Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 66 (2024)

HOW RANDOM CSPS FOOL HIERARCHIES

SIU ON CHAN*, HIU TSUN NG†, AND SIJIN PENG‡

Abstract. Relaxations for the constraint satisfaction problem (CSP) include bounded width, linear program (LP), semidefinite program (SDP), affine integer program (AIP), and the combined LP+AIP of Brakensiek, Guruswami, Wrochna, and Živný (SICOMP 2020). Tightening relaxations systematically leads to hierarchies and stronger algorithms. For the LP+AIP hierarchy, a constant level lower bound for approximate graph coloring was given by Ciardo and Živný (STOC 2023).

We prove the first linear (and hence optimal) level lower bound for LP+AIP and its stronger variant, SDP+AIP. For each hierarchy, our bound holds for random instances of a broad class of CSPs that we call τ-*wise neutral*. We extend to other hierarchies the LP lower bound techniques in Benabbas, Georgiou, Magen and Tulsiani (ToC 2012) and Kothari, Mori, O'Donnell, and Witmer (STOC 2017), and simplify the SDP solution construction in the latter.

1. Introduction

Promise constraint satisfaction problem (PCSP), originated from [AGH17], is a variant of constraint satisfaction problem. In PCSP, given the promise that a CSP instance has a solution satisfying all the constraints, an algorithm is only required to find a solution satisfying all the constraints after each constraint is weakened; see [KO22] for a survey. An example is C-vs-K graph coloring ($C \leqslant K$), where an input graph G is guaranteed to have a proper C-coloring, and our algorithm only needs to find a proper K-coloring of G. The complexity of this problem when $C = 3$ and $K = 6$ is still open. 3-vs-K graph coloring was conjectured NP-hard for every constant $K \geqslant 3$.[1] This conjecture holds under the d-to-1 conjecture [GS20]. C-vs-K graph coloring is NP-hard for every constant $C \geqslant 4$ and $K = \binom{C}{\lfloor C/2 \rfloor} - 1$ [KOWŽ23].

Brakensiek, Guruswami, Wrochna, and Živný [BG19, BGWŽ20] proposed the combined linear and affine integer program (LP+AIP) to unify many algorithms for (P)CSP.[2] LP+AIP solves all tractable Boolean CSPs and is quite powerful, leading [BGWŽ20, KO22] to ask if LP+AIP strengthened into a hierarchy solves all tractable CSPs. A hierarchy is a relaxation systematically tightened based on small variable subsets of an instance; level d of a hierarchy is tightened by constraints induced on subsets of variables of size at most d. Hierarchies for LP and SDP include Sherali–Adams, Lovász–Schrijver, Grigoriev, Lasserre, and Parrilo (e.g. [Lau09]).

For the LP+AIP hierarchy, Ciardo and Živný [CŽ23b] showed that constant level fails to solve 3-vs-K graph coloring, for any constant K. This was the only lower bound known for the LP+AIP hierarchy, leaving open whether sublinear-level LP+AIP solves C-vs-K graph coloring for every $C \leqslant K$, or other (P)CSPs. Since the level-d hierarchy can be solved in time roughly $n^{O(d)}$, sublinear-level LP+AIP might still solve every (P)CSP in subexponential time.

In this paper, we give the first linear (and optimal) level lower bound for the LP+AIP hierarchy, as well as the first average case lower bound for LP+AIP (with or without the hierarchy). Our

*sochan@gmail.com.

†Chinese University of Hong Kong. 1155144015@link.cuhk.edu.hk.

‡Tsinghua University. peng-sj21@mails.tsinghua.edu.cn.

[1] This conjecture is a special case of [BG21, Conjecture 1.2] and was misattributed to [GJ76] in [CŽ23b]. [GJ76] conjectured hardness for arbitrarily large K/C without specifying C, which was proved in [LY94].

[2] They called it BLP+AIP, where the B stands for "basic".

ISSN 1433-8092

lower bound applies to a broad class of CSPs different from graph coloring, by generalizing lower bound techniques from other hierarchies. Here are some hierarchies of interest to us:

- Bounded width hierarchy (BW) in the local consistency algorithm, which is "dual" to bounded-width resolution refutation.[3]
- Sherali–Adams linear programming (LP) hierarchy.
- Lasserre semidefinite programming (SDP) hierarchy, which is the Lagrangian dual of the sum-of-squares SDP hierarchy.
- Affine integer programming (AIP) hierarchy.

Barto, Bulín, Krokhin, and Opršal [BBKO21, Section 7] pointed out the similarity of the BW, LP, and AIP relaxations in terms of minions (i.e. minor-closed families of functions) representing the relaxed values and their homomorphisms. Ciardo and Živný [CŽ23c] gave a *common description* of the four hierarchies in terms of minions and tensors. This paper goes in a different direction, giving a framework to prove *lower bounds* for these hierarchies that are optimal in terms of level.

Previously, Benabbas, Georgiou, Magen and Tulsiani [BGMT12] and Kothari, Mori, O'Donnell, and Witmer [KMOW17] gave (optimal) linear level lower bounds for the LP and SDP hierarchies respectively. We generalize their techniques to other hierarchies, including BW and AIP. They considered CSPs that are τ*-wise uniform*, a notion introduced by Austrin and Mossel [AM09] in the context of hardness of approximation. A CSP is τ-wise uniform if every constraint C has a distribution $\eta(C)$ of satisfying assignments, such that the marginal given $\eta(C)$ of every subset of τ variables is uniform. For example, the CSPs k-SAT and k-XOR are both $(k-1)$-wise uniform, because the satisfying assignments of their constraints support the uniform distribution over assignments of even (or odd) parity, which is a $(k-1)$-wise uniform distribution. Inspired by τ-wise uniformity, we introduce an analogous property for each hierarchy: τ*-wise neutrality* (Definition 4.3). Our first main result, proved in Section 8.2, is that most instances of any τ-wise neutral CSP have linear-level hierarchy solutions.

Theorem 1.1. *Let $\tau \geqslant 2$. For each elementary hierarchy (BW, LP, SDP, and AIP), if a k-CSP is τ-wise neutral for the hierarchy, then except with probability $o_{n;k}(1)$, a random instance of the CSP with n variables and Δn constraints has a hierarchy solution of level $\Omega_k(n/(\Delta^{2/(\tau-1)} \log \Delta))$.*

Most instances at large constraint density Δ are far from satisfiable, unless the CSP is trivially satisfiable by a single value (Lemma B.3). Yet Theorem 1.1 shows that many hierarchies fail to certify these instances as unsatisfiable and are thus *fooled*.

While the LP and SDP parts of Theorem 1.1 were already known, the BW and AIP parts are new. Previously BW/resolution hierarchy lower bounds for a broad class of CSPs were known only for $\tau = 2$ as a special case of [CM13, Theorem 1.4(a)], which has a worse dependence on Δ when $\tau > 2$. The AIP part of Theorem 1.1 is also new. Part of our contribution is identifying a generalization of τ-wise uniformity to AIP. τ-wise neutrality for AIP turns out to be fairly general and has a simple sufficient condition: Satisfying assignments include a Hamming ball of radius τ (Proposition 4.5).

Consider LP+AIP lower bounds for k-SAT. Because k-SAT is $(k-2)$-wise neutral for LP+AIP (Theorem 8.21), Theorem 1.1 implies most k-SAT instances fool LP+AIP whenever $k \geqslant 4$. But not for 3-SAT, the most important CSP in complexity theory. The LP+AIP algorithm looks for distributions $\eta(C)$ of satisfying assignments to constraints C, and affine integer weights $w(C) : \mathrm{supp}(\eta(C)) \to \mathbb{Z}$ supported on the support of $\eta(C)$, so that they both have consistent marginals at variables. Previous works on LP and SDP hierarchies [Gri01, Sch08, Tul09, BGMT12, TW13, Cha16, BCK15, MW16, KMOW17] constructed solutions whose distribution $\eta(C)$ in each constraint is the given τ-wise uniform distribution. The only pairwise uniform distribution supported on 3-SAT is the uniform distribution on 3-XOR (Proposition A.11). In other words, previous LP and SDP hierarchy solutions

[3] "Dual" means that the BW hierarchy has no solution if and only if bounded-width resolution refutation exists [AD08, Lemmas 2 and 3].

strengthen every 3-SAT constraint into 3-XOR. However, the AIP algorithm can refute every unsatisfiable 3-XOR instance [BG19, BBKO21].

We overcome this hurdle and construct new LP and SDP solutions that have *full support*. As explained in Section 1.2, this allows us to construct LP solution and AIP solution separately, which can be combined into an LP+AIP solution. Our second main result, proved in Section 12, gives LP+AIP lower bounds for a broad class of CSPs that include 3-SAT. In fact our lower bounds also hold for the stronger SDP+AIP hierarchy, which we introduce to generalize all the aforementioned hierarchies.

Theorem 1.2. *Let $\tau \geqslant 2$. If a k-CSP is τ-wise neutral for both SDP and AIP (separately), then except with probability $o_{n;k}(1)$, a random instance of the CSP with n variables and Δn constraints has an SDP+AIP hierarchy solution of level $\Omega_k(n/(\Delta^{2/(\tau-1)} \log \Delta))$.*

Remark 1.3. Our lower bound of level in Theorems 1.1 and 1.2 has the same dependency on n and Δ as in [KMOW17, Theorem 7.1]. In particular, our lower bound matches the SDP level upper bounds for CSPs of Boolean predicates [AOW15, dT23] up to constants, for every $\tau \geqslant 2$ and every $\Delta = \Delta(n)$. See [KMOW17, Section 1.5] for a discussion on the optimality of this bound.

Feige's Hypothesis [Fei02, Hypothesis 1] says no sound polynomial-time algorithm refutes most 3-SAT instances at any constant constraint density Δ (a refutation algorithm is *sound* if declares every satisfiable instance to be satisfiable). Many such algorithms, such as constant-level resolution [BSW01] and SDP [Sch08], cannot refute most 3-SAT instances for any $\Delta(n) \leqslant n^{1/2-o(1)}$, let alone constant Δ. It was unknown whether LP+AIP (with or without the hierarchy) refutes most 3-SAT instances at constant Δ; Theorem 1.2 now rules out this possibility. A similar Random Exponential Time Hypothesis of Razenshteyn, Song, and Woodruff [RSW16, Assumption 1.3] states that sound algorithms refuting at least half of 4-SAT instances at constant Δ must take $\exp(\Omega_\Delta(n))$ time. Again, Theorem 1.2 lends weight to RSW's Hypothesis, by ruling out sublinear-level LP+AIP that runs in subexponential time. See [KMOW17, Section 1.2] and [RSW16] for the importance of these hypotheses and their applications to cryptography and learning theory.

Surprisingly, Feige, Kim, and Ofek [FKO06] showed that polynomial-time-verifiable refutation exists for most 3-SAT instances when $\Delta(n) = O(n^{0.4})$, even though no polynomial-time algorithm is known to find their refutation. They combined spectral strong refutation for 3-NAE and parity strong refutation for 3-XOR. Are there better combinations of spectral/SDP and parity argument that disprove Feige's Hypothesis? We consider such a refutation based on the SDP+AIP hierarchy. Theorem 1.2 shows that constant-level SDP+AIP poses no harm to Feige's Hypothesis, and is fooled by most 3-SAT instances whenever $\Delta(n) \leqslant n^{1/2-o(1)}$, just like the SDP hierarchy. This shows Feige–Kim–Ofek refutation cannot be captured by constant-level SDP+AIP.

Algorithms for CSPs, independently by Bulatov [Bul17, Bul20] and by Zhuk [Zhu20], are intricate combinations of local consistency and Gaussian elimination, the two main techniques for CSPs. LP+AIP is a much simpler combination.

Question 1.4. Can constant-level LP+AIP solve all tractable (i.e. polynomial-time solvable) CSPs?

Question 1.4 was asked in [BGWŽ20, Section 6] and [KO22, Section 3]. Similar questions were also asked for two related relaxations: CLAP in [CŽ22] and cohomological k-consistency in [OC22, Question 13]. See also a related conjecture [DO23, Conjecture 4.10] about $\mathbb{Z}$-affine k-consistency. Little was known about the limitation of combined relaxations and hierarchies; the only LP+AIP hierarchy lower bound was [CŽ23b] for approximate graph coloring, a likely intractable problem due to d-to-1 hardness. Theorem 1.2 may make progress on Question 1.4, by giving LP+AIP hierarchy lower bounds to many CSPs, including plausibly tractable ones:

Question 1.5. Is there a tractable CSP pairwise neutral for both SDP and AIP but not trivially satisfiable by a single value?

If the answer to Question 1.5 is Yes, then the answer to Question 1.4 is No, by Theorem 1.2 and Lemma B.3. Even if the answer to Question 1.5 turns out to be No, we believe the question itself is a fruitful step towards understanding the power and limitation of combined hierarchies.

Remark 1.6. Our results also apply to PCSP, including many cases whose computational complexity is open. Consider the PCSP($\mathbf{A}, \mathbf{B}$) with relational structures $\mathbf{A}, \mathbf{B}$ such that $\mathbf{A} \to \mathbf{B}$; see [KO22] for the definitions. Suppose CSP($\mathbf{A}$) is pairwise neutral for both SDP and AIP (see Lemma C.4 for an interesting example), and CSP($\mathbf{B}$) is not trivially satisfiable by a single value. Then most instances of this PCSP have a linear level SDP+AIP solution for $\mathbf{A}$ without being $\mathbf{B}$-satisfiable, by Theorem 1.2 and Lemma B.3.

