2+\varepsilon$
MAX NBTW	$2/3$	$2/3+\varepsilon$ [7]	-	$2/3+\varepsilon$
m-OCSP	$1/m!$	$1/m!+\varepsilon$ [13]	-	$1/\lfloor m/2\rfloor!+\varepsilon$

Table 1: Prior results and improvements.

1.1 Results

In this work we obtain improved NP-hardness of approximating various OCSPs. While a complete characterization such as in the UG regime still eludes us, our results improve the knowledge of what we believe are four important flavors of OCSPs; see Table 1 for a summary of the present state of affairs.

We address the two most studied OCSPs: MAS and MAX BTW. For MAS, we show a factor $(14/15+\varepsilon)$-inapproximability improving the factor from $65/66+\varepsilon$ [19]. For MAX BTW, we show a factor $(1/2+\varepsilon)$-inapproximability improving from $47/48+\varepsilon$ [9].

Theorem 1.1. *For every $\varepsilon > 0$, it is NP-hard to distinguish MAS instances with value at least $15/18-\varepsilon$ from instances with value at most $14/18+\varepsilon$.*

Theorem 1.2. *For every $\varepsilon > 0$, it is NP-hard to distinguish MAX BTW instances with value at least $1-\varepsilon$ from instances with value at most $1/2+\varepsilon$.*

The above two results are inferior to what is known assuming the UGC and in particular do not prove approximation resistance. We introduce the MAXIMUM NON-BETWEENNESS (MAX NBTW) problem which accepts the *complement* of the predicate in MAX BTW. This predicate accepts four of the six permutations on three elements and thus a random ordering satisfies two thirds of the constraints in expectation. We show that this is optimal up to smaller-order terms.

Theorem 1.3. *For every $\varepsilon > 0$, it is NP-hard to distinguish MAX NBTW instances with value at least $1-\varepsilon$ from instances with value at most $2/3+\varepsilon$.*

Finally, we address the approximability of a generic width-m OCSP as a function of the width m. In the CSP world, the generic version is called m-CSP and we call the ordering version m-OCSP. We devise a simple predicate, "$2t$-Same Order" $2t$-SO, on $m=2t$ variables that is satisfied only if the first t elements are relatively ordered exactly as the last t elements. A random ordering satisfies only a fraction $1/t!$ of the constraints and we prove that this is essentially optimal, implying a $(1/\lfloor m/2\rfloor!+\varepsilon)$-factor inapproximability of m-OCSP.

Theorem 1.4. *For every $\varepsilon > 0$ and integer $m \geq 2$, it is NP-hard to distinguish m-OCSP instances with value at least $1-\varepsilon$ from value at most $1/\lfloor m/2\rfloor!+\varepsilon$.*

1.2 Proof overviews

With the exception of MAS, our results follow a route which is by now standard in inapproximability: starting from the optimization problem LABEL COVER (LC), we give reductions using *dictatorship-test* gadgets, also known as long-code tests. We describe these reductions in the context of MAX NBTW to highlight the new techniques in this paper.

The reduction produces an instance $\mathcal{I}$ of MAX NBTW from an instance $\mathcal{L}$ of LC such that $\mathsf{val}(\mathcal{I}) > 1-\varepsilon$ if $\mathsf{val}(\mathcal{L}) = 1$ while $\mathsf{val}(\mathcal{I}) < 2/3+\varepsilon$ if $\mathsf{val}(\mathcal{L}) \le \eta$ for some $\eta = \eta(\varepsilon)$. By the PCP Theorem and the Parallel Repetition Theorem [3, 2, 20], it is NP-hard to distinguish between $\mathsf{val}(\mathcal{L}) = 1$ and $\mathsf{val}(\mathcal{L}) \le \eta$ for every constant $\eta > 0$ and thus we obtain the result in Theorem 1.3. The core component in this paradigm is the design of a dictatorship test: a MAX NBTW instance on $[q]^L \cup [q]^R$, for integers q and label sets L and R. Let π be a map $R \to L$. Each constraint is a tuple $(\vec{\mathbf{x}},\vec{\mathbf{y}},\vec{\mathbf{z}})$ where $\vec{\mathbf{x}} \in [q]^L$, while $\vec{\mathbf{y}},\vec{\mathbf{z}} \in [q]^R$. The distribution of tuples is obtained as follows. First, pick $\vec{\mathbf{x}}$, and $\vec{\mathbf{y}}$ uniformly at random from $[q]^L$, and $[q]^R$. Set $z_j = y_j + x_{\pi(j)} \mod q$. Finally, add noise by independently replacing each coordinate x_i, y_j and z_j with a uniformly random element from $[q]$ with probability γ.

This test instance has canonical assignments that satisfy almost all the constraints. These are obtained by picking an arbitrary $j \in [R]$, and partitioning the variables into q sets $S_0, \dots S_{q-1}$ where

$$S_t = \{\vec{\mathbf{x}} \in [q]^R \mid x_{\pi(j)} = t\} \cup \{\vec{\mathbf{y}} \in [q]^L \mid y_j = t\}\,.$$

If a constraint $(\vec{\mathbf{x}},\vec{\mathbf{y}},\vec{\mathbf{z}})$ is so that $\vec{\mathbf{x}} \in S_t$, $\vec{\mathbf{y}} \in S_u$ then $\vec{\mathbf{z}} \notin S_v$ for any $v \in \{t+1,\dots,u-1\}$ except with probability $O(\gamma)$. This is because $(a+b) \mod q$ is never strictly between a and b. Further, the probability that any two of $\vec{\mathbf{x}}$, $\vec{\mathbf{y}}$, and $\vec{\mathbf{z}}$ fall in the same set S_i is simply the probability that any two of $x_{\pi(j)}, y_j$, and z_j are assigned the same value, which is at most $O(1/q)$. Thus, ordering the variables such that $S_0 \prec S_1 \prec \dots \prec S_{q-1}$ with an arbitrary ordering of the variables within a set satisfies a fraction $1-O(1/q)-O(\gamma)$ constraints.

The proof of Theorem 1.3 requires a partial converse of the above: every ordering that satisfies more than a fraction $2/3+\varepsilon$ of the constraints is more-or-less an ordering that depends on a few coordinates j as above. This proof involves three steps. First, we show that there is a $\Gamma = \Gamma(q,\gamma,\beta)$ such that every ordering $\mathcal{O}$ of $[q]^L$ and $[q]^R$ can be broken into Γ sets $S_0,\dots,S_{\Gamma-1}$ such that one achieves expected value at least $\mathsf{val}(\mathcal{O}) - \beta$ for arbitrarily small β by ordering the sets $S_0 \prec \cdots \prec S_{\Gamma-1}$ and within each set ordering elements arbitrarily. The proof of this "bucketing" uses hypercontractivity of noised functions from a finite domain. We note that a related bucketing argument is used in proving inapproximability of OCSPs assuming the UGC [14, 13]. Somewhat similar to our analysis, they "round" table values to q buckets of equal size and proceed to analyze indicator functions for ranges of order assignments. In particular, the UG-based analysis only needs to consider multiple queries to, and bucketing of, individual tables. For our construction, we have to consider the simultaneous bucketing of two tables. Merely "rounding" the tables to a set $[q]$ does not suffice. Rather, we consider having q equal-sized buckets for each table and analyze functions which do not differentiate values within buckets. More precisely, we introduce a "bucketed payoff" which, given a tuple of bucket indices, samples a random element from the respective indexed buckets of the tables and evaluates the predicate for the resulting tuple. With this setup, we can analyze functions of bounded range, yet the bucketing preserves arbitrarily well the value of an instance even when the tables have influential coordinates.

Similarly to [14, 13], the bucketing argument allows an OCSP to be analyzed as if it were a CSP on a finite domain, enabling us to use powerful techniques developed in this setting. In particular, we show that unless $\mathsf{val}(\mathcal{L}) > \eta$, for some η, then the distribution of constraints $(\vec{\mathbf{x}},\vec{\mathbf{y}},\vec{\mathbf{z}})$ can be regarded as obtained by sampling $\vec{\mathbf{x}}$ independently up to some negligible error in the payoff; in other words, $\vec{\mathbf{x}}$ is "decoupled" from $(\vec{\mathbf{y}},\vec{\mathbf{z}})$. We note that the marginal distribution of the tuple $(\vec{\mathbf{y}},\vec{\mathbf{z}})$ is already symmetric with respect to swaps: $\mathbf{P}(\vec{\mathbf{y}}=y,\vec{\mathbf{z}}=z)=\mathbf{P}(\vec{\mathbf{y}}=z,\vec{\mathbf{z}}=y)$. In order to prove approximation resistance, we combine three of these dictatorship tests: the j^{th} variant has $\vec{\mathbf{x}}$ as the j^{th} component of the triple. We show that the combined instance is symmetric with respect to every swap up to an error $O(\varepsilon)$ unless $\mathsf{val}(\mathcal{L})>\eta$. This implies that the instance has value at most $2/3+O(\varepsilon)$ hence proving approximation resistance of MAX NBTW.

For MAX BTW and MAX $2t$-SO, we do not require the final symmetrization and instead use a dictatorship test based on a different distribution. Finally, the reduction to MAS is a simple gadget reduction from MAX NBTW. For hardness results of width-two predicates, such gadget reductions presently dominate the scene of classical CSPs and also designates the state of affairs for MAS. As an example, the best-known NP-hard approximation hardness of $16/17+\varepsilon$ for MAX CUT is via a gadget reduction from MAX 3-LIN-2 [16, 23]. The preceding best approximation hardness of MAS was also via a gadget reduction from MAX 3-LIN-2 [19], although with the significantly smaller hardness constant $65/66+\varepsilon$. By reducing from a problem more similar to MAS, namely MAX NBTW, we improve to the approximation hardness to $14/15+\varepsilon$. The gadget in question is simple and was given in Section 1.

Organization. Section 2 sets up the notation used in the rest of the article. Section 3 gives a general hardness result based on test distributions, although the proofs are deferred to Section 5. In Section 4, we introduce distributions tailored to the studied predicates and invoke the construction from Section 3 to prove three of the four approximation-hardness results; the fourth result is proven in the same section via a gadget reduction.

Acknowledgments. We would like to thank Johan Håstad for suggesting the topic and for numerous helpful discussions regarding the same, as well as the anonymous referees, and the editors of the publishing issue, who assisted in improving the legibility of the work.

2 Preliminaries

We denote by $[n]$ the integer interval $\{0,\dots,n-1\}$ and by $[n]_1$ the interval $\{1,\dots,n\}$. Given a tuple of reals $\mathbf{x}\in\mathbb{R}^m$, we write $\vec{\sigma}(\mathbf{x})\in S_m$ for the natural-order permutation on $\{1,\dots,m\}$ induced by $\mathbf{x}$. We also denote by $\mathbf{x}_{-i}$ the tuple $(x_1,\dots,x_{i-1},x_{i+1},\dots,x_m)$. For a distribution $\mathcal{D}$ over $\Omega_1\times\cdots\times\Omega_m$, we use $\mathcal{D}_{\le t}$ and $\mathcal{D}_{>t}$ to denote the projection to coordinates up to t and to the remaining coordinates, respectively.

2.1 Ordering-constraint satisfaction problems

We are concerned with predicates $\mathcal{P}:S_m\to[0,1]$ on the symmetric group S_m. Such a predicate specifies a width-m OCSP written as OCSP($\mathcal{P}$). An instance $\mathcal{I}$ of OCSP($\mathcal{P}$) problem is a tuple $(\mathcal{X},\mathcal{C})$ where $\mathcal{X}$ is the set of *variables* and $\mathcal{C}$ is a distribution over ordered m-tuples of $\mathcal{X}$ referred to as the *constraints*.