Remark 1.7. Concurrently, Ciardo and Živný [CŽ24] propose a new relaxation called SDA combining SDP and AIP. When strengthened into a hierarchy, SDA is equivalent to the SDP+AIP hierarchy introduced in this paper, up to a factor of two in level. [CŽ24] prove worst-case lower bounds to the basic SDP+AIP relaxation for approximate graph coloring and approximate graph homomorphism, which are weaker than level-2 SDP+AIP lower bounds for these problems. Their results and ours are incomparable:

- Their lower bounds are for the basic relaxation only. Ours are for linear level of the hierarchy and is optimal.
- Their lower bounds are for the worst case. Ours are average-case, and hence also worst-case.
- Our lower bounds apply to a class of CSPs that excludes approximate graph coloring or approximate graph homomorphism. Via reduction (Appendix C), Theorem 1.2 implicitly implies lower bounds for C-vs-K graph coloring at other parameter regimes than theirs.

1.1. Proof overview: First main theorem.

We lay a common framework for proving lower bounds for various hierarchies. The framework consists of the following high-level steps: For each hierarchy,

(A) Identify a subinstance W_S for each small subset S of variables. Further show that W_S is sparse and small on random instances.
(B) Identify a property shared by certain "nice" CSPs, so that C becomes "invisible" when restricted to any subset of τ variables of a constraint C.
(C) Using the property in (B), construct a local relaxed solution to W_S for every S.
(D) Again using the property in (B), combine the local solutions in (C) into a hierarchy solution.

We now illustrate the above four steps using LP as an example.

Building on [BGMT12, TW13, MW16], Kothari, Mori, O'Donnell, and Witmer [KMOW17] used *closure* (Definitions 5.1 and 5.5) to track constraints in an instance that may affect the local LP solution to a subset S of variables. The closure of S is the subinstance W_S in step (A) above.

Example 1.8. Consider a 3-XOR instance I with two constraints: an even parity constraint C_1 on $\{v_1, v_2, v_3\}$ and an odd parity constraint C_2 on $\{v_3, v_4, v_5\}$. Naturally, the local distribution $\eta(C)$ on each constraint C is the uniform distribution conditioned satisfying the constraint, e.g. $\eta(C_1)$ is the uniform distribution over even parity assignments on $\{v_1, v_2, v_3\}$. As in previous works, the local LP solution to a small and sparse subinstance J is the canonical distribution (Definition 10.2), where an assignment is independently drawn from a distribution $\eta(C)$ for each constraint C in J, conditioned on agreement at common variables. The canonical distribution on $\{C_1, C_2\}$ induces a nonuniform marginal distribution on $S \stackrel{\text{def}}{=} \{v_1, v_2, v_4, v_5\}$, namely the uniform distribution over odd assignments to S. Therefore C_1 and C_2 are included in the closure of S, as they affect the local distribution on S, even though they involve variable(s) outside S.

Example 1.9. Continuing with the previous example, add an extra even parity constraint C_3 on $\{v_5, v_6, v_7\}$ to I. The canonical distribution on $\{C_1, C_2, C_3\}$ induces the same marginal distribution

on $S = \{v_1, v_2, v_4, v_5\}$ as before. Therefore C_3 plays no role in the local distribution on S and is excluded from the closure of S.

As in Example 1.9, if a subinstance J has a constraint C having too many *boundary variables* (i.e. variables belonging to C but not other constraints in J) outside S, C does not belong to the closure of S. [KMOW17] defined closure to capture this intuition.

For step (B), the property of shared by a class of nice CSPs was identified in [BGMT12] to be pairwise uniform (and generalized to τ-wise uniform in [KMOW17]).

For steps (C) and (D), constructing a hierarchy solution amounts to a *scheme* σ of assigning a distribution of satisfying assignments to every small and sparse subinstance J, so that σ is "compatible with closure" (Definition 5.7), which means J and the closure $\mathrm{cl}_S(J)$ of S in J receive distributions with identical marginals on S:

$$\pi_S(\sigma(J)) = \pi_S(\sigma(\mathrm{cl}_S(J))) \qquad \text{for variable subset } S \text{ of subinstance } J.$$

Here π_S means taking the marginal on S given a joint distribution of assignments. We can easily build a hierarchy solution s from a scheme σ compatible with closure: When defining the local distribution $s(S)$ on S, take into account the constraints and variables in the local closure $\mathrm{cl}^t_S(I)$ for S, and finally take the marginal on S given the distribution $\sigma(\mathrm{cl}^t_S(I))$ of satisfying assignments to the closure. Therefore $s(S) \stackrel{\text{def}}{=} \pi_S(\sigma(\mathrm{cl}^t_S(I)))$. Compatibility with closure ensures our hierarchy solution is consistent with projection (Proposition 8.15), which is crucial for step (D).

Provided the CSP is τ-wise uniform, a scheme σ compatible with closure can be built as follows [BGMT12, Lemma 3.2]: Starting with the null subinstance J_0 (that has no variable), keep adding a constraint C_i in J to the current subinstance J_i iteratively for $i = 0, \ldots, |\mathcal{C}(J)| - 1$, such that C_i has at most τ variables shared with previous constraints. At the same time, starting with a local solution $\sigma(J_0)$ for the null subinstance, extend our current local solution $\sigma(J_i)$ iteratively, until we have constructed the local solution $\sigma(J)$ for the whole subinstance J (assuming J has no isolated variables). This works as long as J is small and sparse enough so that J can be decomposed in the said manner (Definition 5.6), completing step (C).

We apply the same framework to BW and AIP. For each hierarchy, we identify a class of "nice" CSPs from which a scheme can be built as above. We call such CSPs τ-*neutral* (Definition 4.3). For BW, the definition of τ-wise neutrality is hinted at in [CM13]: the set satisfying assignments R of each constraint needs to have full projection on any subset of τ variables. For AIP, on the other hand, the correct definition is less obvious. We propose the following definition of τ-wise neutral CSP for AIP:

> For every k-ary constraint of the CSP with satisfying assignments $R \subseteq D^k$, every $a \in D^k$, some AIP solution $w : R \to \mathbb{Z}$ supported on R has $\pi_T(w) = \pi_T(\mathbb{1}_a)$ for every subset $T \subseteq [k]$ of size at most τ.

(Here D is the domain of the CSP.) Our definition is fairly general: it holds for CSPs whose satisfying assignments contain a Hamming ball of radius τ (Proposition 4.5). This completes step (B).

Steps (A) and (D) are the same for BW, AIP, and LP. It remains to complete step (C) and come up with a scheme for each hierarchy. Devising a scheme for AIP (Section 7) is more challenging and requires more ideas than for BW (Section 6) or LP (Section 10), in part due to AIP lacking the convenient probabilistic interpretation of LP solutions.

Not every hierarchy lower bound technique falls into our framework. Previous BW/resolution lower bounds often used the subadditive complexity measure technique [BSW01]; our framework instead uses the closure and sparsity technique of [KMOW17] that also works for other hierarchies.

The above, together with previous SDP lower bounds, yields Theorem 1.1. This also easily implies lower bounds for random 4-SAT, because there is a pairwise uniform distribution supported on all satisfying assignments of 4-SAT, making 4-SAT pairwise neutral for SDP+AIP (Theorem 8.21).

However, this simple approach falls short of SDP+AIP lower bounds for random 3-SAT, which is pairwise neutral for SDP and AIP separately but not pairwise neutral for SDP+AIP.

1.2. **Proof overview: Second main theorem.**

SDP+AIP lower bounds for random 3-SAT calls for our second main result, based on new LP and SDP solutions with full support on small closures. More precisely, for every small subset K of variables, we construct an LP solution that has full support on all satisfying assignments of the K-closure (Definition 8.19). Such an LP solution imposes no additional constraint on the AIP solution, apart from the existing constraints in the instance (Lemma 8.20). This cleanly decouples our LP/SDP construction from the AIP construction, and is a significant departure from the symmetry-based LP+AIP construction in [CŽ23b].

To ensure full support, to every constraint C in the K-closure $\mathcal{E} \stackrel{\text{def}}{=} \mathrm{cl}_K^t(\mathcal{C})$, its distribution $\eta^{\mathcal{E}}(C)$ is *planted* to be the uniform distribution over all satisfying assignments of C; and to every constraint C outside $\mathcal{E}$, its distribution $\eta^{\mathcal{E}}(C)$ is the given τ-wise uniform distribution. This yields a family $\eta^{\mathcal{E}}$ of distributions of satisfying assignments (Definition 12.1). Constructing LP and SDP solutions based on $\eta^{\mathcal{E}}$ calls for subtle changes to [BGMT12, KMOW17]: Because constraints C in $\mathcal{E}$ can have distributions $\eta^{\mathcal{E}}(C)$ that are not τ-wise uniform (and may as well be 0-wise uniform), these constraints can also affect the local distribution on any subset S. We have to *augment* the closure of every S to take into account the constraints in $\mathcal{E}$ when constructing the local distribution of S (Definition 9.1).

To ensure the distribution on an augmented closure J is well defined, every constraint C in J has to be satisfied by an assignment in the support $\mathrm{supp}(\eta^{\mathcal{E}}(C))$ of $\eta^{\mathcal{E}}(C)$ (Definition 10.3). Why are augmented closures satisfiable like that? The simple but crucial observation is that $\mathrm{supp}(\eta^{\mathcal{E}}(C))$ is τ-wise neutral for BW. Augmented closures are sparse enough to have a BW hierarchy solution, so they must be $\eta^{\mathcal{E}}$-satisfiable (Lemma 12.4). Compared with [BGMT12, KMOW17], our LP and SDP solutions gain full support, with their level halved only.

Along the way, we also simplify the SDP construction of [KMOW17]. SDP construction is more challenging than BW/LP/AIP because SDP vectors are more "global", and S-closure alone does not account for all constraints affecting the SDP vectors on a subset S. [KMOW17, Theorem 6.20] defined the *witness* of S for this purpose, i.e. W_S in step (A) of the above framework. Its complete definition depends on a multi-stage Gram–Schmidt procedure and is complicated. We instead work with *ancestor closure* (Section 11.1), a vast simplification of witness. Ancestor closure depends only on the underlying hypergraph of the instance, and not on any Gram–Schmidt procedure, e.g. not on any total ordering of variable subsets, or the constraints, or previous stages of the procedure. We believe our definition ancestor closure is the "right" one concerning SDP vectors coming from τ-wise uniform CSPs.

We also replace the global Gram–Schmidt procedures in [BCK15, KMOW17] with a new orthogonal decomposition (Section 11.2), distilling [KMOW17] into its essence. Our decomposition is inspired by the Efron–Stein decomposition of functions on a product space. Our decomposition is also inspired by previous SDP vector decompositions for k-XOR [Gri01, Sch08] and pairwise uniform subgroups [Tul09, Cha16], generalizing them to τ-wise uniform CSPs. Previous decompositions are based on equivalence classes of linear equations coming from the abelian group structure of the CSP predicate. The equivalence classes Γ in our decomposition are instead based on ancestor closures and do not require the group structure.

1.3. **Paper organization.**

We define all the hierarchies in Section 3, and the τ-wise neutral conditions for each hierarchy in Section 4. Section 5 discusses two important concepts for our lower bounds: closure and scheme. Each hierarchy has its own section on the solution construction: BW in Section 6; AIP in Section 7; LP in Section 10; SDP in Section 11. Section 8 combines the results of previous sections and proves

Theorem 1.1. Section 9 introduces the concept of augmented closure that is crucial for Theorem 1.2. Section 12 constructs SDP solutions with full support and proves Theorem 1.2.

2. Preliminaries

A set is r-small if it has at most r elements. $\mathbb{N} \stackrel{\text{def}}{=} \{0, 1, 2, \dots\}$ denotes the set of natural numbers, and $[k] \stackrel{\text{def}}{=} \{1, 2, \dots, k\}$.

The notation $o_{n;k}(1)$ represents a function $\varepsilon_k(n)$ that, for every fixed k, $\varepsilon_k(n) \to 0$ as $n \to \infty$.

Given a function $a : S \to D$, $a_T : T \to D$ denotes its restriction to $T \subseteq S$.

The notation $f : (x \in X) \to Y_x$ means that f is a dependent function mapping every $x \in X$ to $f(x) \in Y_x$, where the codomain Y_x varies with x.

Given a subset $S \subseteq V$ and assignment $a \in D^S$, denote by $\mathbb{1}_a : D^V \to \mathbb{Z}$ the indicator function of a, defined as $\mathbb{1}_a(b) = 1$ if $b_S = a_S$, and $\mathbb{1}_a(b) = 0$ otherwise.

Given functions $a : S \to D$ and $b : T \to D$ defined on subsets $S, T \subseteq V$ such that $a_{S\cap T} = b_{S\cap T}$, $(a \cup b)$ denotes the combined function: $(a \cup b)(v) = a(v)$ if $v \in S$, and $(a \cup b)(v) = b(v)$ if $v \in T$.

For simplicity, we consider only constraint satisfaction problems with k *distinct* variables per constraint (k-CSP). A constraint satisfaction problem is a pair $(D, \mathfrak{R})$. Its domain D is a nonempty finite set. $\mathfrak{R}$ is a nonempty family of k-ary relations over D, i.e. $R \subseteq D^k$ for every $R \in \mathfrak{R}$. An instance $I = (V, \mathcal{C})$ of the CSP $(D, \mathfrak{R})$ consists of a finite set V of variables and a finite set $\mathcal{C}$ of constraints over V. Every constraint $C \in \mathcal{C}$ is of the form $C = (S, R)$, where the scope $S \in V^k$ of C is a sequence of k distinct variables, and $R \in \mathfrak{R}$ is the set of accepting assignments of C. We also write $v \in C$ if variable v belongs to the scope of C.

A partial assignment $a \in D^{V(C)}$ on a constraint $C = (S, R)$ with scope $S = (v_1, \dots, v_k) \in V^k$ naturally corresponds to an assignment $a' \in D^k$ via $a' \stackrel{\text{def}}{=} a \circ S$, that is, $a'(i) = a(S(i)) = a(v_i)$ for $i \in [k]$. The set of satisfying assignments for C is $A_C \stackrel{\text{def}}{=} \left\{a \in D^{V(C)} \mid a \circ S \in R\right\}$. An assignment $a \in D^V$ satisfies a constraint $C = (S, R)$ if $a_{V(C)} \in A_C$; otherwise a violates C. An assignment $a \in D^V$ satisfies a constraint set $\mathcal{C}$ or an instance $I = (V, \mathcal{C})$ if a satisfies every constraint $C \in \mathcal{C}$; otherwise a violates $\mathcal{C}$ and I.

k-SAT is the CSP $\left(\{0,1\}, \left\{R_a \mid a \in \{0,1\}^k\right\}\right)$ where the relation $R_a \stackrel{\text{def}}{=} \{0,1\}^k \setminus \{a\}$ forbids only the assignment a. k-XOR is the CSP $(\{0,1\}, \{R_{\text{even}}, R_{\text{odd}}\})$, where R_{even} (resp. R_{odd}) consists of strings $a \in \{0,1\}^k$ of even (resp. odd) parity. k-NAE is the CSP $\left(\mathbb{Z}_2, \left\{R_a \mid a \in \mathbb{Z}_2^k\right\}\right)$, where the relation $R_a \stackrel{\text{def}}{=} \mathbb{Z}_2^k \setminus \{a, \overline{a}\}$ forbids only the assignment a and its ones' complement $\overline{a} \stackrel{\text{def}}{=} a \oplus \mathbb{1}$, and $\mathbb{1} \in \{1\}^k$ is the all-one string.

A predicate CSP is a CSP whose domain D is an abelian group, and there is a set $Q \subseteq D^k$ of assignments satisfying a predicate, such that $\mathfrak{R} = \{Q + z \mid z \in D^k\}$, i.e. each constraint is formed from any other by a shift. Examples of such CSPs are k-SAT, k-XOR, and k-NAE.

3. Hierarchy

In this section, we formally define all the hierarchies in this paper.

3.1. Relaxed domain.

Let us define a common generalization of the following four hierarchies: bounded width (BW), linear program (LP), semidefinite program (SDP), and affine integer program (AIP).

Given a CSP with domain D and an instance $I = (V, \mathcal{C})$ of the CSP, a hierarchy is a collection $(\mathcal{D}_S)_S$ of sets, one for each d-small variable subset $S \subseteq V$. $\mathcal{D}_S$ is the *relaxed domain* or the *set of relaxed assignments* on S, and varies across hierarchies.

$A_S \subseteq D^S$ denotes the set of partial assignments on S that satisfy all constraints in I contained in S, where a constraint $C \in \mathcal{C}$ is contained in S if every variable of C belongs to S.

We now define the relaxed domain $\mathcal{D}_S$ of each elementary hierarchy on a subset $S \subseteq V$ of variables.

BW: $\mathcal{D}_S^{\mathrm{BW}} \stackrel{\text{def}}{=} \mathcal{P}(A_S) \setminus \{\emptyset\}$, the family of nonempty subsets over A_S; see e.g. [AD08, Definition 1]. The bounded width hierarchy is also called existential k-pebble game or k-strategy and is "dual" to bounded-width resolution refutation [AD08, Theorem 2].

LP: $\mathcal{D}_S^{\mathrm{LP}} \stackrel{\text{def}}{=} \Delta(A_S)$, the set of distributions over A_S [BGMT12, Lemma 2.3]. The LP hierarchy is known as the Sherali–Adams hierarchy.

SDP: $\mathcal{D}_S^{\mathrm{SDP}} \stackrel{\text{def}}{=} \mathcal{X}^{A_S}$, the set of functions from A_S to an arbitrary inner product space $\mathcal{X}$. Further, the SDP vectors induce distributions $\mu_R \in \Delta(A_R)$ for $2d$-smalls subsets $R \subseteq V$, so that for any d-small $S, T \subseteq V$, any $\alpha_S \in \mathcal{D}_S^{\mathrm{SDP}}, \alpha_T \in \mathcal{D}_T^{\mathrm{SDP}}, a_S \in A_S, a_T \in A_T$,

$$\langle \alpha_S(a_S), \alpha_T(a_T) \rangle_{\mathcal{X}} = \mathop{\mathbb{P}}_{b \sim \mu_{S \cup T}} [b_S = a_S, b_T = a_T]. \tag{1}$$

This combined vector program and LP formulation is [Tul09, Section 2.3]. It is equivalent to the formulation of pseudo-expectation/moment matrix in [KMOW17, Definition 2.7]. The SDP hierarchy is also known as the Lasserre hierarchy and its Lagrangian dual as the sum-of-squares hierarchy; see e.g. [Lau09, Section 6].

AIP: $\mathcal{D}_S^{\mathrm{AIP}} \stackrel{\text{def}}{=} \left\{ w : A_S \to \mathbb{Z} \;\middle|\; \sum_{a \in A_S} w(a) = 1 \right\}$, containing affine integer weights supported on A_S. Our definition here is different from [CŽ23a, CŽ23c]; see Remark 3.2.

3.2. Projection of relaxed assignment.

For each hierarchy, we consider a relaxed assignment $\alpha_S \in \mathcal{D}_S$ to be a function from the set of satisfying assignments A_S on S to a commutative monoid $\mathcal{M}$ (so that addition is defined). In each of the elementary hierarchy (BW/LP/AIP) except SDP, $\mathcal{M}$ is additionally a commutative semiring (so that multiplication is also defined).

BW: $\mathcal{M}_{\mathrm{BW}} \stackrel{\text{def}}{=} \mathbb{B}$. The Boolean algebra $\mathbb{B}$ has join $\vee$ as the addition and meet $\wedge$ as the multiplication. A set $\alpha_S \in \mathcal{P}(A_S)$ is represented by its indicator function $\mathbb{1}_{\alpha_S} : A_S \to \mathbb{B}$, so that $\mathbb{1}_{\alpha_S}(a) = 1$ if $a \in \alpha_S$, and $\mathbb{1}_{\alpha_S}(a) = 0$ otherwise.

LP: $\mathcal{M}_{\mathrm{LP}} \stackrel{\text{def}}{=} \mathbb{R}_+$, the nonnegative reals.

SDP: $\mathcal{M}_{\mathrm{SDP}} \stackrel{\text{def}}{=} \mathcal{X}$.

AIP: $\mathcal{M}_{\mathrm{AIP}} \stackrel{\text{def}}{=} \mathbb{Z}$.

For any subsets $T \subseteq S \subseteq V$, there is a projection $\pi_{S \to T} : \mathcal{M}^{D^S} \to \mathcal{M}^{D^T}$ of functions taking values in $\mathcal{M}$:

$$\pi_{S \to T}(\alpha_S)(b) \stackrel{\text{def}}{=} \sum_{\substack{a \in D^S \\ a_T = b}} \alpha_S(a) \qquad \text{for } \alpha_S \in \mathcal{M}^{D^S}, b \in D^T. \tag{2}$$

Here the sum is over $\mathcal{M}$.

If $\alpha_S : A \to \mathcal{M}$ is a function from $A \subseteq D^S$, we also think of it as a function $\alpha_S : D_S \to \mathcal{M}$ from D_S supported on A:

$$\alpha_S(b) = 0 \qquad \text{for } b \in D^S \setminus A. \tag{3}$$

Here 0 denotes the additive identity of the commutative monoid $\mathcal{M}$. Given any $\alpha : X \to \mathcal{M}$, denote by $\mathrm{supp}(\alpha) \stackrel{\text{def}}{=} \{x \in X \mid \alpha(x) \neq 0\}$ the support of α.

Eq. (2) immediately implies that projection commutes with addition:

$$\pi_{S \to T}(\alpha_S + \beta_S) = \pi_{S \to T}(\alpha_S) + \pi_{S \to T}(\beta_S) \qquad \text{for any } \alpha_S, \beta_S \in \mathcal{M}^{D^S}. \tag{4}$$

It is also easy to verify that composition of compatible projections is a projection:

$$\pi_{S \to R} = \pi_{T \to R} \circ \pi_{S \to T} \qquad \text{for any } R \subseteq T \subseteq S. \tag{5}$$

We sometimes abbreviate $\pi_{S \to T}$ as π_T when S is clear from the context.

Let $\mathcal{S}_d \stackrel{\text{def}}{=} \binom{V}{\leqslant d}$ be the family of d-small variable subsets.

Definition 3.1. A level-d hierarchy solution is a dependent function $s : (S \in \mathcal{S}_d) \to \mathcal{D}_S$ that is consistent with projection:

$$s(T) = \pi_{S \to T}(s(S)) \qquad \text{for every } T \subseteq S \subseteq V, |S| \leqslant d.$$

Every satisfiable instance has a level-$|V|$ hierarchy solution in each of the elementary hierarchy. Indeed, if $a : V \to D$ is a satisfying assignment, then $S \mapsto \mathbb{1}_{a_S}$ is a hierarchy solution for BW, LP, and AIP. Fixing an arbitrary unit vector v in an inner product space $\mathcal{X}$, the functions $\alpha_S(a) \stackrel{\text{def}}{=} \mathbb{1}_{a_S} v$ for $S \subseteq V, a \in D^S$ also form an SDP hierarchy solution.

Remark 3.2. Ciardo and Živný [CŽ23a, CŽ23c] initiated the study of the AIP hierarchy. They defined AIP hierarchy differently [CŽ23a, Section 2]: Their relaxed domain is

$$\mathcal{D}_S^{\text{C}\check{\text{Z}}} \stackrel{\text{def}}{=} \left\{ w : D^S \to \mathbb{Z} \;\middle|\; \textstyle\sum_{a \in D^S} w(a) = 1 \right\},$$

the set of affine integer weights on D^S that need not be supported on satisfying assignments in A_S for a general S. They only require that for every constraint $C \in \mathcal{C}$, every integer weight $w \in \mathcal{D}_{V(C)}^{\text{C}\check{\text{Z}}}$ on $V(C)$ is supported on satisfying assignments of C.

Their AIP hierarchy is strictly weaker than ours when the level d is at least the arity k of the CSP, because any solution to our AIP hierarchy is also a solution to theirs, but not the other way round. All the hierarchies in this paper are refutation-complete at level $d = |V|$, so that every unsatisfiable instance has no level-$|V|$ hierarchy solution. Their AIP hierarchy lacks refutation-completeness, even when $d \geqslant |V|$ [CŽ23a, Theorem 1].

In the terminology of [CŽ23c, Definition 20], our hierarchies are all *conic*, while their AIP hierarchy is not [CŽ23c, Proposition 21]. For consistency with other hierarchies in this paper, our AIP hierarchy does not follow their definition. In any case, when $d \geqslant k$, lower bounds to our AIP hierarchy imply lower bounds to their AIP hierarchy.

The LP and SDP hierarchies are often expressed as a maximization problem with an objective value being the number or fraction of constraints satisfied; see e.g. [BGMT12, Section 2.3] and [Tul09, Section 2.3]. When the objective value is the fraction of constraints satisfied, having an LP or SDP hierarchy solution (Definition 3.1) is equivalent to the LP or SDP value being 1.

Our definitions of hierarchies may differ from those in other works when level d is less than constraint arity k, in which case our hierarchies are all trivial and need not satisfy any constraint. This difference is inessential because our focus is super-constant level lower bounds, in particular $d \geqslant k$ and our definitions agree with (or is stronger than) other works'.

Each of the elementary hierarchy except SDP can be solved in $(|V||D|)^{O(d)}$ time, which is polynomial when d is a constant:

- **BW:** A solution can be found by the local consistency algorithm [BK09, paragraph after Definition 3.3].
- **LP:** This corresponds to a linear program with $(|V||D|)^{O(d)}$ variables and inequalities and can be solved in the aforementioned time. [Sch86, Theorem 13.4]
- **AIP:** This corresponds to an integer program with $(|V||D|)^{O(d)}$ variables and equalities and can be solved in the aforementioned time. [Sch86, Corollary 5.3b]

No polynomial-time algorithm is known to find an exact solution to the SDP hierarchy. Deciding the feasibility of the SDP hierarchy is a special case of the SDP feasibility problem, whose complexity is still open. Often an SDP solution that is approximately feasible (i.e. having small but positive probabilities on assignments that violate some constraints) and approximately optimal is enough for approximation algorithms. Such an approximate solution can be found in $O(|V||D|)^{O(d)}$ time because the SDP hierarchy is a semidefinite program with $(|V||D|)^{O(d)}$ variables and inequalities with the aforementioned bit-complexity [RW17, Theorem 4.1 and Corollary 3.5]; see also [O'D17] which points out the subtlety of bit-complexity. A few level-d SDP hierarchy approximation algorithms,

such as finding a large independent set in a 3-colorable graph with a certain spectral profile, can be implemented in $\mathrm{poly}(|V|, |D|)2^{O(d)}$ time [GS12], which is polynomial whenever $d = O(\log |V|)$.

3.3. Nontriviality.

Instead of taking values in $\mathcal{D}_S$, it is more convenient to define a hierarchy solution as a dependent function $s : (S \in \mathcal{S}_d) \to \mathcal{M}^{A_S}$ taking values in the superset $\mathcal{M}^{A_S} \supseteq \mathcal{D}_S$, so that s satisfies an additional *nontriviality* condition. Note that when $s(S)$ takes values in $\mathcal{M}^{A_S}$, the constant-zero function $s(S)(a) \stackrel{\text{def}}{=} 0 \in \mathcal{M}$ for $a \in A_S$, $S \in \mathcal{S}_d$, is trivially consistent with projection. The nontriviality condition is meant to rule out this invalid hierarchy solution.

Recall that the set $D^\emptyset$ contains a unique assignment, known as the empty function, which we will denote by $\mathbb{0}$. Also $A_\emptyset = D^\emptyset = \{\mathbb{0}\}$, as there is no constraint contained in $\emptyset$ that $\mathbb{0}$ can violate.

The nontriviality condition says that $s(\emptyset)(\mathbb{0})$ equals 1, the multiplicative identity of the commutative semiring $\mathcal{M}$. More explicitly, the nontriviality constraint for each elementary hierarchy is:

BW: $s(\emptyset) = \{\mathbb{0}\}$.
LP: $s(\emptyset)(\mathbb{0}) = 1$.
SDP: SDP inherits the nontriviality condition from the induced LP solution. By Eq. (1), the LP nontriviality is equivalent to the unit length condition $\|s(\emptyset)(\mathbb{0})\| = 1$ in previous works such as [Tul09, Section 2.3].
AIP: $s(\emptyset)(\mathbb{0}) = 1$.