An assignment to $\mathcal{I}$ is an injective map $\mathcal{O}:\mathcal{X}\to\mathbb{Z}$ called an ordering of $\mathcal{X}$. For a tuple $\vec{c}=(v_1,\dots,v_m)$, $\mathcal{O}_{|c}$ denotes the tuple $(\mathcal{O}(v_1),\dots,\mathcal{O}(v_m))$. An ordering is said to satisfy the constraint c when $\mathcal{P}(\vec{\sigma}(\mathcal{O}_{|c}))=1$. The value of an ordering is the probability that a random constraint $c\leftarrow\mathcal{C}$ is satisfied by $\mathcal{O}$ and the value $\mathsf{val}(\mathcal{I})$ of an instance is the maximum value of any ordering. Thus,

$$\mathsf{val}(\mathcal{I}) \stackrel{\text{def}}{=} \max_{\mathcal{O}:\mathcal{X}\to\mathbb{Z}} \mathsf{val}(\mathcal{O};\mathcal{I}) \stackrel{\text{def}}{=} \max_{\mathcal{O}:\mathcal{X}\to\mathbb{Z}} \mathop{\mathbf{E}}_{c\in\mathcal{C}} \left[\mathcal{P}(\mathcal{O}_{|c})\right].$$

To simplify analysis, we extend the predicate to $\mathbb{Z}^m$ and permit assignments which are not necessarily injective as follows: define $\mathcal{P} : \mathbb{Z}^m \to [0,1]$ by setting $\mathcal{P}(a_1,\ldots,a_m) = \mathbf{E}_{\vec{\sigma}}[\mathcal{P}(\vec{\sigma})]$ where $\vec{\sigma}$ is drawn uniformly at random over all permutations in S_m such that $\sigma_i < \sigma_j$ whenever $a_i < a_j$. As an example, for $m = 3$,

$$\mathcal{P}(42,13,13) = \frac{1}{2}\mathcal{P}(3,1,2) + \frac{1}{2}\mathcal{P}(3,2,1).$$

Note that the value of an instance is invariant this extension as there is a complete ordering attaining at least the value of any non-injective assignment.

We define the predicates and problems studied in this work. MAS is exactly OCSP($\{(1,2)\}$). The betweenness predicate BTW is the set $\{(1,3,2),(3,1,2)\}$ and NBTW is $S_3 \setminus$ BTW. We define MAX BTW as OCSP(BTW) and MAX NBTW as OCSP(NBTW). Finally, introduce $2t$-SO as the subset of S_{2t} such that the induced ordering on the first t elements coincides with the ordering of the last t elements,

$$2t\text{-SO} \stackrel{\text{def}}{=} \{\pi \in S_{2t} \mid \sigma(\pi(1),\ldots,\pi(t)) = \sigma(\pi(t+1),\ldots,\pi(2t))\}.$$

This predicate has $(2t)!/t!$ satisfying orderings and will be used in proving the inapproximability of the generic m-OCSP with constraints of width at most m.

2.2 Label cover and inapproximability

The problem LABEL COVER (LC) is a common starting point of strong inapproximability results. An LC instance $\mathcal{L} = (U,V,E,L,R,\Pi)$ consists of a bipartite graph $(U\cup V,E)$ associating with every edge $\{u,v\}$ a projection $\pi_{v,u} : R \to L$ with the goal of labeling the vertices, $\lambda : U\cup V \to L\cup R$, so as to maximize the fraction of projections for which $\pi_{v,u}(\lambda(v)) = \lambda(u)$. For simplicity, we shall additionally assume that every projection has *degree* exactly $d = d(\varepsilon)$. That is, for every LC instance, edge $e \in E$, and label $i \in L$, there are exactly d choices of $j \in R$ for which $\pi_e(j) = i$.

Theorem 2.1. *For every $\varepsilon > 0$, there exists fixed label sets L and R such that it is* NP*-hard to distinguish LC instances of value one from instances of value at most ε. Additionally, one can assume that there exists $d = d(\varepsilon)$ such that every projection of every instance has degree d.*

Proof. Follows from an easy modification [24] of the classical LC construction [2, 20]. □

2.3 Primer on real analysis

We refer to a finite domain Ω along with a distribution μ as a probability space. Given a probability space (Ω,μ), the n^{th} *tensor power* is $(\Omega^n,\mu^{\otimes n})$ where $\mu^{\otimes n}(\omega_1,\ldots,\omega_n) = \mu(\omega_1)\cdots\mu(\omega_n)$. The ℓ_p *norm of* $f : \Omega \to \mathbb{R}$ *w. r. t.* μ is denoted by $\|f\|_{\mu,p}$ and is defined as $\mathbf{E}_{\mathbf{x}\sim\mu}\left[|f(\mathbf{x})|^p\right]^{1/p}$ for real $p \geq 1$ and $\max_{x\in\mathrm{supp}(\mu)} f(x)$ for $p = \infty$. When clear from the context, we omit the distribution μ. The following *noise operator* and its properties play a pivotal role in our analysis.

Definition 2.2. Let (Ω, μ) be a probability space and $f : \Omega^n \to \mathbb{R}$ be a function on the n^{th} tensor power. For a parameter $\rho \in [0,1]$, the *noise operator* T_ρ takes f to $\mathrm{T}_\rho f \to \mathbb{R}$ defined by

$$\mathrm{T}_\rho f(x) \stackrel{\text{def}}{=} \mathbf{E}\left[f(\mathbf{y}) \mid \mathbf{x}\right],$$

where independently for each i, with probability ρ, $\mathbf{y}_i = x_i$ and otherwise a uniform random sample.

The noise operator preserves the mass $\mathbf{E}[f]$ of a function while spreading it in the space. The quantitative bound on this is referred to as hypercontractivity.

Theorem 2.3 ([25]; [18, Theorem 3.16, 3.17])**.** *Let (Ω, μ) be a probability space in which the minimum nonzero probability of any atom is $\alpha < 1/2$. Then, for every $q > 2$ and every $f : \Omega^n \to \mathbb{R}$,*

$$\left\|\mathrm{T}_{\rho(q,\alpha)} f\right\|_q \le \|f\|_2\,,$$

where for $\alpha < 1/2$ we set

$$A = \frac{1-\alpha}{\alpha}\,; \qquad \frac{1}{q'} = 1 - \frac{1}{q}\,; \quad \textit{and} \quad \rho(q,\alpha) = \left(\frac{A^{1/q} - A^{-1/q}}{A^{1/q'} - A^{-1/q'}}\right)^{1/2}.$$

For $\alpha = 1/2$, we set $\rho(q,\alpha) = (q-1)^{-1/2}$.

For a fixed probability space, the above theorem says that for every $\gamma > 0$, there is a $q > 2$ such that $\left\|\mathrm{T}_{1-\gamma} f\right\|_q \le \|f\|_2$. For our application, we need the easy corollary that the reverse direction also holds: for every $\gamma > 0$, there exists a $q > 2$ such that hypercontractivity to the ℓ_2-norm holds.

Lemma 2.4. *Let (Ω, μ) be a probability space in which the minimum nonzero probability of any atom is $\alpha \le 1/2$. Then, for every $f : \Omega^n \to \mathbb{R}$, small enough $\gamma > 0$,*

$$\left\|\mathrm{T}_{1-\gamma} f\right\|_{2+\delta} \le \|f\|_2$$

for any

$$0 < \delta \le \delta(\gamma,\alpha) = 2\frac{\log((1-\gamma)^{-2}) - 1}{\log(A)} \quad \textit{with} \quad A = \frac{1-\alpha}{\alpha} > 1\,.$$

Further, $\delta(\gamma, 1/2) = \gamma(2-\gamma)(1-\gamma)^{-2}$.

Proof. The estimate for $\delta(\gamma, 1/2)$ follows immediately from the above theorem assuming $\gamma < 1/2$. In the case when $\alpha < 1/2$, solving

$$\rho^2 \stackrel{\text{def}}{=} (1-\gamma)^2 = (A^{1/q} - A^{-1/q})(A^{1-1/q} - A^{1/q-1})^{-1}$$

for q gives, for $\gamma < 1 - A^{-1/2}$,

$$\delta = q - 2 = \frac{2\log(A)}{\log\left(\frac{1+\rho^2 A}{1+\rho^2/A}\right)} - 2 \ge 2\frac{\log\left((1-\gamma)^{-2}\right) - 1}{\log(A)}\,. \qquad \square$$

Efron-Stein Decompositions. Our proofs make use of the *Efron-Stein decomposition* which has useful properties akin to Fourier decompositions.

Definition 2.5. Let $f : \Omega^{(1)} \times \cdots \times \Omega^{(n)} \to \mathbb{R}$ and μ a measure on $\prod \Omega^{(t)}$. The *Efron-Stein decomposition* of f with respect to μ is defined as $\{f_S\}_{S \subseteq [n]}$ where

$$f_S(\mathbf{x}) \stackrel{\text{def}}{=} \sum_{T \subseteq S} (-1)^{|S \setminus T|} \mathbf{E}\big[f(\mathbf{x}') \,|\, \mathbf{x}'_T = \mathbf{x}_T\big] .$$

Lemma 2.6 (Efron and Stein [10]; Mossel [18])**.** *Assuming $\{\Omega^{(t)}\}_t$ are independent, the Efron-Stein decomposition $\{f_S\}_S$ of $f : \prod \Omega^{(t)} \to \mathbb{R}$ satisfies:*

- *$f_S(\mathbf{x})$ depends only on $\mathbf{x}_S$,*
- *for any $S, T \subseteq [m]$, and $\mathbf{x}_T \in \prod_{t \in T} \Omega^{(t)}$ such that $S \setminus T \neq \emptyset$, $\mathbf{E}[f_S(\mathbf{x}') \,|\, \mathbf{x}'_T = \mathbf{x}_T] = 0$.*

We use the standard notion of influence and noisy influence as in previous work.

Definition 2.7. Let $f : \Omega^n \to \mathbb{R}$ be a function on a probability space. The *influence* of the $1 \leq i^{\text{th}} \leq n$ coordinate is

$$\mathrm{Inf}_i(f) \stackrel{\text{def}}{=} \mathop{\mathbf{E}}_{\{\Omega_j\}_{j \neq i}} \left[\mathsf{Var}_{\Omega_i}(f)\right] .$$

Additionally, given a noise parameter γ, the *noisy influence* of the i^{th} coordinate is

$$\mathrm{Inf}_i^{(1-\gamma)}(f) \stackrel{\text{def}}{=} \mathop{\mathbf{E}}_{\{\Omega_j\}_{j \neq i}} \left[\mathsf{Var}_{\Omega_i}(\mathrm{T}_{1-\gamma} f)\right] .$$

Influences have simple expressions in terms of Efron-Stein decompositions:

$$\mathrm{Inf}_i(f) = \sum_{S \ni i} \mathbf{E}\left[f_S^2\right] \qquad \text{and} \qquad \mathrm{Inf}_i^{(1-\gamma)}(f) = \sum_{S \ni i} (1-\gamma)^{2|S|} \mathbf{E}\left[f_S^2\right] .$$

The following well-known bounds on noisy influences are handy in the analysis.

Lemma 2.8. *For every $\gamma > 0$, natural numbers i and n such that $1 \leq i \leq n$, and every $f : \Omega^n \to \mathbb{R}$,*

$$\mathrm{Inf}_i^{(1-\gamma)}(f) \leq \mathsf{Var}(f) \qquad \textit{and} \qquad \sum_i \mathrm{Inf}_i^{(1-\gamma)}(f) \leq \frac{\mathsf{Var}(f)}{\gamma} .$$

Proof. The former inequality is immediate from the influences in terms of Efron-Stein and Parseval's Identity for the decomposition: $\sum_S \mathbf{E}\left[f_S^2\right] = \mathsf{Var}(f)$. Expanding the LHS of the second inequality,

$$\begin{aligned} \sum_i \mathrm{Inf}_i^{(1-\gamma)}(f) &= \sum_{k \geq 1} k(1-\gamma)^{2k} \sum_{S : |S| = k} \mathbf{E}\left[F_S^2\right] \leq \mathsf{Var}(f) \max_{k \geq 1} k(1-\gamma)^{2k} \\ &\leq \mathsf{Var}(f) \sum_{k \geq 1} (1-\gamma)^{2k} \leq \mathsf{Var}(f) \frac{1}{(1-\gamma)^2} \leq \frac{\mathsf{Var}(f)}{\gamma} . \end{aligned}$$

□

Let the *total (noisy) influence* of a function $f : \Omega^n \to \mathbb{R}$ be defined as

$$\mathrm{TotInf}(f) \stackrel{\text{def}}{=} \sum_{i\in[n]_1} \mathrm{Inf}_i(f) \qquad \text{and} \qquad \mathrm{TotInf}^{(1-\gamma)}(f) \stackrel{\text{def}}{=} \sum_{i\in[n]_1} \mathrm{Inf}_i^{(1-\gamma)}(f)\,.$$

Finally, we introduce the notion of *cross influence* between functions subject to a projection. This notion is implicitly prevalent in modern LC reductions, either for noised functions or for, analytically similar, degree-bounded functions. Given finite sets Ω, L, R; an injective function $\pi : L \to R$; and functions $f : \Omega^L \to \mathbb{R}; g : \Omega^R \to \mathbb{R}$; define

$$\mathrm{CrInf}_{\pi}^{(1-\gamma)}(f,g) \stackrel{\text{def}}{=} \sum_{(i,j)\in\pi} \mathrm{Inf}_i^{(1-\gamma)}(f)\,\mathrm{Inf}_j^{(1-\gamma)}(g)\,.$$

Note that our definition differs somewhat from the cross-influence, denoted "XInf," used by Samorodnitsky and Trevisan [22].