The nontriviality condition can be written succinctly as $s(\emptyset) = \mathbb{1}_{\mathbb{0}}$.

Lemma 3.3. *For each elementary hierarchy except SDP, a dependent function $s : (S \in \mathcal{S}_d) \to \mathcal{M}^{A_S}$ that is consistent with projection and satisfies the nontriviality condition is a hierarchy solution (i.e. takes values in $\mathcal{D}_S$).*

Proof. Since s is consistent with projection, for any $S \in \mathcal{S}_d$, $s(S)$ sums to 1, i.e., $\sum_{a \in A_S} s(S)(a) = s(\emptyset)(\mathbb{0}) = 1$. This in turn is equivalent to the following statements:

BW: $s(S)$ is the indicator function of a nonempty subset of A_S.
LP: $s(S)$ is the probability mass function of a distribution over A_S.
AIP: $s(S)$ satisfies the affine constraint $\sum_{a \in A_S} s(S)(a) = 1$. □

3.4. Combined hierarchy.

Starting with [BGWŽ20], a number of hierarchies were proposed by combining the elementary ones (BW, LP, SDP, and AIP). Each combined hierarchy strengthens AIP by requiring its solution to be contained in the support of the solution of some other hierarchy.

LP+AIP: The LP+AIP relaxation was introduced in [BGWŽ20]; see also [BG19]. The LP+AIP hierarchy was formally defined in [CŽ23b, Section 3.2]. In this hierarchy, the relaxed domain for d-small $S \subseteq V$ is

$$\mathcal{D}_S^{\text{LP+AIP}} \stackrel{\text{def}}{=} \left\{ (\mu_S, w_S) \in \mathcal{D}_S^{\text{LP}} \times \mathcal{D}_S^{\text{AIP}} \;\middle|\; \operatorname{supp}(\mu_S) \supseteq \operatorname{supp}(w_S) \right\}.$$

The commutative semiring $\mathcal{M}_{\text{LP+AIP}}$ is the direct product $\mathcal{M}_{\text{LP}} \times \mathcal{M}_{\text{AIP}}$.

BW+AIP: The BW+AIP hierarchy was introduced by Dalmau and Opršal [DO23] as the $\mathbb{Z}$-affine k-consistency hierarchy. In this hierarchy, the relaxed domain for d-small $S \subseteq V$ is

$$\mathcal{D}_S^{\text{BW+AIP}} \stackrel{\text{def}}{=} \left\{ (\alpha_S, w_S) \in \mathcal{D}_S^{\text{BW}} \times \mathcal{D}_S^{\text{AIP}} \;\middle|\; \operatorname{supp}(\alpha_S) \supseteq \operatorname{supp}(w_S) \right\}.$$

SDP+AIP: The SDP+AIP hierarchy is introduced in this paper, and is equivalent to the SDA hierarchy in [CŽ24]. For every d-small $S \subseteq V$, there is $\alpha_S \in \mathcal{D}_S^{\text{SDP}}$, and for every $2d$-small $R \subseteq V$, there is $(\mu_R, w_R) \in \mathcal{D}_R^{\text{LP+AIP}}$, so that $(\alpha_S)_S$ induces $(\mu_R)_R$ as in Eq. (1).

Even though our AIP hierarchy is stronger than the AIP hierarchy in [CŽ23a, CŽ23c] when $d \geqslant k$ (Remark 3.2), our LP+AIP hierarchy is equivalent to the LP+AIP hierarchy in [CŽ23b, Section 3.2] when $d \geqslant k$. The LP+AIP hierarchy can be solved in $O(|V||D|)^{O(d)}$ time [CŽ23b, Section 3.2] by first finding a solution in the relative interior of the LP hierarchy, and then solving the AIP hierarchy with the support restricted to that of the LP solution.

By contrast, we do not know how to solve the SDP+AIP hierarchy efficiently, because finding an exact solution in the relative interior of an SDP is harder than the SDP feasibility problem. Even though a polynomial-time algorithm for SDP+AIP or finding a refutation is not known (due to the SDP feasibility obstacle), a refutation for SDP+AIP that can be verified in $O(|V||D|)^{O(d)}$ time may sometimes exist.

One possible refutation consists of a certificate for the "support" Z of the SDP hierarchy. The "support" $Z : (S \in \mathcal{S}_{2d}) \to A_S$ represents local assignments not ruled out by the SDP hierarchy, and is defined for $2d$-small $T \subseteq V$ to be

$$Z(T) \stackrel{\text{def}}{=} \left\{ a \in D^T \;\middle|\; \mu_T(a) > 0 \text{ for some LP solution } (\mu_R)_R \text{ induced by some SDP solution} \right\}.$$

A possible certificate for the support Z consists of, for each $2d$-small $T \subseteq V$ and $a \in D^T \setminus Z(T)$, a sum-of-squares proof showing $\mu_T(a) \leqslant 0$ for every LP solution induced by any SDP solution. Finally, the AIP algorithm can further certify that no AIP solution supported on Z exists.

The SDP+AIP hierarchy is a common generalization of all other hierarchies in this section, and is at least as strong as any one of them. Lower bounds for SDP+AIP unify and imply lower bounds for all these hierarchies.

4. τ-wise neutral CSP

In this section, we define τ-wise neutrality for each hierarchy. As mentioned in Section 1.1, this concept is meant the capture the notion that a constraint C becomes "invisible" to the hierarchy on subsets of at most τ variables in C. In Section 4.3 we further discuss τ-wise neutrality for AIP.

4.1. **Neutral solution.**

Each elementary hierarchy except SDP comes with a natural way to combine $\alpha_S : D^S \to \mathcal{M}$ and $\alpha_T : D^T \to \mathcal{M}$ defined on disjoint variable subsets S and T into their product $\alpha_S \otimes \alpha_T : D^{S \cup T} \to \mathcal{M}$. This is simply the tensor product $\otimes : \mathcal{M}^{D^S} \times \mathcal{M}^{D^T} \to \mathcal{M}^{D^{S \cup T}}$ of appropriate functions taking values in a commutative semiring $\mathcal{M}$:

$$(u \otimes v)(a, b) \stackrel{\text{def}}{=} u(a)v(b) \quad \text{for } (a, b) \in D^S \times D^T \cong D^{S \cup T}. \tag{6}$$

Here the multiplication is over $\mathcal{M}$. More explicitly, the tensor product for each hierarchy is:

BW: $\otimes_{\text{BW}} \stackrel{\text{def}}{=} \otimes_{\mathbb{B}}$.

LP: $\otimes_{\text{LP}} \stackrel{\text{def}}{=} \otimes_{\mathbb{R}_+}$. Note that $\alpha_S \otimes_{\mathbb{R}_+} \alpha_T$ is the product distribution of α_S and α_T in the usual sense of probability theory, i.e. sampling from α_S and α_T independently.

AIP: $\otimes_{\text{AIP}} \stackrel{\text{def}}{=} \otimes_{\mathbb{Z}}$.

To construct a solution fooling a hierarchy, first fix a *neutral* solution ν mapping every $v \in V$ to $\nu(v) \in \mathcal{D}_{\{v\}}$. Then lift $\nu(v)$ to a relaxed assignment ν^S on every subset $S \subseteq V$ by

$$\nu^S \stackrel{\text{def}}{=} \bigotimes_{v \in S} \nu(v). \tag{7}$$

We now define a neutral solution $\nu : (v \in V) \to \mathcal{D}_{\{v\}}$ for each hierarchy except SDP.

BW: $\nu_{\text{BW}}(v) \stackrel{\text{def}}{=} D^{\{v\}}$, the set of functions from $\{v\}$ to D.

LP: $\nu_{\text{LP}}(v)$ is the uniform distribution over $D^{\{v\}}$.

AIP: First fix an assignment $a : V \to D$ (such as assigning an arbitrary fixed element $0 \in D$ to every variable). Then $\nu_{\mathrm{AIP}}(v) = \mathbb{1}_{a_v}$. In other words, $\nu_{\mathrm{AIP}}(v)$ assigns weight 1 to the partial assignment $v \mapsto a(v)$ and weight 0 to every other partial assignment on $\{v\}$.

Consequently, the lifted relaxed assignment ν^S for any $S \subseteq V$ will be:

BW: $\nu^S_{\mathrm{BW}} = D^S$.
LP: ν^S_{LP} is the uniform distribution over the set D^S of all assignments from S to D.
AIP: if ν_{AIP} is the neutral solution with respect to $a : V \to D$, then $\nu^S_{\mathrm{AIP}} = \mathbb{1}_{a_S}$, i.e. ν^S_{AIP} assigns weight 1 to a_S and weight 0 to every other partial assignment on S.

The function $S \mapsto \nu^S$ is a solution to the hierarchy if I is the empty instance containing no constraint, but not in general if I contains some constraints.

4.2. τ-wise neutral CSP.

Definition 4.1. Given a neutral solution $\nu : (v \in [k]) \to \mathcal{D}_{\{v\}}$ and a natural number $\tau \in \mathbb{N}$, a relaxed assignment $\alpha : D^k \to \mathcal{M}$ is *τ-wise ν-neutral* if $\tau \leqslant k$ and

$$\pi_{V \to T}(\alpha) = \nu^T \quad \text{for every } \tau\text{-small } T \subseteq [k]. \tag{8}$$

Note that a τ-wise ν-neutral relaxed assignment α is necessarily nontrivial:

$$\pi_{[k] \to \emptyset}(\alpha) \overset{(8)}{=} \nu^{\emptyset} \overset{(7)}{=} \mathbb{1}_{\mathbb{0}}.$$

Definition 4.2. A k-CSP $(D, \mathfrak{R})$ is *τ-wise ν-neutral* if every relation $R \in \mathfrak{R}$ has a *τ-wise ν-neutral* relaxed assignment $\alpha^R : D^k \to \mathcal{M}$ supported on R.

A τ-wise ν-neutral CSP is called as such because the neutral solution ν is independent of (i.e. neutral to) the choice of the set R of satisfying assignments in a constraint.

Definition 4.3. For each elementary hierarchy, a k-CSP is *τ-wise neutral* if the CSP is:

BW: τ-wise ν-neutral for $\nu = v \mapsto D^{\{v\}}$.
LP: τ-wise ν-neutral, where $\nu(v)$ is the uniform distribution on $D^{\{v\}}$ for every v. Such a CSP is also known as τ-wise uniform (or τ-wise independent in some earlier works).
SDP: τ-wise neutral in the LP sense.
AIP: τ-wise ν_a-neutral for every $a \in D^k$, where $\nu_a \overset{\text{def}}{=} v \mapsto \mathbb{1}_{a_v}$. (We also write τ-wise ν_a-neutral as τ-wise $\mathbb{1}_a$-neutral.)

In the previous definition, AIP is somewhat unusual. For other hierarchies, τ-wise neutrality means τ-wise ν-neutral for *one specific* ν; but for AIP, it requires *a collection of ν's*. We will need a full collection of ν's to construct the AIP solution later.

A CSP is τ-wise neutral for LP+AIP if it is τ-wise $(\nu_{\mathrm{LP}}, \mathbb{1}_a)$-neutral for every $a \in D^k$. In other words, every relation R of the CSP supports a τ-wise uniform distribution μ^R of satisfying assignments, and there is a τ-wise $\mathbb{1}_a$-neutral AIP relaxed assignment supported on $\mathrm{supp}(\mu^R)$ for every $a \in D^k$. If a CSP is τ-wise neutral for LP+AIP, we also say that it is τ-wise neutral for SDP+AIP.

4.3. τ-wise neutral CSP for AIP.

In this subsection, we give examples of CSPs that are τ-wise neutral for AIP. We first show that k-SAT is τ-wise neutral for AIP.

Given $S \subseteq V$, the parity function $\chi_S : \{0,1\}^V \to \{+1, -1\} \subseteq \mathbb{Z}$ is defined as follows: $\chi_S(x) = +1$ if x_S has an even number of 1's; and $\chi_S(x) = -1$ if x_S has an odd number of 1's.

Proposition 4.4. *k-SAT is $(k-1)$-wise neutral for AIP for every $k \geqslant 1$.*

Proof. Consider the set $R = \{0,1\}^k \setminus \{b\}$ of satisfying assignments of a k-SAT constraint that forbids $b \in \{0,1\}^k$. We now show that every $a \in \{0,1\}^k$ has a $(k-1)$-wise $\mathbb{1}_a$-neutral relaxed assignment $\alpha_a : D^k \to \mathbb{Z}$ supported on R.

If $a \neq b$, then $\alpha_a \stackrel{\text{def}}{=} \mathbb{1}_a$ works. Otherwise $a = b$, and the relaxed assignment $\alpha_a \stackrel{\text{def}}{=} \mathbb{1}_b - \chi_{[k]}(b)\chi_{[k]}$ works. Indeed, this α_a is supported on R because $\alpha_a(b) = \mathbb{1}_b(b) - \chi_{[k]}(b)\chi_{[k]}(b) = 1 - 1 = 0$. This α_a is also $(k-1)$-wise $\mathbb{1}_a$-neutral, because $\pi_T(\chi_{[k]}) \equiv 0$ for $(k-1)$-small $T \subseteq [k]$ by Proposition A.2. □

Is there a simple sufficient condition for a CSP to be τ-wise neutral for each hierarchy? For BW, the necessary and sufficient is easy to find (Proposition 6.1). For LP and SDP, the necessary and sufficient condition is well known [AH11, Lemma 3.1]. For AIP, we now give a simple sufficient condition for a CSP to be τ-wise neutral: every relation contains a Hamming ball of radius τ.