3 A general hardness result

In this section, we prove a general inapproximability result for OCSPs which, similar to results for classic CSPs, permit us to deduce hardness of approximation based on the existence of certain simple distributions. The proof is via a scheme of reductions from LC to OCSPs. For an m-width predicate $\mathcal{P}$, we instantiate this scheme with a distribution $\mathcal{D}$ over $Q_1^t \times Q_2^{m-t}$; for some parameter $1 \le t < m$ and where Q_1 and Q_2 are finite integer sets. As an example of a classical distribution, Q_1 and Q_2 could be the set $\{-1,1\}$; $t = 1, m = 3$; and $\mathcal{D}$ could choose the two first arguments independently, setting the third to be their product. To get appropriate properties for ordering problems, we may however need Q_1 and Q_2 to be larger ranges than the classical bits, and the sizes of tables in the resulting OCSP instance will similarly depend on these ranges.

When the predicate permits introducing distributions satisfying certain properties, akin to pairwise independence, then the hardness claims in this section follow and, together with the hardness of LC and other properties of the distributions, yield the desired inapproximability results.

The reduction itself is composed of pieces known as dictatorship test which is described in the next section. Section 3.2 uses this test to construct the overall reduction and also contains the properties of this reduction. For concrete examples of distributions $\mathcal{D}$, we refer the reader to Section 4.

3.1 Dictatorship test

The dictatorship test uses the following test distribution (Distribution 1) parametrized by γ, and π and is denoted by $\mathcal{T}_{\pi}^{(\gamma)}(\mathcal{D})$. The reader should think of the support of the argument distribution $\mathcal{D}$ as essentially being contained in the predicate $\mathcal{P}$.

Recall our convention of extending $\mathcal{P}$ to a predicate $\mathcal{P} : \mathbb{Z}^m \to [0,1]$. For a pair of functions (f,g), let $(f,g)\circ(\mathbf{X},\mathbf{Y})$ denote the tuple $(f(\vec{\mathbf{x}}^{(1)}),\dots,f(\vec{\mathbf{x}}^{(t)}),\ g(\vec{\mathbf{y}}^{(t+1)}),\dots,g(\vec{\mathbf{y}}^{(m)}))$. Then, the *acceptance probability* on $\mathcal{T}_{\pi}^{(\gamma)}(\mathcal{D})$ for a pair of functions (f,g) where $f : Q_1^L \to \mathbb{Z}$ and $g : Q_2^R \to \mathbb{Z}$ is

$$\mathsf{Acc}_{f,g}(\mathcal{T}_{\pi}^{(\gamma)}(\mathcal{D})) \stackrel{\text{def}}{=} \mathop{\mathbf{E}}_{(\mathbf{X},\mathbf{Y})\leftarrow\mathcal{T}_{\pi}^{(\gamma)}(\mathcal{D})}\left[\mathcal{P}((f,g)\circ(\mathbf{X},\mathbf{Y}))\right]. \tag{3.1}$$

Distribution 1: $\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})$

Parameters:

- distribution $\mathcal{D}$ over $\overbrace{Q_1 \times \cdots \times Q_1}^{t} \times \overbrace{Q_2 \times \cdots \times Q_2}^{m-t}$;
- noise parameter $\gamma > 0$;
- projection map $\pi : R \to L$.

Output: Distribution $\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})$ over

$$(\vec{\mathbf{x}}^{(1)}, \ldots, \vec{\mathbf{x}}^{(t)}, \vec{\mathbf{y}}^{(t+1)}, \ldots, \vec{\mathbf{y}}^{(m)}) \in \left(Q_1^L \times \cdots \times Q_1^L\right) \times \left(Q_2^R \times \cdots \times Q_2^R\right).$$

1. pick a random $|L| \times t$ matrix $\mathbf{X}$ over Q_1 by letting each row be a sample from $\mathcal{D}_{\leq t}$, independently.
2. pick a random $|R| \times (m-t)$ matrix $\mathbf{Y} \stackrel{\text{def}}{=} (\vec{\mathbf{y}}^{(t+1)}, \ldots, \vec{\mathbf{y}}^{(m)})$ over Q_2 by letting the i^{th} row be a sample from $\mathcal{D}_{>t}$ conditioned on $\mathbf{X}_{\pi(i)} = (\mathbf{x}_{\pi(i)}^{(1)}, \ldots, \mathbf{x}_{\pi(i)}^{(t)})$
3. for each entry of $\mathbf{X}$ and $\mathbf{Y}$, independently with probability γ, replace the entry with a new sample from Q_1 and Q_2, resp.
4. output $(\mathbf{X}, \mathbf{Y})$.

This definition is motivated by the overall reduction described in the next section. The distribution is designed so that functions (f, g) that are dictated by a single coordinate have a high acceptance probability, justifying the name of the test.

Lemma 3.1. *Let $f : Q_1^L \to \mathbb{Z}$ and $g : Q_2^R \to \mathbb{Z}$ be defined by $g(\mathbf{y}) = y_i$ and $f(\mathbf{x}) = x_{\pi(i)}$ for some $i \in R$. Then,*

$$\mathsf{Acc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})) \geq \mathop{\mathbf{E}}_{\vec{\mathbf{x}}^{()} \sim \mathcal{D}}[\mathcal{P}(\mathbf{x})] - \gamma m.$$

Proof. The vector $(f(\vec{\mathbf{x}}^{(1)}), \ldots, f(\vec{\mathbf{x}}^{(t)}), g(\vec{\mathbf{y}}^{(t+1)}), \ldots, g(\vec{\mathbf{y}}^{(m)}))$ simply equals the $\pi(i)^{\text{th}}$ row of $\mathbf{X}$ followed by the i^{th} row of $\mathbf{Y}$. With probability $(1-\gamma)^m \geq 1 - \gamma m$ this is a sample from $\mathcal{D}$ and is hence accepted by $\mathcal{P}$ with probability at least $\mathbf{E}_{\mathbf{x} \sim \mathcal{D}}[\mathcal{P}(\mathbf{x})] - \gamma m$. $\square$

We prove a partial converse of the above: unless f and g have influential coordinates i and j such that $\pi(j) = i$, the distribution $\mathcal{D}$ can be replaced by a product of two distributions with a negligible loss in the acceptance probability. We define this product distribution below and postpone the analysis to Section 5.2 in order to complete the description of the reduction.

Definition 3.2. Given the base distribution $\mathcal{D}$, the *decoupled distribution* $\mathcal{D}^{\perp}$ is obtained by taking independent samples $\mathbf{x}$ from $\mathcal{D}_{\leq t}$ and $\mathbf{y}$ from $\mathcal{D}_{>t}$.

Reduction 2: $R^{(\mathcal{P})}_{\mathcal{D},\gamma}$

Parameters: distribution $\mathcal{D}$ over $Q_1^t \times Q_2^{m-t}$ and noise parameter $\gamma > 0$.

Input: a Label Cover instance $\mathcal{L} = (U, V, E, L, R, \Pi)$.

Output: a weighted OCSP($\mathcal{P}$) instance $\mathcal{I} = (\mathcal{X}, \mathcal{C})$ where $\mathcal{X} = (U \times Q_1^L) \cup (V \times Q_2^R)$. The distribution of constraints in $\mathcal{C}$ is obtained by sampling a random edge $e = (u, v) \in E$ with projection π_e and $(\mathbf{X}, \mathbf{Y})$ from $\mathcal{T}^{(\gamma)}_{\pi_e}(\mathcal{D})$; the constraint is the predicate $\mathcal{P}$ applied to the tuple

$$((u, \vec{\mathbf{x}}^{(1)}), \ldots, (u, \vec{\mathbf{x}}^{(t)}), (v, \vec{\mathbf{y}}^{(t+1)}), \ldots, (v, \vec{\mathbf{y}}^{(m)}))\,.$$

3.2 Reduction from Label Lover

An assignment to $\mathcal{I}$ is seen as a collection of functions, $\{f_u\}_{u\in U} \cup \{g_v\}_{v\in V}$, where $f_u : Q_1^L \to \mathbb{Z}$, $g_v : Q_2^R \to \mathbb{Z}$ and, e. g., $f_u(\mathbf{x})$ is identified with the variable $(u, \mathbf{x})$. The value of an assignment is

$$\mathop{\mathbf{E}}_{\substack{\mathbf{e}=(\mathbf{u},\mathbf{v})\in E;\\ (\mathbf{X},\mathbf{Y})\in\mathcal{T}^{(\gamma)}_{\pi_\mathbf{e}}(\mathcal{D})}} \mathcal{P}\left((f_\mathbf{u}, g_\mathbf{v}) \circ (\mathbf{X}, \mathbf{Y})\right) = \mathop{\mathbf{E}}_{\mathbf{e}=(\mathbf{u},\mathbf{v})\in E} [\mathsf{Acc}_{f_\mathbf{u}, g_\mathbf{v}}(\mathcal{T}^{(\gamma)}_{\pi_\mathbf{e}}(\mathcal{D}))]\,.$$

Lemma 3.1 now implies that if $\mathcal{L}$ is satisfiable, then the value of the output instance is also high.

Lemma 3.3. *If λ is a labeling of $\mathcal{L}$ satisfying a fraction c of its constraints, then the ordering assignment $f_u(\mathbf{x}) = x_{\lambda(u)}$, $g_v(\mathbf{y}) = y_{\lambda(v)}$ satisfies at least a fraction*

$$c \cdot \left(\mathop{\mathbf{E}}_{\mathbf{x}\sim\mathcal{D}}[\mathcal{P}(\mathbf{x})] - \gamma m\right)$$

of the constraints of $R^{(\mathcal{P})}_{\mathcal{D},\gamma}(\mathcal{L})$. In particular, there is an ordering of $R^{(\mathcal{P})}_{\mathcal{D},\gamma}(\mathcal{L})$ which does not depend on $\mathcal{D}$ and attains a value

$$\mathsf{val}(\mathcal{L}) \cdot \left(\mathop{\mathbf{E}}_{\mathbf{x}\sim\mathcal{D}}[\mathcal{P}(\mathbf{x})] - \gamma m\right).$$

On the other hand, we also extend the decoupling property of the dictatorship test to the instance output when $\mathsf{val}(\mathcal{L})$ is small. This is the technical core of the paper and is proven in Section 5, more specifically via Lemma 5.9.

Theorem 3.4. *Suppose that $\mathcal{D}$ over $Q_1^t \times Q_2^{m-t}$ satisfies the following properties:*

- *$\mathcal{D}$ has uniform marginals;*
- *for every $i > t$, $\mathcal{D}_i$ is independent of $\mathcal{D}_{\le t}$.*

Then, for every $\varepsilon > 0$ and $\gamma > 0$ there exists $\varepsilon_{LC} > 0$ such that if $\mathsf{val}(\mathcal{L}) \le \varepsilon_{LC}$ then for every assignment $A = \{f_u\}_{u\in U} \cup \{g_v\}_{v\in V}$ to $\mathcal{I}$ it holds that

$$\mathsf{val}(A; R^{(\mathcal{P})}_{\mathcal{D},\gamma}(\mathcal{L})) \le \mathsf{val}(A; R^{(\mathcal{P})}_{\mathcal{D}^\perp,\gamma}(\mathcal{L})) + \varepsilon\,.$$

In particular,

$$\mathsf{val}(R^{(\mathcal{P})}_{\mathcal{D},\gamma}(\mathcal{L})) \le \mathsf{val}(R^{(\mathcal{P})}_{\mathcal{D}^\perp,\gamma}(\mathcal{L})) + \varepsilon\,.$$

4 Applications of the general result

In this section, we prove the inapproximability of MAX BTW, MAX NBTW, and MAX $2t$-SO, using the general hardness result of Section 3. We also present a gadget reduction from MAX NBTW to MAS.

4.1 Hardness of MAXIMUM BETWEENNESS

For $q \in \mathbb{Z}^{\ge 1}$, define the distribution $\mathcal{D}$ over $\{-1,q\} \times [q] \times [q]$ by picking $\mathbf{x}_1 \sim \{-1,q\}$, $\mathbf{y}_2 \sim [q]$, and

$$\mathbf{y}_3 = \begin{cases} \mathbf{y}_2 + 1 \bmod q & \text{if } \mathbf{x}_1 = q, \\ \mathbf{y}_2 - 1 \bmod q & \text{if } \mathbf{x}_1 = -1. \end{cases}$$

The way to think of this distribution is that $\mathbf{x}_1$ will always take on either the smallest or the largest value which in turn dictates whether $\mathbf{y}_3$ should be smaller or greater than $\mathbf{y}_2$ in order to satisfy BTW, which will be satisfied when drawn as above except perhaps the unlikely $\mathbf{y}_2 \in \{0, q-1\}$.

Proposition 4.1. *Let* $(\mathbf{x}_1, \mathbf{y}_2, \mathbf{y}_3) \sim \mathcal{D}$. *Then the following holds:*

1. $\mathcal{D}$ *has uniform marginals.*
2. $\mathbf{y}_2$ *and* $\mathbf{y}_3$ *are separately independent of* $\mathbf{x}$.
3. $(\mathbf{y}_2, \mathbf{y}_3)$ *has the same distribution as* $(\mathbf{y}_3, \mathbf{y}_2)$.
4. $\mathbf{E}_{\mathbf{x}_1,\mathbf{y}_2,\mathbf{y}_3 \sim \mathcal{D}}[\mathrm{BTW}(\mathbf{x}_1,\mathbf{y}_2,\mathbf{y}_3)] \ge 1 - 1/q$.

Let $\gamma > 0$ be a noise parameter and $\mathcal{D}^\perp$ be the decoupled distribution of $\mathcal{D}$ which draws the first coordinate independently of the remaining. Consider applying Reduction 2 to a given LC instance $\mathcal{L}$ with test distributions $\mathcal{D}$ and $\mathcal{D}^\perp$, obtaining MAX BTW instances

$$\mathcal{I} = R^{\mathrm{BTW}}_{\mathcal{D},\gamma}(\mathcal{L}) \quad \text{and} \quad \mathcal{I}^\perp = R^{\mathrm{BTW}}_{\mathcal{D}^\perp,\gamma}(\mathcal{L})\,.$$

Lemma 4.2 (Completeness)**.** *If* $\mathsf{val}(\mathcal{L}) = 1$, *then* $\mathsf{val}(\mathcal{I}) \ge 1 - 1/q - 3\gamma$.

Proof. This is an immediate corollary of Lemma 3.3 and Proposition 4.1. □

Lemma 4.3 (Soundness)**.** *For every* $\varepsilon > 0$, $\gamma > 0$, q, *there is an* $\varepsilon_{LC} > 0$ *such that if* $\mathsf{val}(\mathcal{L}) \le \varepsilon_{LC}$, *then* $\mathsf{val}(\mathcal{I}) \le 1/2 + \varepsilon$.

Proof. We note that the first two items of Proposition 4.1 asserts that $\mathcal{D}$ satisfies the conditions of Theorem 3.4 and it suffices to show $\mathsf{val}(\mathcal{I}^\perp) \le 1/2$. Let $\{f_u : \{0,1\}^L \to \mathbb{Z}\}_{u \in U}, \{g_v : [q]^R \to \mathbb{Z}\}_{v \in V}$ be an arbitrary assignment to $\mathcal{I}^{\perp}$. Fix an LC edge $\{u,v\}$ with projection π and consider the mean value of constraints produced for this edge by the construction:

$$\mathop{\mathbf{E}}_{\vec{\mathbf{x}}^{(1)},\mathbf{y}^{(2)},\mathbf{y}^{(3)}\leftarrow\mathcal{T}_\pi^{(\gamma)}\left(\mathcal{D}^{\perp}\right)}\left[\mathrm{BTW}\left(f_u(\vec{\mathbf{x}}^{(1)}),g_v(\vec{\mathbf{y}}^{(2)}),g_v(\vec{\mathbf{y}}^{(3)})\right)\right]. \tag{4.1}$$

As noted in Proposition 4.1, $(\vec{\mathbf{y}}^{(2)},\vec{\mathbf{y}}^{(3)})$ has the same distribution as $(\vec{\mathbf{y}}^{(3)},\vec{\mathbf{y}}^{(2)})$ when drawn from $\mathcal{D}$. Consequently, when drawing arguments from the decoupled test distribution, the probability of a specific outcome $(\mathbf{x}^{(1)},\mathbf{y}^{(2)},\mathbf{y}^{(3)})$ equals the probability of $(\mathbf{x}^{(1)},\mathbf{y}^{(3)},\mathbf{y}^{(2)})$. For strict orderings, possibly having broken ties, at most one of the two can satisfy the predicate BTW. Thus, the expression in (4.1), and in effect $\mathsf{val}(\mathcal{I}^{\perp})$, is bounded by $1/2$. $\square$

Theorem 1.2 is now an immediate corollary of Lemmas 4.2 and 4.3, taking $q=\lceil 2/\varepsilon\rceil$ and $\gamma=\varepsilon/6$.

4.2 Hardness of MAXIMUM NON-BETWEENNESS

For an implicit parameter q, define a distribution $\mathcal{D}$ over $[q]^3$ by picking $\mathbf{x}_1,\mathbf{y}_2\sim[q]$ and setting $\mathbf{y}_3=\mathbf{x}_1+\mathbf{y}_2 \bmod q$.

Proposition 4.4. *The distribution $\mathcal{D}$ satisfies the following:*

1. *$\mathcal{D}$ is pairwise independent with uniform marginals,*
2. *and $\mathbf{E}_{\mathbf{x}_1,\mathbf{x}_2,\mathbf{x}_3\sim\mathcal{D}}[\mathrm{NBTW}(\mathbf{x}_1,\mathbf{y}_2,\mathbf{y}_3)]\geq 1-3/q$.*

A straightforward application of the general inapproximability with $t=1$ shows that $\mathbf{x}_1$ is decoupled from $\mathbf{y}_2$ and $\mathbf{y}_3$ unless $\mathsf{val}(\mathcal{L})$ is large. Furthermore, pairwise independence implies that the decoupled distribution is simply the uniform distribution over $[q]^3$. However, this does not suffice to prove approximation resistance and in fact the value could be greater than $2/3$. To see this, note that if $\{f_u\}_{u\in U},\{g_v\}_{v\in V}$ is an ordering of the instance from the reduction, then the first coordinate of every NBTW constraint is a variable of the form $f_u(\cdot)$ while the rest are $g_v(\cdot)$. Thus, ordering the $f_u(\cdot)$ variables in the middle and randomly ordering $g_v(\cdot)$ on both sides satisfies a fraction $3/4$ of the constraints.

To remedy this issue and prove approximation resistance, we permute $\mathcal{D}$ by swapping the last coordinate with each of the remaining coordinates and overlay the instances obtained by the reduction using these respective distributions. More specifically, for $1\leq j\leq 3$, define $\mathcal{D}^{(j)}$ as the distribution over $[q]^3$ obtained by first sampling from $\mathcal{D}$ and then swapping the j^{th} and third coordinate, i. e., the j^{th} coordinate is the sum of the other two which are picked independently at random. Similarly, define $\mathrm{NBTW}^{(j)}$ as the ordering predicate which is true if the j^{th} argument does not lie between the other two, e. g., $\mathrm{NBTW}^{(3)}=\mathrm{NBTW}$.

As in the previous section, take an LC instance $\mathcal{L}$ and consider applying Reduction 2 to $\mathcal{L}$ with the distributions $\{\mathcal{D}^{(j)}\}_{j=1}^3$, and write

$$\mathcal{I}^{(j)}=R_{\mathcal{D}^{(j)},\gamma}^{\mathrm{NBTW}^{(j)}}(\mathcal{L}).$$

Similarly let

$$\mathcal{I}^{\perp(j)}=R_{\mathcal{D}^{\perp(j)},\gamma}^{\mathrm{NBTW}^{(j)}}(\mathcal{L})$$

be the corresponding decoupled instances.

As the distributions $\mathcal{D}^{(j)}$ are over the same domain $[q]^3$, the instances $\mathcal{I}^{(1)}, \mathcal{I}^{(2)}, \mathcal{I}^{(3)}$ are over the same variables. We define a new instance $\mathcal{I}$ over the same variables as the "sum"

$$\frac{1}{3}\sum_{j\in[3]} \mathcal{I}^{(j)},$$

defined by taking all constraints in $\mathcal{I}^{(1)}, \mathcal{I}^{(2)}, \mathcal{I}^{(3)}$ with multiplicities and normalizing their weights by $1/3$. In other words, up to noise, we have an acceptance probability of

$$\mathbf{E}\left[\frac{2}{3}\mathrm{NBTW}(f^{\mathbf{u}}(\mathbf{x}), g^{\mathbf{v}}(\mathbf{y}), g^{\mathbf{v}}(\mathbf{x}^{\pi} +_q \mathbf{y})) + \frac{1}{3}\mathrm{NBTW}(g^{\mathbf{v}}(\mathbf{y}), g^{\mathbf{v}}(\mathbf{x}^{\pi} -_q \mathbf{y}), f^{\mathbf{u}}(\mathbf{x}))\right].$$

Lemma 4.5 (Completeness)**.** *If* $\mathsf{val}(\mathcal{L}) = 1$*, then* $\mathsf{val}(\mathcal{I}) \geq 1 - 3/q - 3\gamma$.

Proof. This is an immediate corollary of Lemma 3.3 and Proposition 4.4. □

Lemma 4.6 (Soundness)**.** *For every* $\varepsilon > 0$*,* $\gamma > 0$*,* q*, there is an* $\varepsilon_{LC} > 0$ *such that if* $\mathsf{val}(\mathcal{L}) \leq \varepsilon_{LC}$*, then* $\mathsf{val}(\mathcal{I}) \leq 2/3 + \varepsilon$.

Proof. Again our goal is to use Theorem 3.4 and we start by bounding $\mathsf{val}(\mathcal{I}^{\perp})$. To do this, note that the decoupled distributions $\{\mathcal{D}^{\perp(j)}\}_j$ are in fact the uniform distributions over $[q]^3$ and in particular do not depend on j. This means that the distributions of variables which $\mathrm{NBTW}^{(j)}$ is applied to in $\mathcal{I}^{\perp(j)}$ is independent of j, e. g., if $\mathcal{I}^{\perp(1)}$ contains the constraint $\mathrm{NBTW}^{(1)}(z_1, z_2, z_3)$ with weight w then $\mathcal{I}^{\perp(2)}$ contains the constraint $\mathrm{NBTW}^{(2)}(z_1, z_2, z_3)$ with the same weight. In other words, $\mathcal{I}^{\perp}$ can be thought of as having constraints of the form

$$\mathop{\mathbf{E}}_{j}\left[\mathrm{NBTW}^{(j)}(z_1, z_2, z_3)\right].$$

It is readily verified that

$$\mathop{\mathbf{E}}_{j}\left[\mathrm{NBTW}^{(j)}(a, b, c)\right] \leq \frac{2}{3}$$

for every a, b, c.

Getting back to the main task of bounding $\mathsf{val}(\mathcal{I})$, fix an arbitrary assignment

$$A = \{f_u : [q]^L \to \mathbb{Z}\}_{u\in U} \cup \{g_v : [q]^R \to \mathbb{Z}\}_{v\in V}$$

of $\mathcal{I}$. By Theorem 3.4, $\mathsf{val}(A; \mathcal{I}^{(j)}) \leq \mathsf{val}(A; \mathcal{I}^{\perp(j)}) + \varepsilon$ for $j \in [3]$. It follows that $\mathsf{val}(A; \mathcal{I}) \leq \mathsf{val}(A; \mathcal{I}^{\perp}) + \varepsilon$ and therefore, since A was arbitrary, it holds that $\mathsf{val}(\mathcal{I}) \leq \mathsf{val}(\mathcal{I}^{\perp}) + \varepsilon \leq 2/3 + \varepsilon$, as desired. □

4.3 Hardness of MAXIMUM ACYCLIC SUBGRAPH

The inapproximability of MAS is via a simple gadget reduction from MAX NBTW. We claim the following properties of the directed graph shown in Fig. 2, defined formally as follows.

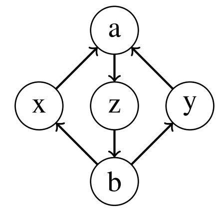

Figure 2: The gadget reducing NBTW to MAS.