Proposition 4.5. *Suppose every $R \in \mathfrak{R}$ contains some Hamming ball $B(c_R, \tau)$ of radius τ around $c_R \in D^k$, where $B(a,\tau) \stackrel{\text{def}}{=} \{b \in D^k \mid d(a,b) \leqslant \tau\}$ and $d(a,b) \stackrel{\text{def}}{=} |\{i \in [k] \mid a_i \neq b_i\}|$ is the Hamming distance. Then the k-CSP $(D, \mathfrak{R})$ is τ-wise neutral for AIP.*

Proof. Fix any center $c \in D^k$ of a Hamming ball $B(c,\tau)$. For $a \in D^k$, we construct by induction on $d(c,a)$ a τ-wise $\mathbb{1}_a$-neutral relaxed assignment α_a supported on $B(c,\tau)$.

Base Case: $d(c,a) \leqslant \tau$. Then $\alpha_a \stackrel{\text{def}}{=} \mathbb{1}_a$ is supported on $\{a\} \subseteq B(c,\tau)$ and is τ-wise $\mathbb{1}_a$-neutral.

Induction Step: $d(c,a) > \tau$. Let $Q(c,a) \stackrel{\text{def}}{=} \left\{ b \in D^k \;\middle|\; b_i \in \{a_i, c_i\} \text{ for every } i \in [k] \right\}$. Define the generalized parity function $\underline{\chi} : Q(c,a) \to \{+1,-1\} \subseteq \mathbb{Z}$ as follows: $\underline{\chi}(b) = +1$ if $d(a,b)$ is even; and $\underline{\chi}(b) = -1$ if $d(a,b)$ is odd. Consider the relaxed assignment

$$\alpha_a \stackrel{\text{def}}{=} - \sum_{b \in Q(c,a) \setminus \{a\}} \underline{\chi}(b)\alpha_b.$$

Every $b \in Q(c,a) \setminus \{a\}$ has $d(c,b) < d(c,a)$, so every α_b is supported on $B(c,\tau)$ by Induction Hypothesis, and therefore so is α_a.

Finally, for any τ-small $T \subseteq [k]$, we now verify that $\pi_T(\alpha_a) = \mathbb{1}_{a_T}$. Indeed,

$$\pi_T(\alpha_a) \stackrel{(4)}{=} - \sum_{b \in Q(c,a) \setminus \{a\}} \underline{\chi}(b)\pi_T(\alpha_b) \stackrel{\text{I.H.}}{=} - \sum_{b \in Q(c,a) \setminus \{a\}} \underline{\chi}(b)\mathbb{1}_{b_T}$$
$$\stackrel{(*)}{=} \mathbb{1}_{a_T} - \sum_{b \in Q(c,a)} \underline{\chi}(b)\mathbb{1}_{b_T} = \mathbb{1}_{a_T} - \sum_{b' \in \pi_T(Q(c,a))} \pi_T(\underline{\chi})(b')\mathbb{1}_{b'},$$

where $(*)$ uses $\underline{\chi}(a) = 1$. Let $S \stackrel{\text{def}}{=} \{i \in [k] \mid a_i \neq c_i\}$. Since $|S| = d(c,a) > \tau$ and T is τ-small, $S \setminus T \neq \emptyset$. It remains to show that $\pi_T(\underline{\chi})(b') = 0$ for any $b' \in \pi_T(Q(a,c))$. This holds because extensions $b \in Q(c,a)$ of b' correspond to choosing $b_i \in \{a_i, c_i\}$ for $i \in S \setminus T$, and half of these extensions have even $d(a,b)$ and $\underline{\chi}(b) = 1$, while the other half have odd $d(a,b)$ and $\underline{\chi}(b) = -1$. □

How few satisfying assignments can there be in the relations R of a τ-wise neutral CSP for each hierarchy? For LP and SDP, Austrin and Håstad [AH11, Theorems 1.1 and 1.2] showed that for every $\tau \geqslant 2$, most k-ary relations $R \in \binom{D^k}{t}$ are τ-wise uniform for some $t \lesssim_{|D|,\tau} k^\tau \log k$, and the $\log k$ factor can be avoided when $\tau = 2$. For AIP, Proposition 4.5 above implies that some relation $R \in \binom{D^k}{t}$ is τ-wise neutral for AIP when $t \lesssim_{|D|,\tau} k^\tau$, similar to Austrin and Håstad's bound for LP and SDP.

We also have examples of CSPs over $D = \{0,1\}$ that are pairwise neutral for AIP without containing any Hamming ball of radius 2, but not included in this paper.

5. Closure and scheme

5.1. Closure.

Following [BGMT12, KMOW17], we define S-closure to track constraints affecting the relaxed assignment on a variable subset S. Our definition is almost identical to [KMOW17, Definition 5.3], which in turn is a simplification of advice set in [BGMT12, Section 3.1]. The closure as defined in [KMOW17] will be called *local closure* instead in this work (Definition 8.7).

Given a constraint set $\mathcal{C}$, denote by $V(\mathcal{C}) \stackrel{\text{def}}{=} \{v \in V \mid v \in C \text{ for some } C \in \mathcal{C}\}$ its variable set. Its boundary $B(\mathcal{C})$ is the set of its variables that belong to a unique constraint in $\mathcal{C}$:

$$B(\mathcal{C}) \stackrel{\text{def}}{=} \{v \in V(\mathcal{C}) \mid v \text{ belongs to a unique } C \in \mathcal{C}\}.$$

Denote by $\overline{B}(\mathcal{C}) \stackrel{\text{def}}{=} V(\mathcal{C}) \setminus B(\mathcal{C})$ the set of variables belonging to multiple constraints in $\mathcal{C}$.

Definition 5.1 (S-closure)**.** Given $S \subseteq V$, constraint set $\mathcal{C}$ over V, and $\tau \in \mathbb{N}$, $\mathcal{C}$ is *S-closed* if $|V(C) \cap (S \cup \overline{B}(\mathcal{C}))| > \tau$ for every $C \in \mathcal{C}$. The *S-closure* $\mathrm{cl}_S(\mathcal{C})$ of $\mathcal{C}$ is the union of S-closed $\mathcal{C}' \subseteq \mathcal{C}$.

For our applications in subsequent sections, the parameter τ in the definition of S-closure will be chosen so that the CSP is τ-wise neutral.

Lemma 5.2. *The union $\bigcup_i \mathcal{C}_i$ of S-closed constraint sets $\mathcal{C}_i$ is S-closed. In particular, the S-closure $\mathrm{cl}_S(\mathcal{C})$ is also S-closed.*

Proof. Every $C \in \bigcup_i \mathcal{C}_i$ belongs to some $\mathcal{C}_i$, whose S-closeness implies $|V(C) \cap (S \cup \overline{B}(\mathcal{C}_i))| > \tau$. Now $B(\mathcal{C}_i) \cap V(C) \supseteq B(\bigcup_i \mathcal{C}_i) \cap V(C)$ by Lemma A.1, so $|V(C) \cap (S \cup \overline{B}(\bigcup_i \mathcal{C}_i))| > \tau$ as well. □

Lemma 5.3. *For constraint sets $\mathcal{C}' \subseteq \mathcal{C}$, $\mathrm{cl}_S(\mathcal{C}') \subseteq \mathrm{cl}_S(\mathcal{C})$.*

Proof. $\mathrm{cl}_S(\mathcal{C}') \subseteq \mathcal{C}' \subseteq \mathcal{C}$ and $\mathrm{cl}_S(\mathcal{C}')$ is S-closed by Lemma 5.2. □

We now prove that S-closure admits an equivalent definition: Start with $\mathcal{C}$; Remove $C \in \mathcal{C}$ from $\mathcal{C}$ if C violates the S-closeness property, i.e. $|V(C) \cap (S \cup \overline{B}(\mathcal{C}))| \leqslant \tau$. Keep removing constraints this way until no more constraints can be removed. We will show that the resulting set coincides with the S-closure in Definition 5.1.

Formally, define the set of "S-internal" constraints

$$R_S(\mathcal{C}) \stackrel{\text{def}}{=} \{C \in \mathcal{C} \mid |V(C) \cap (S \cup \overline{B}(\mathcal{C}))| > \tau\}, \tag{9}$$

and

$$R_S^*(\mathcal{C}) \stackrel{\text{def}}{=} \begin{cases} \mathcal{C} & \text{if } R_S(\mathcal{C}) = \mathcal{C} \\ R_S^*(R_S(\mathcal{C})) & \text{if } R_S(\mathcal{C}) \subsetneq \mathcal{C} \end{cases}. \tag{10}$$

Proposition 5.4. $R_S^*(\mathcal{C}) = \mathrm{cl}_S(\mathcal{C})$.

Proof. We first prove by induction on $\mathcal{C}$ that $\mathcal{C}' \subseteq R_S^*(\mathcal{C})$ for every S-closed $\mathcal{C}' \subseteq \mathcal{C}$.

Base Case: $\mathcal{C} = R_S(\mathcal{C})$. Then $\mathcal{C}' \subseteq \mathcal{C} \stackrel{(10)}{=} R_S^*(\mathcal{C})$.

Induction Step: $\mathcal{C} \supsetneq R_S(\mathcal{C})$. Consider any S-closed $\mathcal{C}' \subseteq \mathcal{C}$. We now show that $\mathcal{C}' \subseteq R_S(\mathcal{C})$. Indeed, every constraint $C \in \mathcal{C}'$ satisfies $|V(C) \cap (S \cup \overline{B}(\mathcal{C}'))| > \tau$, which implies $|V(C) \cap (S \cup \overline{B}(\mathcal{C}))| > \tau$ because $V(C) \cap B(\mathcal{C}) \subseteq V(C) \cap B(\mathcal{C}')$ by Lemma A.1. This means $C \in R_S(\mathcal{C})$ by Eq. (9). Thus $\mathcal{C}' \subseteq R_S(\mathcal{C})$, which implies (by Induction Hypothesis for $R_S(\mathcal{C})$) that $\mathcal{C}' \subseteq R_S^*(R_S(\mathcal{C})) \stackrel{(10)}{=} R_S^*(\mathcal{C})$.

Having completed the induction proof, apply its conclusion to the S-closed set $\mathcal{C}' = \mathrm{cl}_S(\mathcal{C})$, and get $\mathrm{cl}_S(\mathcal{C}) \subseteq R_S^*(\mathcal{C})$.

For the converse inclusion, note that $R_S(R_S^*(\mathcal{C})) = R_S^*(\mathcal{C})$ by Eq. (10), and hence $R_S^*(\mathcal{C})$ is S-closed by Eq. (9), and therefore $R_S^*(\mathcal{C}) \subseteq \mathrm{cl}_S(\mathcal{C})$. □

The equivalence in Proposition 5.4 does not appear in [KMOW17] or [BGMT12]. Both definitions of closure are useful: The removal viewpoint helps construct hierarchy solutions recursively (Lemma 5.11), while the union viewpoint helps show local closures to be small and sparse (Lemma 8.10).

Given any constraint set $\mathcal{C}$, any variable subset $S \subseteq V$, let $\overline{R}_S(\mathcal{C}) \stackrel{\text{def}}{=} \mathcal{C} \setminus R_S(\mathcal{C})$ be the set of "S-exterior" constraints.

We also define the S-closure of an instance as the subinstance (i.e. variables and constraints) affecting the relaxed assignment on S.

Definition 5.5. Given an instance $I = (V, \mathcal{C})$ and a subset $S \subseteq V$, denote by $\mathrm{cl}_S(I) \stackrel{\text{def}}{=} S \cup \mathrm{cl}_S(\mathcal{C})$ the subinstance with variable set $S \cup V(\mathrm{cl}_S(\mathcal{C}))$ and constraint set $\mathrm{cl}_S(\mathcal{C})$.

As is well known, small subinstances of a random instance are sparse. Our next definition, implicit in [BGMT12, Lemma 3.2] and [KMOW17, Theorem 5.12], highlights the key combinatorial property enjoyed by these sparse subinstances.

Definition 5.6 (Dismissible)**.** A constraint set $\mathcal{C}$ is *dismissible* if $\mathrm{cl}_\emptyset(\mathcal{C}) = \emptyset$. An instance I is *dismissible* if its constraint set $\mathcal{C}(I)$ is.

In other words, a dismissible constraint set can be completely removed by iteratively taking away $\emptyset$-exterior constraints.

5.2. Scheme.

In this subsection, we lay a common framework to construct hierarchy solutions for each elementary hierarchy except SDP, generalizing previous LP constructions in [BGMT12, Lemma 3.2] and [KMOW17, Theorem 5.12]. The framework involves a *scheme* assigning a relaxed assignment to every (small and sparse) subinstance.

Definition 5.7. A *scheme* σ for a family $\boldsymbol{\mathcal{I}}$ of instances is a function mapping $I \in \boldsymbol{\mathcal{I}}$ to $\sigma(I) \in \mathcal{M}^{A_I}$, where A_I is the set of assignments $a \in D^{V(I)}$ that satisfy all constraints in I. A scheme is *compatible with closure*[4] if for every $I = (U, \mathcal{C}) \in \boldsymbol{\mathcal{I}}$ and $S \subseteq U$,

$$\pi_{U \to S}(\sigma(I)) = \pi_{V(\mathrm{cl}_S(I)) \to S}(\sigma(\mathrm{cl}_S(I))). \tag{11}$$

A scheme σ for $\boldsymbol{\mathcal{I}}$ is *nontrivial* if $\sigma(I_\emptyset) = \mathbb{1}_\emptyset$, where $I_\emptyset \stackrel{\text{def}}{=} (\emptyset, \emptyset)$ is the *null instance* without any variable or constraint. We will later turn a nontrivial scheme into a hierarchy solution (Proposition 8.15).

When assigning a relaxed assignment to a subinstance, what really matter are the constraints in the subinstance. The isolated variables play essentially no role. Our next definition focuses on the simpler setting of assigning a relaxed assignment to the constraints of the subinstance, and considers a scheme for a *family of constraint sets* (namely, the family of constraints appearing in the family of small subinstances). At the same time, we also simplify the "compatibility with closure" condition. Recall that the S-closure can be obtained by iteratively removing S-exterior constraints (Proposition 5.4). Instead of directly requiring every instance to share the same projection as its S-closure, we instead ask for the same projection with the subinstance obtained after removing *any single* S-exterior constraint.

Definition 5.8. Let $\boldsymbol{\mathcal{C}}$ be a family of constraint sets over a variable set V. A *scheme* σ for $\boldsymbol{\mathcal{C}}$ is a function mapping $\mathcal{C} \in \boldsymbol{\mathcal{C}}$ to $\sigma(\mathcal{C}) \in \mathcal{M}^{A_\mathcal{C}}$, where $A_\mathcal{C}$ is the set of assignments $a \in D^{V(\mathcal{C})}$ that satisfy all constraints in $\mathcal{C}$. A scheme σ for $\boldsymbol{\mathcal{C}}$ is *compatible with a neutral solution* $\nu : (v \in V) \to \mathcal{D}_{\{v\}}$ if for every $\mathcal{C} \in \boldsymbol{\mathcal{C}}$, $S \subseteq V(\mathcal{C})$, $C \in \overline{R}_S(\mathcal{C})$,

$$\pi_{V(\mathcal{C}) \to S}(\sigma(\mathcal{C})) = \pi_{V(\mathcal{C}) \to S}\left(\sigma(\mathcal{C}') \otimes \nu^{V(C) \setminus V(\mathcal{C}')}\right), \tag{12}$$

[4]This definition only makes sense if $\boldsymbol{\mathcal{I}}$ is closed under closure, which is the case in our applications.

where $\mathcal{C}' \stackrel{\text{def}}{=} \mathcal{C} \setminus \{C\}$.

We will show in Lemma 5.11 below that, roughly speaking, the new compatible condition in Eq. (12) implies the original one in Eq. (11). The new condition is easier to verify in practice, because it only involves removing one constraint at a time.

Given a family $\boldsymbol{\mathcal{C}}$ of constraint sets over V, let $\boldsymbol{\mathcal{I}}(\boldsymbol{\mathcal{C}})$ be the family of instances $I = (U, \mathcal{C})$ with a constraint set in $\boldsymbol{\mathcal{C}}$, that is, $U \subseteq V$ and $\mathcal{C} \in \boldsymbol{\mathcal{C}}$.

Definition 5.9. Given a scheme σ for $\boldsymbol{\mathcal{C}}$ compatible with a neutral solution ν, its *extended scheme* σ^* for $\boldsymbol{\mathcal{I}}(\boldsymbol{\mathcal{C}})$ is

$$\sigma^*(I) \stackrel{\text{def}}{=} \sigma(\mathcal{C}) \otimes \nu^{U \setminus V(\mathcal{C})} \quad \text{for } I = (U, \mathcal{C}) \in \boldsymbol{\mathcal{I}}(\boldsymbol{\mathcal{C}}). \tag{13}$$

In the above definition, since σ is supported on satisfying assignments of $\mathcal{C} = \mathcal{C}(I)$, σ^* is indeed supported on satisfying assignments of I.

A scheme σ for $\boldsymbol{\mathcal{C}}$ is *nontrivial* if $\sigma(\emptyset) = \mathbb{1}_{\mathbb{0}}$.

Lemma 5.10. *Let σ be a scheme for a family $\boldsymbol{\mathcal{C}}$ of constraint sets compatible with a neutral solution. If σ is nontrivial, then so is its extended scheme σ^*.*

Proof. $\sigma^*(I_\emptyset) \stackrel{(13)}{=} \sigma(\emptyset) = \mathbb{1}_{\mathbb{0}}$. □

Lemma 5.11. *Let σ be a scheme for a family $\boldsymbol{\mathcal{C}}$ of constraint sets compatible with a neutral solution ν. Then its extended scheme σ^* for $\boldsymbol{\mathcal{I}}(\boldsymbol{\mathcal{C}})$ is compatible with closure.*

Proof. We prove by induction on $\mathcal{C} \in \boldsymbol{\mathcal{C}}$ that for any instance $I = (U, \mathcal{C})$, any $S \subseteq U$, $\pi_S(\sigma^*(I)) = \pi_S(\sigma^*(\mathrm{cl}_S(I)))$.

Base Case: $\mathcal{C} = \mathrm{cl}_S(\mathcal{C})$. The result follows because $\mathcal{C} = \mathrm{cl}_S(\mathcal{C})$, so $I = \mathrm{cl}_S(I)$.

Induction Step: $\mathcal{C} \supsetneq \mathrm{cl}_S(\mathcal{C})$. By Proposition 5.4 and Eq. (10), $\mathcal{C} \supsetneq R_S(\mathcal{C}) \supseteq \mathrm{cl}_S(\mathcal{C})$. Fix an arbitrary $C \in \overline{R}_S(\mathcal{C})$. Let $\mathcal{C}' \stackrel{\text{def}}{=} \mathcal{C} \setminus \{C\}$ and $T \stackrel{\text{def}}{=} S \cap V(\mathcal{C})$. Then $C \in \overline{R}_T(\mathcal{C})$. Since σ is compatible with ν,

$$\begin{aligned}
\pi_S(\sigma^*(I)) &\stackrel{(13)}{=} \pi_S\left(\sigma(\mathcal{C}) \otimes \nu^{U \setminus V(\mathcal{C})}\right) \stackrel{(52)}{=} \pi_T(\sigma(\mathcal{C})) \otimes \nu^{S \setminus V(\mathcal{C})} \\
&\stackrel{(12)}{=} \pi_T\left(\sigma(\mathcal{C}') \otimes \nu^{V(C) \setminus V(\mathcal{C}')}\right) \otimes \nu^{S \setminus V(\mathcal{C}')} \\
&\stackrel{(52)}{=} \pi_S\left(\sigma(\mathcal{C}') \otimes \nu^{(S \cup V(C)) \setminus V(\mathcal{C}')}\right) \stackrel{(13)}{=} \pi_S(\sigma^*(I')) \stackrel{\text{I.H.}}{=} \pi_S(\sigma^*(\mathrm{cl}_S(I'))),
\end{aligned}$$

where $I' = (U, \mathcal{C}')$. The desired result now follows from the fact that $\mathrm{cl}_S(\mathcal{C}') = \mathrm{cl}_S(\mathcal{C})$ (and hence $\mathrm{cl}_S(I') = \mathrm{cl}_S(I)$), which holds because (1) $\mathcal{C}' \supseteq R_S(\mathcal{C}) \supseteq \mathrm{cl}_S(\mathcal{C})$ and $\mathrm{cl}_S(\mathcal{C})$ is S-closed, so $\mathrm{cl}_S(\mathcal{C}') \supseteq \mathrm{cl}_S(\mathcal{C})$; and (2) of Lemma 5.3. □

6. BW scheme

This section concerns the bounded width hierarchy. Recall that the neutral solution $\nu \stackrel{\text{def}}{=} \nu_{\mathrm{BW}}$ maps every $v \in V$ to $D^{\{v\}}$.

Proposition 6.1. *Let $R \subseteq D^k$ be a relation. Consider an integer τ such that $0 \leqslant \tau \leqslant k$. Then some τ-wise neutral relaxed assignment $\alpha \subseteq D^k$ is supported on R if and only if*

$$\pi_T(R) = D^T \quad \textit{for } \tau\textit{-small } T \subseteq [k]. \tag{14}$$

Proof. If Eq. (14) holds, then the relaxed assignment $\alpha \stackrel{\text{def}}{=} R$ is τ-wise ν-neutral because $\pi_T(\alpha) = \pi_T(R) = D^T = \nu^T$ for τ-small $T \subseteq [k]$. Also, this α is supported on R, that is, $\alpha \subseteq R$.

Conversely, if some τ-wise neutral relaxed assignment $\alpha \subseteq R$ is supported on R, then $\pi_T(R) \supseteq \pi_T(\alpha) = \nu^T = D^T$ for τ-small $T \subseteq [k]$. But $\pi_T(R)$ is trivially a subset of D^T. This implies Eq. (14). □

Consequently, a CSP $(D, \mathfrak{R})$ is τ-wise neutral if and only if Eq. (14) holds for every $R \in \mathfrak{R}$.

k-SAT, k-NAE, k-XOR are all examples of $(k-1)$-wise neutral CSPs for BW when $k \geqslant 1$.

When $\tau = 2$, our pairwise neutral condition here is closely related to (but stronger than) the null-constraining condition in [CM13, Definition 2.6], which means some lower bounds in [CM13] do not follow from our lower bounds here. As an example, [AD22, Theorem 3] morally follows from [CM13, Theroem 1.4(a)] (and Lemma B.3) but not from Theorem 1.1.[5]

Consider the function σ mapping constraint sets $\mathcal{C}$ to the set $\sigma(\mathcal{C}) \stackrel{\text{def}}{=} A_{\mathcal{C}} \subseteq D^{V(\mathcal{C})}$ of satisfying assignments of $\mathcal{C}$.

Proposition 6.2. *σ is a nontrivial scheme for constraint sets and is compatible with ν.*

Proof. σ is supported on satisfying assignments of $\mathcal{C}$ by definition.

σ is nontrivial, because $\sigma(\emptyset) = A_\emptyset = \{\mathbb{0}\}$, which is nonempty.

We now prove that for any $S \subseteq V(\mathcal{C})$, $C \in \overline{R}_S(\mathcal{C})$, if $\mathcal{C}' \stackrel{\text{def}}{=} \mathcal{C} \setminus \{C\}$, $U \stackrel{\text{def}}{=} V(\mathcal{C}')$, then

$$\pi_S(\sigma(\mathcal{C})) = \pi_S\left(\sigma(\mathcal{C}') \otimes \nu^{V(C)\setminus U}\right).$$

For any $b \in \pi_S(\sigma(\mathcal{C}))$, some assignment $a \in V(\mathcal{C})$ satisfies all constraints of $\mathcal{C}$ and $a_S = b$. In particular a_U satisfies all constraints in $\mathcal{C}'$ and $a \in A_{\mathcal{C}'} \times D^{V(C)\setminus U} = \sigma(\mathcal{C}') \otimes \nu^{V(C)\setminus U}$, so $b \in \pi_S\left(\sigma(\mathcal{C}') \otimes \nu^{V(C)\setminus U}\right)$.

Conversely, if $b \in \pi_S\left(\sigma(\mathcal{C}') \otimes \nu^{V(C)\setminus U}\right)$, then some $a \in D^{V(\mathcal{C})}$ satisfies $a_S = b$ and a_U satisfies all constraints in $\mathcal{C}'$. Let $T \stackrel{\text{def}}{=} V(C) \cap (S \cup U) = V(C) \cap (S \cup \overline{B}(\mathcal{C}))$. Since $C \in \overline{R}_S(\mathcal{C})$, $|T| \leqslant \tau$. Since Eq. (14) holds for the relation R of C, some satisfying assignment $c \in A_C$ has $c_T = a_T$. Then the combined assignment $a' \stackrel{\text{def}}{=} a_{S\cup U} \cup c \in D^{V(\mathcal{C})}$ satisfies all constraints in $\mathcal{C}$ and has projection $a'_S = a_S = b$, so $b \in \pi_S(\sigma(\mathcal{C}))$. □

Corollary 6.3. *If a dismissible constraint set $\mathcal{C}$ is τ-wise neutral, then $\mathcal{C}$ is satisfiable.*

Proof. By Proposition 6.2 and Lemmas 5.10 and 5.11, the function $\sigma \stackrel{\text{def}}{=} I \mapsto A_I$ is a nontrivial extended scheme for instances, and σ is compatible with closure. Let $I \stackrel{\text{def}}{=} (V(\mathcal{C}), \mathcal{C})$ be the instance consisting of $\mathcal{C}$ and $I_\emptyset \stackrel{\text{def}}{=} (\emptyset, \emptyset)$ be the null instance. Then $\pi_\emptyset(\sigma(I)) \stackrel{(11)}{=} \pi_\emptyset(\sigma(\mathrm{cl}_\emptyset(I))) = \pi_\emptyset(\sigma(I_\emptyset)) \neq \emptyset$ because σ is nontrivial. Therefore $\sigma(I) \neq \emptyset$ and any assignment $a \in \sigma(I) = A_I$ satisfies $\mathcal{C}$. □

7. AIP scheme

This section concerns the AIP hierarchy. Given any τ-wise neutral CSP and any $a \in D^V$, we wish to construct a scheme for constraint sets compatible with $\mathbb{1}_a$ (or more precisely, compatible with $\nu_a \stackrel{\text{def}}{=} v \mapsto \mathbb{1}_{a_v}$). To this end, we will recursively define a function $M_a(\mathcal{C})$ that outputs an integer weight $w : D^{V(C)} \to \mathbb{Z}$ by "mending" the integer weight $\mathbb{1}_{a_{V(\mathcal{C})}}$ according to $\mathcal{C}$.

Call a family $\left\{\alpha_a \in \mathbb{Z}^{D^k} \;\middle|\; a \in D^k\right\}$ of relaxed assignments *τ-wise neutral* if α_a is τ-wise $\mathbb{1}_a$-neutral for every $a \in D^k$.

Definition 7.1. Given a family $\left\{\alpha_a \;\middle|\; a \in D^k\right\}$ of τ-wise neutral relaxed assignments, the *derived relaxed assignments* are

$$\alpha_b \stackrel{\text{def}}{=} \sum_{\substack{a \in D^k \\ a_Q = b}} \alpha_a \quad \text{for } b \in D^Q, Q \subseteq [k].$$

[5] The reason is that for any relational structure $\mathbf{S}$ satisfying the condition "$U, V \in \mathscr{G}(S)$ implies $U \circ V = S^2$" of [AD22, Theorem 3], the CSP has no forbidding path of length at least 2 and its domain S is null-constraining. However, $\mathbf{S}$ is not necessarily pairwise neutral for BW, and an example is graph 3-coloring.