Definition 4.7. Define the MAS gadget H from a constraint $\mathrm{NBTW}(x,y,z)$ with auxiliary variables $\{a,b\}$ as the directed graph $H \stackrel{\text{def}}{=} (V,A)$ where $V = \{x,y,z,a,b\}$ and A consists of the walk

$$b \to x \to a \to z \to b \to y \to a\,.$$

Lemma 4.8. *Consider an ordering $\mathcal{O}$ of x, y, z. Then,*

1. *if* $\mathrm{NBTW}(\mathcal{O}(x),\mathcal{O}(y),\mathcal{O}(z)) = 1$, *then* $\max_{\mathcal{O}'} \mathsf{val}(\mathcal{O}';H) = 5/6$ *where the* max *is over all extensions* $\mathcal{O}' : V \to \mathbb{Z}$ *of* $\mathcal{O}$ *to* V.

2. *if* $\mathrm{NBTW}(\mathcal{O}(x),\mathcal{O}(y),\mathcal{O}(z)) = 0$, *then* $\max_{\mathcal{O}'} \mathsf{val}(\mathcal{O}';H) = 4/6$ *where the* max *is over all extensions* $\mathcal{O}'$ *of* $\mathcal{O}$ *to* V.

Proof. To find the value of the gadget H, we individually consider the optimal placement of a and b relative x,y, and z. There are three edges in which the respective variables appear: a appears in (x,a), (y,a) and (a,z); while b appears in (z,b), (b,x), and (b,y).

From this, we gather that two out of the three respective constraints can always be satisfied by placing a after x,y,z and similarly placing b before x,y,z. We also see that all three constraints involving a can be satisfied if and only if z comes after both x and y. Similarly, satisfying all three constraints involving b is possible if and only if z is ordered before both x and y. From this, one concludes that if $\mathrm{NBTW}(x,y,z) = 1$, i. e., if z comes first or last, then we can satisfy five out of the six constraints, whereas if z is the middle element of $\mathcal{O}$, we can satisfy only four out of the six constraints. □

The proof of Theorem 1.1 is now a routine application of the MAS gadget.

Proof of Theorem 1.1. Given an instance $\mathcal{I}$ of MAX NBTW, construct a directed graph G by replacing each constraint $\mathrm{NBTW}(x,y,z)$ of $\mathcal{I}$ with a MAS gadget H, identifying x,y,z with the vertices x,y,z of H and using two new vertices a,b for each constraint of $\mathcal{I}$. By Lemma 4.8, it follows that

$$\mathsf{val}(G) = \frac{5}{6}\mathsf{val}(\mathcal{I}) + \frac{4}{6}(1-\mathsf{val}(\mathcal{I}))\,.$$

By Theorem 1.3, it is NP-hard to distinguish between $\mathsf{val}(\mathcal{I}) \geq 1-\varepsilon$ and $\mathsf{val}(\mathcal{I}) \leq 2/3+\varepsilon$ for every $\varepsilon > 0$, implying that it is NP-hard to distinguish $\mathsf{val}(G) \geq 5/6 - \varepsilon/6$ from $\mathsf{val}(G) \leq 7/9+\varepsilon/6$, providing a hardness gap of

$$\frac{7/9}{5/6} + \varepsilon' = 14/15 + \varepsilon'\,.$$ □

4.4 Hardness of MAXIMUM $2t$-SAME ORDER

We establish the hardness of MAX $2t$-SO, Theorem 1.4, via the approximation resistance of the relatively sparse predicate $2t$-SO. The proof is similar to the that of MAX BTW in Section 4.1. Let $q_1 < q_2$ be integer parameters and define the base distribution $\mathcal{D}$ over $[q_1]^t \times [q_2]^t$ as follows: draw $\mathbf{x}_1, \ldots, \mathbf{x}_t$ uniformly at random from $[q_1]$, draw $\mathbf{z}$ uniformly at random from $[q_2]$, and for $1 \le j \le t$ set $\mathbf{y}_j = \mathbf{x}_j + \mathbf{z} \bmod q_2$. The distribution of $(\mathbf{x}_1, \ldots, \mathbf{x}_t, \mathbf{y}_1, \ldots, \mathbf{y}_t)$ defines $\mathcal{D}$. For a permutation $\vec{\sigma} \in S_t$, let $\mathbf{1}_\sigma(\cdot)$ be the ordering predicate which is 1 on $\vec{\sigma}$ and 0 on all other inputs.

Proposition 4.9. *$\mathcal{D}$ satisfies the following properties.*

1. *$\mathcal{D}$ has uniform marginals.*
2. *For every $i > t$, $\mathcal{D}_i$ is independent of $\mathcal{D}_{\le t}$.*
3. *For every $\vec{\sigma} \in S_t$, $\mathbf{E}\left[\mathbf{1}_\sigma(\mathbf{x}_1, \ldots, \mathbf{x}_t)\right] = \mathbf{E}\left[\mathbf{1}_\sigma(\mathbf{y}_{t+1}, \ldots, \mathbf{y}_m)\right] = 1/t!$.*
4. $\displaystyle \mathop{\mathbf{E}}_{\mathbf{x}_1,\ldots,\mathbf{x}_t,\mathbf{y}_{t+1},\ldots,\mathbf{y}_m \sim \mathcal{D}}\left[2t\text{-SO}(\mathbf{x}_1, \ldots, \mathbf{x}_t, \mathbf{y}_{t+1}, \ldots, \mathbf{y}_m)\right] \ge 1 - \frac{t^2}{2q_1} - \frac{q_1}{q_2}.$

Proof. The first three properties are immediate from the construction and recalling the extension of predicates to non-unique values. For the last property, note that $2t\text{-SO}(\mathbf{x}_1, \ldots, \mathbf{x}_t, \mathbf{y}_{t+1}, \ldots, \mathbf{y}_m) = 1$ if $\mathbf{x}_1, \ldots, \mathbf{x}_t$ are distinct and $\mathbf{z} < q_2 - q_1$. The former event occurs with probability at least $1 - t^2/(2q_1)$ and the latter independently with probability at least $1 - q_1/q_2$; a union bound implies the claim. $\square$

As in the proof of Theorem 1.2, let γ be some noise parameter, take an LC instance $\mathcal{L}$, and let

$$\mathcal{I} = R^{2t\text{-SO}}_{\mathcal{D},\gamma}(\mathcal{L}) \quad \text{and} \quad \mathcal{I}^\perp = R^{2t\text{-SO}}_{\mathcal{D}^\perp,\gamma}(\mathcal{L})$$

be the respective instances produced by Reduction 2 using the base distribution $\mathcal{D}$ and the decoupled version $\mathcal{D}^\perp$.

The following lemma is an immediate corollary of Lemma 3.3, Proposition 4.9, and Item 4.

Lemma 4.10 (Completeness)**.** *If* $\mathsf{val}(\mathcal{L}) = 1$, *then* $\mathsf{val}(\mathcal{I}) \ge 1 - \dfrac{t^2}{2q_1} - \dfrac{q_1}{q_2} - 2t\gamma$.

For the soundness, we have the following.

Lemma 4.11 (Soundness)**.** *For every $\varepsilon > 0, \gamma > 0$, and $1 \le q_1 \le q_2$, there is an $\varepsilon_{LC} > 0$ such that if* $\mathsf{val}(\mathcal{L}) \le \varepsilon_{LC}$, *then* $\mathsf{val}(I) \le 1/t! + \varepsilon$.

Proof. As in the proof of Lemma 4.3, it suffices to prove that $\mathsf{val}(\mathcal{I}^\perp) \le 1/t!$. Let $\{f_u : [q_1]^L \to \mathbb{Z}\}_{u \in U}$, $\{g_v : [q_2]^R \to \mathbb{Z}\}_{v \in V}$ be an arbitrary assignment to $\mathcal{I}^\perp$. Set $m = 2t$. Fix an arbitrary edge $\{u, v\}$ of $\mathcal{L}$ with projection π. The value of constraints corresponding to $\{u,v\}$ satisfied by the assignment is

$$
\begin{aligned}
&\mathop{\mathbf{E}}_{(\mathbf{X},\mathbf{Y})\in\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}^\perp)}\left[2t\text{-SO}(f_u(\vec{\mathbf{x}}^{(1)}),\dots,f_u(\vec{\mathbf{x}}^{(t)}),g_v(\vec{\mathbf{y}}^{(t+1)}),\dots,g_v(\vec{\mathbf{y}}^{(2t)}))\right] \\
&\quad=\sum_{\vec{\sigma}\in S_t}\mathop{\mathbf{E}}_{(\mathbf{X},\mathbf{Y})\in\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}^\perp)}\left[\mathbf{1}_{\vec{\sigma}}\left(f_u(\vec{\mathbf{x}}^{(1)}),\dots,f_u(\vec{\mathbf{x}}^{(t)})\right)\mathbf{1}_{\vec{\sigma}}\left(g_v(\vec{\mathbf{y}}^{(t+1)}),\dots,g_v(\vec{\mathbf{y}}^{(2t)})\right)\right] \\
&\quad=\sum_{\vec{\sigma}\in S_t}\mathop{\mathbf{E}}_{\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}^\perp)}\left[\mathbf{1}_{\vec{\sigma}}\left(f_u(\vec{\mathbf{x}}^{(1)}),\dots,f_u(\vec{\mathbf{x}}^{(t)})\right)\right]\mathop{\mathbf{E}}_{\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}^\perp)}\left[\mathbf{1}_{\vec{\sigma}}\left(g_v(\vec{\mathbf{y}}^{(t+1)}),\dots,g_v(\vec{\mathbf{y}}^{(2t)})\right)\right] \\
&\quad=1/t!\,,
\end{aligned}
$$

where the penultimate step uses the independence of $\mathbf{X}$ and $\mathbf{Y}$ in the decoupled distribution, and the final step uses Item 3 of Proposition 4.9. □

Theorem 1.4 is a corollary of Lemmas 4.10 and 4.11, taking, e. g.,

$$
q_1=\left\lceil\frac{2t^2}{\varepsilon}\right\rceil,\qquad q_2=\left\lceil\frac{2q_1}{\varepsilon}\right\rceil,\quad\text{and}\quad\gamma=\frac{\varepsilon}{8t}\,.
$$

5 Analysis of the reduction

In this section we prove Theorem 3.4 which bounds the value of the instance generated by the reduction in terms of the decoupled distribution. Throughout, we fix an LC instance $\mathcal{L}$, a predicate $\mathcal{P}$, an OCSP instance $\mathcal{I}$ obtained by the procedure $R_{\mathcal{D},\gamma}^{(\mathcal{P})}$ for a distribution $\mathcal{D}$ and noise-parameter γ, and finally an assignment $A=\{f_u\}_{u\in U}\cup\{g_v\}_{v\in V}$.

The proof involves three major steps. First, we show that the assignment functions, which are $\mathbb{Z}$-valued, can be approximated by functions on finite domains via *bucketing* (see Section 5.1). This approximation makes the analyzed instance value susceptible to tools developed in the context of finite-domain CSPs [24, 6] which are used in Section 5.2 to prove the decoupling property of the dictatorship test. Finally, the decoupling is applied to the reduction in Section 5.3 and related to the value of the reduced-from LC instance.

5.1 Bucketing

For an integer Γ, we approximate the possibly non-injective function $f_u: Q_1^L\to\mathbb{Z}$ by partitioning the range into Γ multisets. Set $q_1=|Q_1|$ and partition the cardinality-Q_1^L multiset range of f_u into sets $B_1^{(f_u)},\dots,B_\Gamma^{(f_u)}$ of respective size q_1^L/Γ such that if $x\in B_i^{(f_u)}$ and $y\in B_j^{(f_u)}$ for some $i<j$ then $x\le y$. Note that this is possible as long as the parameter Γ divides q_1^L, which is the case in this treatise. Let $F_u: Q_1^L\to[\Gamma]$ specify the mapping of points to the bucket containing it, and $F_u^{(a)}: Q_1^L\to\{0,1\}$ the indicator of points assigned to $B_a^{(f_u)}$. Partition $g_v: Q_2^R\to\mathbb{Z}$ similarly into buckets $\{B_a^{(g_v)}\}$ obtaining $G_v: Q_2^R\to[\Gamma]$ and $G_v^{(a)}: Q_2^R\to\{0,1\}$.