Lemma 7.2. *The derived relaxed assignment α_b in Definition 7.1 is τ-wise $\mathbb{1}_b$-neutral:*

$$\pi_{[k]\to T}(\alpha_b) = \pi_{[k]\to T}(\mathbb{1}_b) \quad \text{for } \tau\text{-small subset } T \subseteq [k]. \tag{15}$$

Proof.

$$\pi_T(\alpha_b) \overset{(4)}{=} \sum_{\substack{a\in D^k\\ a_Q=b}} \pi_T(\alpha_a) \overset{(*)}{=} \sum_{\substack{a\in D^k\\ a_Q=b}} \pi_T(\mathbb{1}_a) \overset{(4)}{=} \pi_T(\mathbb{1}_b),$$

where $(*)$ is by Definition 4.1. (Here $\mathbb{1}_b : D^k \to \mathbb{Z}$.) □

Lemma 7.3. *If the relaxed assignments α_a in Definition 7.1 are all supported on $R \subseteq D^k$, then so is every derived relaxed assignment α_b.*

Proof. For every $c \in D^k \setminus R$, $\alpha_b(c) = \sum_{\substack{a\in D^k\\ a_Q=b}} \alpha_a(c) = 0$. □

Given a τ-wise neutral CSP $(D, \mathfrak{R})$, let $\left\{\alpha_a^R \;\middle|\; a \in D^k\right\}$ be a family of τ-wise neutral relaxed assignments supported on R for every $R \in \mathfrak{R}$, and $\left\{\alpha_b^R \;\middle|\; b \in D^Q, Q \subseteq [k]\right\}$ be the derived τ-wise neutral relaxed assignments for $R \in \mathfrak{R}$. They further yield relaxed assignments $\left\{\alpha_b^C \in \mathbb{Z}^{A_C} \;\middle|\; b \in D^Q, Q \subseteq V(C)\right\}$ on every constraint C of the CSP, where α_b^C is α_b^R transferred to C (Definition 7.4 below) and R is the k-ary relation of satisfying assignments of C:

Definition 7.4. Given a relaxed assignment $\alpha : D^k \to \mathcal{M}$ and a constraint C with scope S, the relaxed assignment $\alpha^C : D^{V(C)} \to \mathcal{M}$ transferred to C is $\alpha^C(a) \overset{\text{def}}{=} \alpha(a \circ S)$ for $a \in D^{V(C)}$.

Since we only consider constraints C having no repeated variables (i.e. its scope S consists of k *distinct* variables), α_b^C inherits τ-wise neutrality from α_b^R.

Definition 7.5. Given subsets S, T of V, define the *generalized tensor product* $\tilde{\otimes} : \mathbb{Z}^{D^S} \times \mathbb{Z}^{D^T} \to \mathbb{Z}^{D^{S\cup T}}$ by

$$(u \,\tilde{\otimes}\, v)(a) \overset{\text{def}}{=} u(a_S)v(a_T) \quad \text{for } a \in D^{S\cup T}, u \in \mathbb{Z}^{D^S}, v \in \mathbb{Z}^{D^T}. \tag{16}$$

It is easy to verify that $\tilde{\otimes}$ is commutative and associative. In the above definition, the subsets S and T need not be disjoint. When they are, the generalized tensor product coincides with the usual tensor product in Eq. (6).

We are now ready to define the mending function $M_a : (\mathcal{C} \in \boldsymbol{\mathcal{C}}) \to \mathbb{Z}^{D^{V(\mathcal{C})}}$ for the family of dismissible constraint sets $\boldsymbol{\mathcal{C}}$. Define $M_a(\emptyset) \overset{\text{def}}{=} \mathbb{1}_{\mathbb{0}}$. For a nonempty dismissible $\mathcal{C}$, $M_a(\mathcal{C})$ is defined recursively via inclusion-exclusion by

$$M_a(\mathcal{C}) \overset{\text{def}}{=} \sum_{\substack{\mathcal{C}'\subseteq \overline{R}_\emptyset(\mathcal{C})\\ \mathcal{C}'\neq\emptyset}} (-1)^{|\mathcal{C}'|+1} M'_a(\mathcal{C}, \mathcal{C}'), \tag{17}$$

where

$$M'_a(\mathcal{C}, \mathcal{C}') \overset{\text{def}}{=} M_a(\mathcal{C} \setminus \mathcal{C}') \otimes \mathbb{1}_{a_{V(\mathcal{C}')\setminus(B(\mathcal{C})\cup V(\mathcal{C}\setminus\mathcal{C}'))}} \tilde{\otimes} \widetilde{\bigotimes_{C\in\mathcal{C}'}} \alpha^C_{a_{V(C)\cap B(\mathcal{C})}}. \tag{18}$$

Proposition 7.6. *M_a is a nontrivial scheme for constraint sets.*

Proof. Nontriviality follows by definition of $M_a(\emptyset) = \mathbb{1}_{\mathbb{0}}$.

We prove by induction on $\mathcal{C}$ that $M_a(\mathcal{C})$ is supported on satisfying assignments of $\mathcal{C}$.

Base Case: $\mathcal{C} = \emptyset$. The result holds because no constraint can be violated.

Induction Step: $\mathcal{C} \neq \emptyset$. Then $M_a(\mathcal{C})$ is a signed sum of $M'_a(\mathcal{C}, \mathcal{C}')$ over nonempty $\mathcal{C}' \subseteq \overline{R}_\emptyset(\mathcal{C})$. For every such term, let $\overline{\mathcal{C}'} \stackrel{\text{def}}{=} \mathcal{C} \setminus \mathcal{C}'$. For any $b \in D^{V(\mathcal{C})}$,

$$M'_a(\mathcal{C}, \mathcal{C}')(b) \stackrel{(18)}{=} M_a(\overline{\mathcal{C}'}) \left(b_{V(\overline{\mathcal{C}'})}\right) \mathbb{1}_{a_X}(b_X) \prod_{C \in \mathcal{C}'} \alpha^C_{a_{V(C) \cap B(\mathcal{C})}} \left(b_{V(C)}\right), \tag{19}$$

where $X \stackrel{\text{def}}{=} V(\mathcal{C}') \setminus (B(\mathcal{C}) \cup V(\overline{\mathcal{C}'}))$. If b violates some constraint in $\overline{\mathcal{C}'} = \mathcal{C} \setminus \mathcal{C}'$, Eq. (19) is zero by Induction Hypothesis for $\overline{\mathcal{C}'} \subsetneq \mathcal{C}$. If b violates $C \in \mathcal{C}'$, Eq. (19) is also zero by Lemma 7.3. ☐

The next lemma says that given $y : D^Y \to \mathbb{Z}$ and $w : D^W \to \mathbb{Z}$, their generalized tensor product has the same projection to S as if one of them, say y, is first partially projected to $T \stackrel{\text{def}}{=} (S \cup W) \cap Y$.

Lemma 7.7. *For any variable subsets S, Y, W such that $S \subseteq Y \cup W$, any $y : D^Y \to \mathbb{Z}, w : D^W \to \mathbb{Z}$,*

$$\pi_S(y \mathbin{\tilde{\otimes}} w) = \pi_S(\pi_T(y) \mathbin{\tilde{\otimes}} w), \tag{20}$$

where $T \stackrel{\text{def}}{=} (S \cup W) \cap Y$.

Proof. For any $a \in D^S$,

$$\pi_S(y \mathbin{\tilde{\otimes}} w)(a) \stackrel{(2)}{=} \sum_b (y \mathbin{\tilde{\otimes}} w)(b) \stackrel{(16)}{=} \sum_b y(b_Y) w(b_W),$$

where the sum is over all $b \in D^{Y \cup W}$ such that $b_S = a$. Write $c \stackrel{\text{def}}{=} b_{T \cup W}$ and $c' \stackrel{\text{def}}{=} b_{Y \setminus T}$. Our assumptions on S and T imply $c_S = b_S$ and that $Y \cup W$ is the disjoint union of $T \cup W$ and $Y \setminus T$. The sum is equivalently over $c \in D^{T \cup W}$ and $c' \in D^{Y \setminus T}$ such that $c_S = a$. The sum becomes

$$\sum_c w(c_W) \sum_{c'} y(c_T \cup c') \stackrel{(2)}{=} \sum_c w(c_W) \pi_T(y)(c_T) \stackrel{(16)}{=} \sum_c (\pi_T(y) \mathbin{\tilde{\otimes}} w)(c) \stackrel{(2)}{=} \pi_S(\pi_T(y) \mathbin{\tilde{\otimes}} w). \qquad \square$$

For variable subsets $Q \subseteq R$, write $\mathbb{1}^R_b : D^R \to \mathbb{Z}$ for the indicator function of $b \in D^Q$. Given a constraint C, write $\mathbb{1}^C_b \stackrel{\text{def}}{=} \mathbb{1}^{V(C)}_b$ for $b \in D^Q$, $Q \subseteq V(C)$.

Lemma 7.8. *For any variable subsets S, W, any constraint C such that $S \subseteq W \cup V(C)$ and $|V(C) \cap (S \cup W)| \leqslant \tau$, any $Q \subseteq V(C)$, $w : D^W \to \mathbb{Z}$, $b \in D^Q$,*

$$\pi_S\left(w \mathbin{\tilde{\otimes}} \alpha^C_b\right) = \pi_S\left(w \mathbin{\tilde{\otimes}} \mathbb{1}^C_b\right). \tag{21}$$

Proof. $T \stackrel{\text{def}}{=} V(C) \cap (S \cup W)$ is τ-small. Then

$$\pi_S\left(w \mathbin{\tilde{\otimes}} \alpha^C_b\right) \stackrel{(20)}{=} \pi_S\left(w \mathbin{\tilde{\otimes}} \pi_T\left(\alpha^C_b\right)\right) \stackrel{(15)}{=} \pi_S\left(w \mathbin{\tilde{\otimes}} \pi_T\left(\mathbb{1}^C_b\right)\right) \stackrel{(20)}{=} \pi_S\left(w \mathbin{\tilde{\otimes}} \mathbb{1}^C_b\right). \qquad \square$$

Theorem 7.9. *The scheme M_a is compatible with $\mathbb{1}_a$: For any constraint set $\mathcal{C}$, variable subset $S \subseteq V(\mathcal{C})$, S-exterior constraint $C \in \overline{R}_S(\mathcal{C})$,*

$$\pi_S(M_a(\mathcal{C})) = \pi_S\left(M_a(\mathcal{C} \setminus \{C\}) \otimes \mathbb{1}_{a_{V(C) \cap B(\mathcal{C})}}\right). \tag{22}$$

Proof. We prove by induction on $\mathcal{C}$ that for every $S \subseteq V(\mathcal{C})$, $C \in \overline{R}_S(\mathcal{C})$, Eq. (22) holds.

Base Case: $\mathcal{C} = \mathrm{cl}_S(\mathcal{C})$. Eq. (22) holds vacuously because $R_S(\mathcal{C}) = \mathcal{C}$ and no $C \in \overline{R}_S(\mathcal{C})$ exists.

Induction Step: Fix any $S \subseteq V(\mathcal{C})$ and any $C \in \overline{R}_S(\mathcal{C})$. Let $\overline{C} \stackrel{\text{def}}{=} \mathcal{C} \setminus \{C\}$. Using Eq. (17), expand $M_a(\mathcal{C})$ as a signed sum of $M'_a(\mathcal{C}, \mathcal{C}')$ over nonempty $\mathcal{C}' \subseteq \overline{R}_\emptyset(\mathcal{C})$.

Consider the summand for $\mathcal{C}' = \{C\}$. After projecting to S, this summand becomes

$$\begin{aligned} \pi_S\left(M'_a(\mathcal{C}, \{C\})\right) &\stackrel{(18)}{=} \pi_S\left(M_a(\overline{C}) \mathbin{\tilde{\otimes}} \alpha^C_{a_{V(C) \setminus V(\overline{C})}}\right) \\ &\stackrel{(21)}{=} \pi_S\left(M_a(\overline{C}) \mathbin{\tilde{\otimes}} \mathbb{1}^C_{a_{V(C) \setminus V(\overline{C})}}\right) = \text{RHS of (22)}. \end{aligned}$$

We can apply Lemma 7.8 with $W \stackrel{\text{def}}{=} V(\overline{C})$ since $C \in \overline{R}_S(\mathcal{C})$ implies $|V(C) \cap (S \cup W)| \leqslant \tau$.

It remains to show that all other terms in Eq. (17) cancel each other after projecting to S. These other terms can be paired: For every nonempty $\mathcal{C}_0 \subseteq \overline{R}_\emptyset(\mathcal{C}) \setminus \{C\}$, pair up the term for $\mathcal{C}_0$ and the term for $\mathcal{C}_1 \stackrel{\text{def}}{=} \mathcal{C}_0 \cup \{C\}$. Let $\overline{\mathcal{C}}_i \stackrel{\text{def}}{=} \mathcal{C} \setminus \mathcal{C}_i$ for $i = 0, 1$.