We show that a bucketed version approximates well the acceptance probability of the dictatorship test for any edge $e = (u,v)$ from Section 3,

$$\mathsf{Acc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})) \stackrel{\text{def}}{=} \mathop{\mathbf{E}}_{(\mathbf{X},\mathbf{Y})\leftarrow\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})}[\mathcal{P}((f,g)\circ(\mathbf{X},\mathbf{Y}))] .$$

Fix an edge $e = (u,v)$ and put $f = f_u$, $g = g_v$. As before, we concisely denote by $(\mathbf{X},\mathbf{Y})$ and $(f,g)\circ(\mathbf{X},\mathbf{Y})$ the query tuple and assignment tuple,

$$(\vec{\mathbf{x}}^{(1)},\dots,\vec{\mathbf{x}}^{(t)},\vec{\mathbf{y}}^{(t+1)},\dots,\vec{\mathbf{y}}^{(m)}) \qquad \text{and} \qquad (f(\vec{\mathbf{x}}^{(1)}),\dots,f(\vec{\mathbf{x}}^{(t)}),g(\vec{\mathbf{y}}^{(t+1)}),\dots,g(\vec{\mathbf{y}}^{(m)})) .$$

Define the *bucketed payoff function* $\wp^{(f,g)} : [\Gamma]^m \to [0,1]$ with respect to f and g as

$$\wp^{(f,g)}(a_1,\dots,a_m) = \mathop{\mathbf{E}}_{\substack{\vec{\mathbf{x}}^{(i)}\leftarrow B_{a_i}^{(f)};\ i\le t\\ \vec{\mathbf{y}}^{(j)}\leftarrow B_{a_j}^{(g)};\ j>t}} \left[\mathcal{P}(\vec{\mathbf{x}}^{(1)},\dots,\vec{\mathbf{x}}^{(t)},\vec{\mathbf{y}}^{(t+1)},\dots,\vec{\mathbf{y}}^{(m)})\right] \tag{5.1}$$

and the *bucketed acceptance probability*,

$$\mathsf{BAcc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})) = \mathop{\mathbf{E}}_{(\mathbf{X},\mathbf{Y})\leftarrow\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})}\left[\wp^{(f,g)}((F,G)\circ(\mathbf{X},\mathbf{Y}))\right] . \tag{5.2}$$

In other words, bucketing corresponds to generating a tuple $\vec{a} = (F,G)\circ(\mathbf{X},\mathbf{Y})$ and replacing each coordinate a_i with a random value from the bucket which a_i fell in. We show that above is close to the true acceptance probability $\mathsf{Acc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}))$.

Theorem 5.1 (Bucketing preserves value). *For every predicate $\mathcal{P}$, noise parameter $\gamma > 0$, projection $\pi : R \to L$, distribution $\mathcal{D}$ with uniform marginals, pair of functions $f : Q_1^L \to \mathbb{Z}$ and $g : Q_2^R \to \mathbb{Z}$, and bucketing parameter Γ,*

$$\left|\mathsf{Acc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})) - \mathsf{BAcc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}))\right| \le m^2\Gamma^{-\delta},$$

for some $\delta = \delta(\gamma,Q) > 0$ with $Q \ge \max\{|Q_1|,|Q_2|\}$.

To prove this, we show that for every choice of f and g, including $f = g$, there is only a few overlapping pairs of buckets, and the probability of hitting any particular pair is small. For each $a \in [\Gamma]$, let $R_a^{(f)}$ be the smallest interval in $\mathbb{Z}$ containing $B_a^{(f)}$; and similarly $R_a^{(g)}$ for g.

Lemma 5.2 (Few buckets overlap). *For every integer Γ there are at most 2Γ choices of pairs $(a,b) \in [\Gamma]\times[\Gamma]$ such that $R_a^{(f)} \cap R_b^{(g)} \ne \emptyset$.*

Proof. Construct the bipartite intersection graph $G_I = (U_I \cup V_I, E_I)$ where the vertex sets are disjoint copies of $[\Gamma]$, and there is an edge $(a,b) \in U_I \times V_I$ iff $R_a^{(f)} \cap R_b^{(g)} \ne \emptyset$. By construction, the intervals $\{R_a^{(f)}\}_a$ are disjoint and similarly $\{R_b^{(g)}\}_b$. Consequently, G_I must be a forest as it does not contain any pair of distinct edges $(u,v),(u',v')$ such that $u < u'$ and $v > v'$. We conclude that that the number of edges and hence intersections is at most $|U_I \cup V_I| = 2\Gamma$. □

Next, we prove a bound on the probability that a fixed pair of the m queries fall in a specific pair of buckets. For a distribution $\mathcal{D}$ over $Q_1^L \times Q_2^R$, define $\mathcal{D}^{(\gamma)}$ as the distribution that samples from $\mathcal{D}$ and for each of the $|L|+|R|$ coordinates independently with probability γ replaces it with a new sample from $\mathcal{D}$. The distribution $\mathcal{D}^{(\gamma)}$ is representative of the projection of $\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})$ to two specific coordinates and we show that noise prevents the buckets from intersecting with good probability.

Lemma 5.3 (Bucket collisions are unlikely). *Let $\mathcal{D}$ be a distribution over $Q_1^L \times Q_2^R$ whose marginals are uniform in Q_1^L and Q_2^R and $\mathcal{D}^{(\gamma)}$ be as defined above. For every integer Γ and every pair of functions $F : Q_1^L \to \{0,1\}$ and $G : Q_2^R \to \{0,1\}$ such that $\mathbf{E}[F(\vec{\mathbf{x}})] = \mathbf{E}[G(\vec{\mathbf{y}})] = 1/\Gamma$,*

$$\mathop{\mathbf{E}}_{(\vec{\mathbf{x}},\vec{\mathbf{y}})\in\mathcal{D}^{(\gamma)}}[F(\vec{\mathbf{x}})G(\vec{\mathbf{y}})] \le \Gamma^{-(1+\delta)}$$

for some $\delta = \delta(\gamma, Q) > 0$ where $Q \ge \min\{|Q_1|,|Q_2|\}$.

Proof. Without loss of generality, let $|Q_1| = \max\{|Q_1|,|Q_2|\}$. Set $q = 2+\delta' > 2$ as in Lemma 2.4, $1/q' = 1-1/q$, and define

$$H(\vec{\mathbf{x}}) \stackrel{\text{def}}{=} \mathop{\mathbf{E}}_{\vec{\mathbf{y}}|\vec{\mathbf{x}}}\left[T_{1-\gamma}G(\vec{\mathbf{y}})\right] .$$

Then,

$$\begin{aligned}\mathop{\mathbf{E}}_{(\vec{\mathbf{x}},\vec{\mathbf{y}})\in\mathcal{D}^{(\gamma)}}[F(\vec{\mathbf{x}})G(\vec{\mathbf{y}})] &= \mathop{\mathbf{E}}_{\vec{\mathbf{x}}}\left[T_{1-\gamma}F(\vec{\mathbf{x}})H(\vec{\mathbf{x}})\right] \le \left\|T_{1-\gamma}F\right\|_q \|H\|_{q'} \\ &\le \|F\|_2\|H\|_{q'} \le \|F\|_2\left\|T_{1-\gamma}G\right\|_{q'} \le \|F\|_2\|G\|_{q'} \\ &= \Gamma^{-(1/2+1/q')} = \Gamma^{-\left(1+\delta'/2(2+\delta')\right)},\end{aligned}$$

using the hypercontractivity from Lemma 2.4 and Jensen's inequality. □

Note that the above lemma applies to queries to the same function as well, setting $F = G$, etc. To complete the proof of Theorem 5.1, we apply the above lemma to every distinct pair of the m queries made in $\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})$ and each of the at most 2Γ pairs of overlapping buckets.

Proof of Theorem 5.1. Note that the bucketed payoff $\wp^{(f,g)}((F,G)\circ(\mathbf{X},\mathbf{Y}))$ is equal to the true payoff $\mathcal{P}((f,g)\circ(\mathbf{X},\mathbf{Y}))$ except possibly when at least two elements in $(F,G)\circ(\mathbf{X},\mathbf{Y})$ fall in an overlapping pair of buckets. Fix a pair of inputs, say $\mathbf{x}^{(i)}$ and $\mathbf{y}^{(j)}$ – the proofs are identical when choosing two inputs from $\mathbf{X}$ or two from $\mathbf{Y}$. Let $a = F(x^{(i)})$ and $b = G(y^{(j)})$. By Lemma 5.2 there are at most 2Γ possible values (a,b) such that the buckets indexed by a and b are overlapping. From Lemma 5.3, the probability that $F(\vec{\mathbf{x}}^{(i)}) = a$ and $G(\vec{\mathbf{y}}^{(j)}) = b$ is at most $\Gamma^{-1-\delta}$. By a union bound, the two outputs $F(\vec{\mathbf{x}}^{(i)})$, $G(\vec{\mathbf{y}}^{(j)})$ consequently fall in overlapping buckets with probability at most $2\Gamma^{-\delta}$. As there are at most $\binom{m}{2} \le m^2/2$ pairs of outputs, the proof is complete. □

5.2 Soundness of the dictatorship test

We now reap the benefits of bucketing and prove the decoupling property of the dictatorship test alluded to in Section 3.

Theorem 5.4. *For every predicate $\mathcal{P}$ and distribution $\mathcal{D}$ satisfying the conditions of Theorem 3.4, and any noise rate $\gamma > 0$, projection $\pi : R \to L$, and bucketing parameter Γ, the following holds. For any functions $f : Q_1^L \to \mathbb{Z}$, $g : Q_2^R \to \mathbb{Z}$ with bucketing functions $F : Q_1^L \to [\Gamma]$, $G : Q_2^R \to [\Gamma]$,*

$$\left|\mathsf{BAcc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})) - \mathsf{BAcc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}^\perp))\right| \leq \gamma^{-1/2} m^{1/2} 4^m \Gamma^m \sum_{a,b\in[\Gamma]} \mathrm{CrInf}_\pi^{(1-\gamma)}\left(F^{(a)}, G^{(b)}\right)^{1/2}.$$

Recall that the decoupled version $\mathcal{D}^\perp$ of a base distribution $\mathcal{D}$ is obtained by combining two independent samples of $\mathcal{D}$: one for the first t coordinates and one for the remaining. A similar claim, as above, for the true acceptance probabilities of the dictatorship test is now a simple corollary of the above theorem and Theorem 5.1. This will be used later in extending the decoupling property to our general inapproximability reduction.

Lemma 5.5. *For every predicate $\mathcal{P}$ and distribution $\mathcal{D}$ satisfying the conditions of Theorem 3.4, and any noise rate $\gamma > 0$, projection $\pi : R \to L$, and bucketing parameter Γ, the following holds. For any functions $f : Q_1^L \to \mathbb{Z}$, $g : Q_2^R \to \mathbb{Z}$ with bucketing functions $F : Q_1^L \to [\Gamma]$, $G : Q_2^R \to [\Gamma]$,*

$$\begin{aligned}&\left|\mathsf{Acc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})) - \mathsf{Acc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}^\perp))\right| \\ &\qquad \leq \gamma^{-1/2} m^{1/2} 4^m \Gamma^m \sum_{a,b\in[\Gamma]} \mathrm{CrInf}_\pi^{(1-\gamma)}\left(F^{(a)}, G^{(b)}\right)^{1/2} + 2\Gamma^{-\delta} m^2 .\end{aligned} \tag{5.3}$$

The proof of the theorem is via an invariance-style theorem and uses techniques found in the works of Mossel [18], Samorodnitsky and Trevisan [22], and Wenner [24]. Frankly, it mostly involves introducing new notation to apply existing machinery. The notion of *lifted* functions may be the most alien: a large-side table $g : [q_2]^R \to \mathbb{R}$ may equivalently be seen as the function $g'^{\pi} : \Omega'^L \to \mathbb{R}$ where $\Omega' = [q_2]^d$ contains the values of all d coordinates in R projecting to the same coordinate in L. g'^{π} is called the lifted analogue of g with respect to the projection π.

The first lemma proved in this section essentially says that if a product of functions is influential, then at least one of the functions is also influential. The second says that if the lifted analogue of a function g is influential for some coordinate $i \in L$, then g is influential for a coordinate projecting to i. The lemma will be used to—after massaging the expression—decouple the small-side table from the lifted analogue of the large-side table as a function of their cross influence.

Lemma 5.6 (Mossel [18, Lemma 6.5])**.** *Let $f_1, \ldots, f_t : \Omega^n \to [0,1]$ be arbitrary functions. Then for any $j \in [n]$, $\mathrm{Inf}_j\left(\prod_{r=1}^t f_r\right) \leq t \sum_{r=1}^t \mathrm{Inf}_j(f_r)$.*

Lemma 5.7. *Consider a function $g : \Omega^R \to \mathbb{R}$, a map $\pi : R \to L$, and an arbitrary bijection $\varsigma : R \leftrightarrow ((\ell) \times [|\pi^{-1}(\ell)|])_{\ell \in L}$ such that for all (i, j), $\exists_t \varsigma(j) = (i,t)$ iff $\pi(j) = i$. Introduce $\Omega'_\ell \stackrel{\text{def}}{=} \Omega^{|\pi^{-1}(\ell)|}$ and define $g' : \prod_L \Omega'_\ell \to \mathbb{R}$ as $g'(\vec{z}) = g(\vec{y})$ where $y_j = z_{\varsigma(j)}$. Then for every $i \in L$,*

$$\mathrm{Inf}_i\left(g'\right) \leq \sum_{j \in \pi^{-1}(i)} \mathrm{Inf}_j(g) .$$

Proof. Using definitions and the law of total conditional variance,

$$\mathrm{Inf}_i\left(g'\right) = \mathop{\mathbf{E}}_{\Omega'_{-i}}\left[\mathop{\mathrm{Var}}_{\Omega'_i} g'\right] = \mathop{\mathbf{E}}_{\{\Omega_\ell\}_{\ell:\nexists_j \varsigma(i,j)=\ell}}\left[\mathop{\mathrm{Var}}_{\{\Omega_\ell\}_{\ell:\exists_j \varsigma(i,j)=\ell}} g\right] \leq \sum_{\ell:\exists_j \varsigma(i,j)=\ell} \mathop{\mathbf{E}}_{\Omega_{-\ell}}\left[\mathop{\mathrm{Var}}_{\Omega_\ell} g\right] = \sum_{\ell:\exists_j \varsigma(i,j)=\ell} \mathrm{Inf}_\ell\left(g\right). \quad \square$$

Theorem 5.8 (Wenner [24, Theorem 3.21])**.** *Consider functions $\{f^{(r)} \in \mathrm{L}^\infty(\Omega_r^n)\}_{r\in[m]_1}$ on a probability space $\mathcal{P} = (\prod_{i=1}^m \Omega_i, \mathrm{P})^{\otimes n}$, a set $M \subsetneq [m]_1$, and a collection $\mathcal{C}$ of minimal sets*[1] *$C \subseteq [m]_1, C \nsubseteq M$ such that the spaces $\{\Omega_i\}_{i\in C}$ are dependent. Then,*

$$\left|\mathbf{E}\left[\prod_{r=1}^m f^{(r)}\right] - \prod_{r\notin M}\mathbf{E}\left[f^{(r)}\right]\mathbf{E}\left[\prod_{r\in M} f^{(r)}\right]\right| \\ \leq 4^m \max_{C\in\mathcal{C}} \sqrt{\min_{r'\in C} \mathrm{TotInf}\left(f^{(r')}\right) \sum_\ell \prod_{r\in C\setminus\{r'\}} \mathrm{Inf}_\ell\left(f^{(r)}\right)} \prod_{r\notin C} \left|\left|f^{(r)}\right|\right|_\infty.$$

Proof of Theorem 5.4

Proof. We massage the expression $\mathsf{BAcc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}))$ to a form suitable for applying Theorem 5.8. Recall that $F^{(a)}$ denotes the indicator of "$F(\mathbf{x}) = a$" and similarly $G^{(a)}$ of "$G(\mathbf{y}) = a$." In terms of these indicators, $\mathsf{BAcc}_{f,g}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}))$ equals

$$\sum_{\vec{a}\in[\Gamma]^t, \vec{b}\in[\Gamma]^{m-t}} \wp(\vec{a},\vec{b}) \mathop{\mathbf{E}}_{\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})}\left[\prod_{r=1}^t F^{(a_r)}(\mathbf{x}^{(r)}) \prod_{r=t+1}^m G^{(b_{r-t})}(\mathbf{y}^{(r)})\right].$$

Consequently, $\left|\mathsf{BAcc}_{f,g}(\mathcal{D}) - \mathsf{BAcc}_{f,g}(\mathcal{D}^\perp)\right|$ may be bounded from above by

$$\sum_{\vec{a},\vec{b}} \wp(\vec{a},\vec{b}) \left| \mathop{\mathbf{E}}_{\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})}\left[\prod_{r=1}^t F^{(a_r)}(\mathbf{x}^{(r)}) \prod_{r=t+1}^m G^{(b_{r-t})}(\mathbf{y}^{(r)})\right] - \mathop{\mathbf{E}}_{\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}^\perp)}\left[\prod_{r=1}^t F^{(a_r)}(\mathbf{x}^{(r)}) \prod_{r=t+1}^m G^{(b_{r-t})}(\mathbf{y}^{(r)})\right]\right|. \tag{5.4}$$

We note that $0 \leq \wp \leq 1$ and proceed to bound the difference in the summation for fixed (a,b). To this end, we make a slight change of notation as discussed previously. The new notation may seem cumbersome; the high-level picture is that we group the first set of functions into a single function, receiving a single argument over a larger domain, and redefine the latter functions to take arguments indexed by L instead of R. Also recall that we intend to reduce from LC instances with projection degrees exactly d for some $d = d(\varepsilon_{\mathrm{LC}})$.

Define $m' = m - t + 1$, $\Omega_1 = [q_1]^t$, $\Omega_2 = \cdots = \Omega_{m'} = [q_2]^d$. Let ς be a bijection $L \times [d] \leftrightarrow R$ such that $\pi(\varsigma(i,j')) = i$; i. e., for each $i \in L$, we group together the d coordinates in R mapping to i. Introduce the distribution

$$\Omega_1^L \times \cdots \times \Omega_{m'}^L \ni (\vec{\mathbf{w}}, \mathbf{z}^{(2)}, \ldots, \mathbf{z}^{(m')}) \sim \mathcal{R}(\mu)$$

[1] Minimal sets of dependent outcome spaces in the sense that if $C \in \mathcal{C}$, then the outcomes spaces of every strict subset of C are independent.

which samples $(\mathbf{x}^{(1)},\ldots,\mathbf{x}^{(t)},\mathbf{y}^{(t+1)},\ldots,\mathbf{y}^{(m)})$ from $\mathcal{T}_{\pi}^{(\gamma)}(\mathcal{D})$, setting

$$\mathbf{w}_{i,r}=\mathbf{x}_i^{(r)} \quad \text{and} \quad \mathbf{z}_i^{(r)}=(\mathbf{y}^{(r)}_{\varsigma^{-1}(i,j')})_{j'=0}^{d-1}.$$

Let

$$W(\vec{w})\stackrel{\text{def}}{=}\prod_{r=1}^{t}\left(\mathrm{T}_{1-\gamma}F^{(a_r)}\right)(\mathbf{x}^{(r)})$$

where $x_i^{(r)}=w_{i,r}$ and similarly, for $2\le r\le m'$, call the lifted function $H^{(r)}:\Omega_r^L\to\mathbb{R}$ defined as

$$H^{(r)}(\mathbf{z})=\left(\mathrm{T}_{1-\gamma}G^{(b_{r-1})}\right)(\mathbf{y})$$

where $y_{\varsigma(i,j')}=z_{i,j'}$. In the new notation, the difference within the summation in (5.4) is

$$\left|\mathop{\mathbf{E}}_{\mathcal{R}(\mathcal{D})}\left[W(\vec{\mathbf{w}})\prod_{r=2}^{m'}H^{(r)}(\mathbf{z}^{(r)})\right]-\mathop{\mathbf{E}}_{\mathcal{R}(\mathcal{D}^{\perp})}\left[W(\vec{\mathbf{w}})\prod_{r=2}^{m'}H^{(r)}(\mathbf{z}^{(r)})\right]\right|. \tag{5.5}$$

We note that $\mathcal{R}$ is a product distribution $\mathcal{R}=\mu^{\otimes L}$ for some μ and for any $2\le r\le m'$, Ω_r is independent of Ω_1 due to the premise that $\mathcal{D}_r$ independent of $\mathcal{D}_{\le t}$. Choosing $M=\{2,\ldots,m'\}$, minimal indices $\mathcal{C}$ of dependent sets in μ not contained in M contains 1 and at least two elements from M, i. e., $C\in\mathcal{C}$ implies $1,e,e'\in C$ for some $e\neq e'\in\{2,\ldots,m'\}$.

Applying Theorem 5.8 and choosing $r'\neq 1$ bounds the difference (5.5) by

$$4^m\sqrt{\max_{C\in\mathcal{C}}\min_{1\neq e\in C}\mathrm{TotInf}(H^{(e)})\sum_i\mathrm{Inf}_i(W)\prod_{t\in C\setminus\{1,e\}}\mathrm{Inf}_i\left(H^{(t)}\right)\prod_{r\notin C}\left\|H^{(r)}\right\|_\infty}. \tag{5.6}$$

As we assumed the codomain of the studied functions $\{G^{(r)}\}_r$ to be $[0,1]$ the same holds for $\{H^{(r)}\}_r$ and consequently the influences and infinity norms in (5.6) are upper-bounded by one on account of Lemma 2.8, yielding at most

$$(5.6)\le 4^m\left(\max_{e\neq 1}\mathrm{TotInf}(H^{(e)})\cdot\max_{e\neq 1}\sum_i\mathrm{Inf}_i(W)\,\mathrm{Inf}_i\left(H^{(e)}\right)\right)^{1/2}. \tag{5.7}$$

We recall that $W=\prod_{r=1}^{t}\mathrm{T}_{1-\gamma}F^{(a_r)}$ and hence by Lemma 5.6,

$$\mathrm{Inf}_i(W)\le t\sum_{r=1}^{t}\mathrm{Inf}_i\left(\mathrm{T}_{1-\gamma}F^{(a_r)}\right)=t\sum\mathrm{Inf}_i^{(1-\gamma)}\left(F^{(a_r)}\right).$$

Similarly, Lemma 5.7 implies that

$$\mathrm{Inf}_i\left(H^{(e)}\right)\le\sum_{j\in\pi^{-1}(i)}\mathrm{Inf}_j\left(\mathrm{T}_{1-\gamma}G^{(b_{e-1})}\right)=\sum_{j\in\pi^{-1}(i)}\mathrm{Inf}_j^{(1-\gamma)}\left(G^{(b_{e-1})}\right).$$

Returning to Equation (5.7), we have the bound

$$(5.7) \le 4^m \left(t \max_{e>t} \mathrm{TotInf}^{(1-\gamma)}(G^{(b_e)}) \cdot \max_{e>t} \sum_i \sum_{r=1}^{t} \mathrm{Inf}_i^{(1-\gamma)}\left(F^{(a_r)}\right) \sum_{j\in\pi^{-1}(i)} \mathrm{Inf}_j^{(1-\gamma)}\left(G^{(b_e)}\right)\right)^{1/2}. \quad (5.8)$$

Using $t \le m$; $\max_e \cdots \le \sum_e \cdots$ for non-negative expressions; bounding the total noisy influence by γ^{-1} from Lemma 2.8; and identifying the inner sum as a cross influence, we establish the desired bound on the considered difference as

$$(5.5) \le (5.8) \le 4^m \left(t\gamma^{-1} \sum_{r,e} \mathrm{CrInf}_\pi^{(1-\gamma)}\left(F^{(a_r)}, G^{(b_e)}\right)\right)^{1/2} \le \gamma^{-1/2} m^{1/2} 4^m \sum_{r,r'} \mathrm{CrInf}_\pi^{(1-\gamma)}\left(F^{(a_r)}, G^{(b_{r'})}\right)^{1/2}.$$

Finally, as we noted before, since $0 \le \wp \le 1$, and since there are at most Γ^m terms in the summation, (5.4) is at most

$$\gamma^{-1/2} m^{1/2} 4^m \Gamma^m \sum_{a,b\in[\Gamma]} \mathrm{CrInf}_\pi^{(1-\gamma)}\left(F^{(a)}, G^{(b)}\right)^{1/2}. \qquad \square$$

5.3 Soundness of the reduction

With the soundness for the dictatorship test in place, proving the soundness of the reduction (Theorem 3.4) is a relatively standard task of constructing noisy-influence decoding strategies.

The proof follows immediately from the more general estimate given in the following lemma by taking

$$\Gamma = \left\lceil (4m^2/\varepsilon)^{1/\delta} \right\rceil \quad \text{and then} \quad \varepsilon_{LC} = \left(\frac{\varepsilon\gamma^{3/2}}{m^{1/2}4^m\Gamma^{m+1}}\right)^2.$$

Lemma 5.9. *Given an LC instance $\mathcal{L} = (U,V,E,L,R,\Pi)$ and a collection of functions, $f_u : Q_1^L \to \mathbb{Z}$ for $u \in U$; $g_v : Q_2^R \to \mathbb{Z}$ for $v \in V$, and Γ, γ, δ as in this section,*

$$\mathop{\mathbf{E}}_{u,v\sim E}\left[\left|\mathsf{Acc}_{f_u,g_v}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D})) - \mathsf{Acc}_{f_u,g_v}(\mathcal{T}_\pi^{(\gamma)}(\mathcal{D}^\perp))\right|\right] \le \gamma^{-1.5} m^{1/2} 4^m \Gamma^{m+1} \mathsf{val}(\mathcal{L})^{1/2} + 2\Gamma^{-\delta} m^2.$$

Proof. For a function $f : Q_1^L \to \mathbb{Z}$ define a distribution $\Psi(f)$ over L as follows. First pick $a \sim [\Gamma]$ uniformly, then pick $\ell \in L$ with probability $\gamma \cdot \mathrm{Inf}_\ell^{(1-\gamma)}(F_v^{(a)})$ and otherwise an arbitrary label. Note that by Lemma 2.8,

$$\sum_{\ell\in L} \mathrm{Inf}_\ell^{(1-\gamma)}(F_u^{(a)}) \le 1/\gamma$$

and so picking $\ell \in L$ with the given probabilities is possible. Define $\Psi(g)$ over R for $g : Q_2^R \to \mathbb{Z}$ similarly. Now define a labeling of $\mathcal{L}$ by, for each $u \in U$ (and $v \in V$), sampling a label from $\Psi(f_u)$ ($\Psi(g_v)$, resp.), independently.

For an edge $e=(u,v)\in E$, the probability that e is satisfied by the labeling equals $\mathbf{P}(\pi_e(\Psi(f_u))=\Psi(g_v))$, which can be lower-bounded by

$$\sum_{i,j\,:\,\pi_e(j)=i}\gamma^2\mathop{\mathbf{E}}_{a,b\in[\Gamma]}\left[\mathrm{Inf}_i^{(1-\gamma)}\left(F_u^{(a)}\right)\mathrm{Inf}_j^{(1-\gamma)}\left(G_v^{(b)}\right)\right]=(\gamma/\Gamma)^2\sum_{a,b}\mathrm{CrInf}_{\pi_e}^{(1-\gamma)}\left(F_u^{(a)},G_v^{(b)}\right).$$

Taking the expectation over all edges of $\mathcal{L}$, we get that the fraction of satisfied constraints is

$$(\gamma/\Gamma)^2\mathop{\mathbf{E}}_{e=(u,v)}\left[\sum_{a,b}\mathrm{CrInf}_{\pi_e}^{(1-\gamma)}\left(F_u^{(a)},G_v^{(b)}\right)\right]\le\mathsf{val}(\mathcal{L}),$$

and by concavity of the $\sqrt{\cdot}$ function, this implies that

$$\mathop{\mathbf{E}}_{e=(u,v)}\left[\sum_{a,b}\mathrm{CrInf}_{\pi_e}^{(1-\gamma)}\left(F_u^{(a)},G_v^{(b)}\right)^{1/2}\right]\le\Gamma\gamma^{-1}\mathsf{val}(\mathcal{L})^{1/2}.$$

Plugging this bound on the total cross influence into the soundness for the test, Lemma 5.5, we obtain

$$\mathop{\mathbf{E}}_{e=(u,v)}\left[\left|\mathsf{Acc}_{f_u,g_v}(\mathcal{T}_{\pi_e}^{(\gamma)}(\mathcal{D}))-\mathsf{Acc}_{f_u,g_v}(\mathcal{T}_{\pi_e}^{(\gamma)}(\mathcal{D}^{\perp}))\right|\right]\le\gamma^{-1.5}m^{1/2}4^m\Gamma^{m+1}\mathsf{val}(\mathcal{L})^{1/2}+2\Gamma^{-\delta}m^2,$$

as desired. □

References

[1] SANJEEV ARORA, BOAZ BARAK, AND DAVID STEURER: Subexponential algorithms for Unique Games and related problems. In *Proc. 51st FOCS*, pp. 563–572. IEEE Comp. Soc. Press, 2010. [doi:10.1109/FOCS.2010.59] 258

[2] SANJEEV ARORA, CARSTEN LUND, RAJEEV MOTWANI, MADHU SUDAN, AND MARIO SZEGEDY: Proof verification and the hardness of approximation problems. *J. ACM*, 45(3):501–555, 1998. Versions in ECCC and FOCS'92. [doi:10.1145/278298.278306] 260, 262

[3] SANJEEV ARORA AND SHMUEL SAFRA: Probabilistic checking of proofs: A new characterization of NP. *J. ACM*, 45(1):70–122, 1998. Preliminary version in FOCS'92. [doi:10.1145/273865.273901] 260

[4] PER AUSTRIN, RAJSEKAR MANOKARAN, AND CENNY WENNER: On the NP-hardness of approximating Ordering Constraint Satisfaction Problems. In *Proc. 16th Internat. Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'13)*, pp. 26–41, 2013. [doi:10.1007/978-3-642-40328-6_3] 257

[5] BOAZ BARAK, FERNANDO G. S. L. BRANDÃO, ARAM WETTROTH HARROW, JONATHAN A. KELNER, DAVID STEURER, AND YUAN ZHOU: Hypercontractivity, sum-of-squares proofs, and their applications. In *Proc. 44th STOC*, pp. 307–326. ACM Press, 2012. [doi:10.1145/2213977.2214006, arXiv:1205.4484] 258

[6] SIU ON CHAN: Approximation resistance from pairwise independent subgroups. In *Proc. 45th STOC*, pp. 447–456. ACM Press, 2013. Version in ECCC. [doi:10.1145/2488608.2488665] 258, 273

[7] MOSES CHARIKAR, VENKATESAN GURUSWAMI, AND RAJSEKAR MANOKARAN: Every permutation CSP of arity 3 is approximation resistant. In *Proc. 24th IEEE Conf. on Computational Complexity (CCC'09)*, pp. 62–73, 2009. [doi:10.1109/CCC.2009.29] 259

[8] MOSES CHARIKAR, KONSTANTIN MAKARYCHEV, AND YURY MAKARYCHEV: On the advantage over random for Maximum Acyclic Subgraph. In *Proc. 48th FOCS*, pp. 625–633. IEEE Comp. Soc. Press, 2007. Version in ECCC. [doi:10.1109/FOCS.2007.65] 259

[9] BENNY CHOR AND MADHU SUDAN: A geometric approach to betweenness. *SIAM J. Discrete Math.*, 11(4):511–523, 1998. Preliminary version in ESA'95. [doi:10.1137/S0895480195296221] 258, 259

[10] BRADLEY EFRON AND CHARLES STEIN: The Jackknife Estimate of Variance. *Annals of Statistics*, 9(3):586–596, 1981. [doi:10.1214/aos/1176345462] 264

[11] LARS ENGEBRETSEN AND JONAS HOLMERIN: More efficient queries in PCPs for NP and improved approximation hardness of maximum CSP. *Random Struct. Algorithms*, 33(4):497–514, 2008. Preliminary version in STACS'05. [doi:10.1002/rsa.20226] 258

[12] MICHAEL R. GAREY AND DAVID S. JOHNSON: *Computers and Intractability: A Guide to the Theory of* NP*-Completeness*. W. H. Freeman & Co., 1979. 258

[13] VENKATESAN GURUSWAMI, JOHAN HÅSTAD, RAJSEKAR MANOKARAN, PRASAD RAGHAVENDRA, AND MOSES CHARIKAR: Beating the random ordering is hard: Every ordering CSP is approximation resistant. *SIAM J. Comput.*, 40(3):878–914, 2011. Version in ECCC. [doi:10.1137/090756144] 258, 259, 260

[14] VENKATESAN GURUSWAMI, RAJSEKAR MANOKARAN, AND PRASAD RAGHAVENDRA: Beating the random ordering is hard: Inapproximability of Maximum Acyclic Subgraph. In *Proc. 49th FOCS*, pp. 573–582. IEEE Comp. Soc. Press, 2008. [doi:10.1109/FOCS.2008.51] 258, 259, 260

[15] VENKATESAN GURUSWAMI, PRASAD RAGHAVENDRA, RISHI SAKET, AND YI WU: Bypassing UGC from some optimal geometric inapproximability results. In *Proc. 23rd Ann. ACM-SIAM Symp. on Discrete Algorithms (SODA'12)*, pp. 699–717. ACM Press, 2012. Available on ACM DL. Preliminary version in ECCC. 258

[16] JOHAN HÅSTAD: Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001. Preliminary version in STOC'97. [doi:10.1145/502090.502098] 258, 261

[17] SUBHASH KHOT: On the power of unique 2-prover 1-round games. In *Proc. 34th STOC*, pp. 767–775. ACM Press, 2002. Also in CCC'02. [doi:10.1145/509907.510017] 258

[18] ELCHANAN MOSSEL: Gaussian bounds for noise correlation of functions. *Geometric and Functional Analysis*, 19(6):1713–1756, 2010. Preliminary version in FOCS'08. [doi:10.1007/s00039-010-0047-x] 263, 264, 276

[19] ALANTHA NEWMAN: The Maximum Acyclic Subgraph Problem and degree-3 graphs. In *Proc. 4th Internat. Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'01)*, pp. 147–158. Springer, 2001. [doi:10.1007/3-540-44666-4_18] 258, 259, 261

[20] RAN RAZ: A parallel repetition theorem. *SIAM J. Comput.*, 27(3):763–803, 1998. Preliminary version in STOC'95. [doi:10.1137/S0097539795280895] 260, 262

[21] ALEX SAMORODNITSKY AND LUCA TREVISAN: A PCP characterization of NP with optimal amortized query complexity. In *Proc. 32nd STOC*, pp. 191–199. ACM Press, 2000. [doi:10.1145/335305.335329] 258

[22] ALEX SAMORODNITSKY AND LUCA TREVISAN: Gowers uniformity, influence of variables, and PCPs. *SIAM J. Comput.*, 39(1):323–360, 2009. Preliminary version in STOC'06. [doi:10.1137/070681612] 265, 276

[23] LUCA TREVISAN, GREGORY B. SORKIN, MADHU SUDAN, AND DAVID P. WILLIAMSON: Gadgets, approximation, and linear programming. *SIAM J. Comput.*, 29(6):2074–2097, 2000. Preliminary version in FOCS'96. [doi:10.1137/S0097539797328847] 261

[24] CENNY WENNER: Circumventing d-to-1 for approximation resistance of satisfiable predicates strictly containing parity of width at least four. *Theory of Computing*, 9(23):703–757, 2013. Preliminary version in APPROX'12. [doi:10.4086/toc.2013.v009a023] 262, 273, 276, 277

[25] PAWEŁ WOLFF: Hypercontractivity of simple random variables. *Studia Mathematica*, 180(3):219–236, 2007. [doi:10.4064/sm180-3-3] 263

AUTHORS

Per Austrin
associate professor
KTH – Royal Institute of Technology, Stockholm, Sweden
austrin@kth.se
`http://www.csc.kth.se/~austrin`

Rajsekar Manokaran
postdoc
KTH – Royal Institute of Technology, Stockholm, Sweden
rajsekar@csc.kth.se
`http://www.csc.kth.se/~rajsekar`

Cenny Wenner
Ph. D. student
KTH – Royal Institute of Technology, Stockholm, Sweden
cenny@cwenner.net
http://www.cwenner.net

ABOUT THE AUTHORS

PER AUSTRIN graduated from KTH – Royal Institute of Technology in Sweden, 2008. His advisor was Johan Håstad. Following this, he did post-docs at Institut Mittag-Leffler, New York University, the University of Toronto, and Aalto University. In his spare time he enjoys writing tepid biographic blurbs.

RAJSEKAR MANOKARAN received his Ph. D. from Princeton University under the supervision of Sanjeev Arora. He is currently a postdoctoral researcher with Johan Håstad at KTH in Stockholm.

CENNY WENNER is a Ph. D. student at Stockholm University and KTH – Royal Institute of Technology advised by Johan Håstad and Viggo Kann. His research area is Hardness of Approximation and with a particular interest for extending PCP machinery to problems which seem to present fundamental technical obstacles to the same. The author's spare time is occupied by activities such as playing Go, Mahjong, programming, volunteering, and daydreaming about making a difference in the world when he grows up.