Ignoring the sign, and after projecting to S, the term for $\mathcal{C}_0$ contributes

$$\pi_S\left(M'_a(\mathcal{C}, \mathcal{C}_0)\right) \stackrel{(18)}{=} \pi_S\left(M_a(\overline{\mathcal{C}}_0) \otimes \mathbb{1}_{a_X} \mathbin{\tilde{\otimes}} \widetilde{\bigotimes_{C \in \mathcal{C}_0}} \alpha^C_{a_{V(C)\cap B(\mathcal{C})}}\right), \tag{23}$$

where $X \stackrel{\text{def}}{=} V(\mathcal{C}_0) \setminus (B(\mathcal{C}) \cup V(\overline{\mathcal{C}}_0))$. Lemma 7.7 implies the RHS of Eq. (23) equals

$$\pi_S\left(\pi_T\left(M_a(\overline{\mathcal{C}}_0)\right) \otimes \mathbb{1}_{a_X} \mathbin{\tilde{\otimes}} \widetilde{\bigotimes_{C \in \mathcal{C}_0}} \alpha^C_{a_{V(C)\cap B(\mathcal{C})}}\right), \tag{24}$$

where $T \stackrel{\text{def}}{=} V(\overline{\mathcal{C}}_0) \cap (S \cup V(\mathcal{C}_0))$. Since $C \in \overline{R}_T(\overline{\mathcal{C}}_0)$, Induction Hypothesis implies

$$\begin{aligned}
\pi_T\left(M_a(\overline{\mathcal{C}}_0)\right) &\stackrel{\text{I.H.}}{=} \pi_T\left(M_a(\overline{\mathcal{C}}_1) \otimes \mathbb{1}_{a_{V(C)\setminus V(\overline{\mathcal{C}}_1)}}\right) \\
&\stackrel{(*)}{=} \pi_T\left(M_a(\overline{\mathcal{C}}_1) \otimes \mathbb{1}_{a_Y} \mathbin{\tilde{\otimes}} \mathbb{1}^C_{a_{V(C)\cap B(\mathcal{C})}}\right) \\
&\stackrel{(21)}{=} \pi_T\left(M_a(\overline{\mathcal{C}}_1) \otimes \mathbb{1}_{a_Y} \mathbin{\tilde{\otimes}} \alpha^C_{a_{V(C)\cap B(\mathcal{C})}}\right),
\end{aligned}$$

where $Y \stackrel{\text{def}}{=} V(C) \setminus (V(\overline{\mathcal{C}}_1) \cup B(\mathcal{C}))$, and $(*)$ uses $V(C) \setminus V(\overline{\mathcal{C}}_1) = Y \cup (V(C) \cap B(\mathcal{C}))$. We can apply Lemma 7.8 with $W \stackrel{\text{def}}{=} V(\overline{\mathcal{C}}_1) \cup Y$ since $C \in \overline{R}_S(\mathcal{C})$ implies $|V(C) \cap (T \cup W)| \leqslant \tau$. Then

$$(24) \stackrel{(20)}{=} \pi_S\left(M_a(\overline{\mathcal{C}}_1) \otimes \mathbb{1}_{a_Y} \otimes \mathbb{1}_{a_X} \mathbin{\tilde{\otimes}} \widetilde{\bigotimes_{C \in \mathcal{C}_1}} \alpha^C_{a_{V(C)\cap B(\mathcal{C})}}\right) \stackrel{(18)}{=} \pi_S\left(M'_a(\mathcal{C}, \mathcal{C}_1)\right),$$

because $Y \cup X = V(\mathcal{C}_1) \setminus (B(\mathcal{C}) \cup V(\overline{\mathcal{C}}_1))$. Since the terms for $\mathcal{C}_0$ and $\mathcal{C}_1$ have opposite signs in Eq. (17), their projections cancel each other. □

8. Hierarchy solution on expanding instances

Throughout this section, we consider a fixed τ-wise neutral CSP for some $\tau \in \mathbb{N}$. We prove our first main theorem (Theorem 1.1) in Section 8.2. As an application, we also prove SDP+AIP lower bounds for k-SAT for $k \geqslant 4$ in Section 8.3.

8.1. Expansion and local closure.

Following [KMOW17, Definition 4.8], we measure the sparsity of a subinstance as follows:[6]

Definition 8.1 (Sparsity)**.** Given a subset $S \subseteq V$ and an S-closed constraint set $\mathcal{C}$ over V, their *sparsity* is

$$\mathrm{sp}(\mathcal{C}, S) \stackrel{\text{def}}{=} 2|S \cup \overline{B}(\mathcal{C})| + (\tau + 1)|\mathcal{C}| - 2|J(\mathcal{C}, S)|, \tag{25}$$

where

$$J(\mathcal{C}, S) \stackrel{\text{def}}{=} \{(C, v) \in \mathcal{C} \times (S \cup \overline{B}(\mathcal{C})) \mid v \in C\}$$

denotes the constraint-variable incidences between $\mathcal{C}$ and $S \cup \overline{B}(\mathcal{C})$.

[6] Sparsity is called *revenue* in [KMOW17]. Eq. (25) has the factor $\tau + 1$ in place of the factor τ in the revenue. This is because our CSP is τ-wise neutral, whereas [KMOW17, Notation 2.4] is $(\tau - 1)$-wise uniform.

The constraint set $\mathcal{C}$ and variable subset S above represent a "subinstance" $I = (S, \mathcal{C})$ of a τ-wise neutral CSP, where $\mathcal{C}$ is allowed to contain variables outside S in this definition. Definition 8.1 counts how much sparser (i.e. many fewer constraint-variable incidences) I is from being potentially non-dismissible. In other words, how far away from the impossibility of recursively constructing a hierarchy solution for I from scratch. Consider the case of k-XOR, which is τ-wise neutral in the BW or LP hierarchy for $\tau = k - 1$. Consider a k-XOR subinstance $I = (U, \mathcal{C})$ whose variables each participates in exactly two constraints and $U = V(\mathcal{C})$. Such a subinstance may be unsatisfiable, if the sum of parities of all the constraints is odd. And indeed $\mathrm{sp}(\mathcal{C}, U) = 0$, signifying I is not sparse enough to guarantee satisfiability, let alone having a hierarchy solution of level $|U|$. By contrast, any k-XOR instance I all of whose nonempty constraint subsets $\mathcal{C}'$ have positive sparsity $\mathrm{sp}(\mathcal{C}', V(\mathcal{C}'))$ is dismissible.

The sparsity can be equivalently decomposed as the sum

$$\mathrm{sp}(\mathcal{C}, S) = \sum_{C \in \mathcal{C}} \mathrm{sp}_{S \cup \overline{B}(\mathcal{C})}(C) + \sum_{v \in S \cup \overline{B}(\mathcal{C})} \mathrm{sp}_{\mathcal{C}}(v), \tag{26}$$

where

$$\mathrm{sp}_T(C) \stackrel{\text{def}}{=} (\tau + 1) - |V(C) \cap T| \quad \text{for } T \subseteq V, \text{and}$$
$$\mathrm{sp}_{\mathcal{C}}(v) \stackrel{\text{def}}{=} 2 - |\{C \in \mathcal{C} \mid v \in C\}|.$$

Indeed, all the contributions to Eq. (25) and Eq. (26) are:

- Every $v \in S \cup \overline{B}(\mathcal{C})$ contributes 2 to Eq. (25) and to Eq. (26).
- Every $C \in \mathcal{C}$ contributes $\tau + 1$ to Eq. (25) and to Eq. (26).
- Every incidence $(C, v) \in J(\mathcal{C}, S)$ contributes -2 to Eq. (25). It also contributes -1 to $\mathrm{sp}_{S \cup \overline{B}(\mathcal{C})}(C)$ and -1 to $\mathrm{sp}_{\mathcal{C}}(v)$ in Eq. (26).

Remark 8.2. Since sparsity $\mathrm{sp}(\mathcal{C}, S)$ is only defined for an S-closed constraint set $\mathcal{C}$, every term $\mathrm{sp}_{S \cup \overline{B}(\mathcal{C})}(C)$ of Eq. (26) is nonpositive for $C \in \mathcal{C}$. Also, every term $\mathrm{sp}_{\mathcal{C}}(v)$ is nonpositive if $v \in \overline{B}(\mathcal{C})$. The only positive terms are $\mathrm{sp}_{\mathcal{C}}(v) = 1$ for $v \in S \cap B(\mathcal{C})$, and $\mathrm{sp}_{\mathcal{C}}(v) = 2$ for $v \in S \setminus V(\mathcal{C})$.

Remark 8.3. Any S-closed constraint set $\mathcal{C}$ has $\mathrm{sp}(\mathcal{C}, S) \leqslant 2|S|$ by the previous remark.

Definition 8.4 (Sparsity expansion)**.** Let $t \geqslant 0$ and $\gamma > 0$. A constraint set $\mathcal{C}$ over V is (t, γ)-*expanding* if $\mathrm{sp}(\mathcal{C}', S) \geqslant \gamma|\mathcal{C}'|$ for any S-closed t-small constraint subset $\mathcal{C}' \subseteq \mathcal{C}$ and any $S \subseteq V$. An instance I is (t, γ)-*expanding* if its constraint set is.

Remark 8.5. Suppose $\tau = k - 1$. Remark 8.2 and $|V(C)| = k = \tau + 1$ imply $\mathrm{sp}(\mathcal{C}, V(\mathcal{C})) = |B(\mathcal{C})|$. The expansion condition is then equivalent to $|B(\mathcal{C}')| \geqslant \tau|\mathcal{C}'|$ for t-small constraint subsets $\mathcal{C}'$. That means the factor graph of the instance is a (one-sided) unique-neighbor expander, also known as a boundary expander, when $\tau = k - 1$.

Remark 8.6. Sparsity expansion was introduced by [KMOW17] as "Plausibility Assumption." More precisely, (t, γ)-expansion of an instance $I = (V, \mathcal{C})$ is equivalent to the factor graph of I satisfying their "Plausibility Assumption", when their $(2 \cdot \mathrm{SMALL}, \zeta) = (t, \gamma)$. See [KMOW17, Remark 4.10 and "Plausibility Assumption, Restated"] for relevant definitions. Indeed, if Plausibility Assumption holds, then any $S \subseteq V$ and any S-closed t-small $\mathcal{C}' \subseteq \mathcal{C}$ corresponds to the $(\tau + 1)$-factor graph $(\mathcal{C}', V(\mathcal{C}') \cap S)$, whose plausibility implies $\mathrm{sp}(\mathcal{C}', S) \geqslant \mathrm{sp}(\mathcal{C}', V(\mathcal{C}') \cap S) \geqslant \gamma|\mathcal{C}'|$. Conversely, if I is (t, γ)-expanding, then any $(\tau + 1)$-factor graph $H \stackrel{\text{def}}{=} (\mathcal{C}', S)$ with t-small $\mathcal{C}'$ satisfies $\mathrm{sp}(\mathcal{C}', S) \geqslant \gamma|\mathcal{C}'|$ because $\mathcal{C}'$ is S-closed, and hence H is plausible.

We now define local closure [KMOW17, Definition 5.3] and recall some results in [KMOW17].

Definition 8.7 (Local closure)**.** Given a variable subset $S \subseteq V$, a constraint set $\mathcal{C}$ over V and a real number $t \geqslant 0$, the *t-local S-closure* $\mathrm{cl}_S^t(\mathcal{C})$ of $\mathcal{C}$ is the union of t-small S-closed $\mathcal{C}' \subseteq \mathcal{C}$.

Definition 8.8. Given an instance $I = (V, \mathcal{C})$, denote by $\mathrm{cl}_S^t(I) \stackrel{\text{def}}{=} S \cup \mathrm{cl}_S^t(\mathcal{C})$ the subinstance with variable set $S \cup V(\mathrm{cl}_S^t(\mathcal{C}))$ and constraint set $\mathrm{cl}_S^t(\mathcal{C})$.

Lemma 8.9. *For any t, any variable subsets $T \subseteq S$, any constraint set $\mathcal{C}$, $\mathrm{cl}_T^t(\mathcal{C}) \subseteq \mathrm{cl}_S^t(\mathcal{C})$.*

Proof. For any t-small T-closed $\mathcal{C}' \subseteq \mathcal{C}$, $\mathcal{C}' \subseteq \mathrm{cl}_S^t(\mathcal{C})$ by definition of $\mathrm{cl}_S^t(\mathcal{C})$ and S-closedness of $\mathcal{C}'$. Taking the union over all t-small T-closed $\mathcal{C}' \subseteq \mathcal{C}$, we get $\mathrm{cl}_T^t(\mathcal{C}) \subseteq \mathrm{cl}_S^t(\mathcal{C})$. □

Lemma 8.10. *Let $S \subseteq V$, and $\mathcal{C} \stackrel{\text{def}}{=} \mathcal{C}_1 \cup \mathcal{C}_2$ be the union of two t-small S-closed constraint sets $\mathcal{C}_1$ and $\mathcal{C}_2$. If $\mathcal{C}$ is $(2t, 2\gamma)$-expanding and $\mathrm{sp}(\mathcal{C}, S) \leqslant 2r\gamma$ for some $r \leqslant t$, then $\mathcal{C}$ is r-small.*

Proof. $\mathcal{C}_1$ and $\mathcal{C}_2$ are t-small, so $\mathcal{C}$ is $2t$-small. $\mathcal{C}$ is S-closed by Lemma 5.2. $\mathrm{sp}(\mathcal{C}, S) \geqslant 2\gamma|\mathcal{C}|$ by $(2t, 2\gamma)$-expansion of $\mathcal{C}$. Therefore $2r\gamma \geqslant \mathrm{sp}(\mathcal{C}, S) \geqslant 2\gamma|\mathcal{C}|$, implying $|\mathcal{C}| \leqslant r$. □

Lemma 8.11. *For any $r \leqslant t$, $\mathrm{cl}_S^t(\mathcal{C})$ is r-small if $\mathcal{C}$ is $(2t, 2\gamma)$-expanding and S is $r\gamma$-small.*

Proof. By Remark 8.3, $\mathrm{sp}(\mathcal{C}', S) \leqslant 2|S| \leqslant 2r\gamma$ for any S-closed $\mathcal{C}'$. The local closure $\mathrm{cl}_S^t(\mathcal{C})$ is a finite union of t-small S-closed constraint subsets, so an induction using Lemma 8.10 and Lemma 5.2 implies $\mathrm{cl}_S^t(\mathcal{C})$ is r-small and S-closed. □

Lemma 8.12. *For any $(2t, 2\gamma)$-expanding $\mathcal{C}$ and any t-small $\mathcal{C}' \subseteq \mathcal{C}$, $\mathcal{C}'$ is dismissible.*

Proof. $\mathrm{cl}_\emptyset^t(\mathcal{C}) = \emptyset$ by Lemma 8.11 (with $r \stackrel{\text{def}}{=} 0, S \stackrel{\text{def}}{=} \emptyset$). The next lemma implies $\mathrm{cl}_\emptyset(\mathcal{C}') \subseteq \mathrm{cl}_\emptyset^t(\mathcal{C}) = \emptyset$. □

We also relate closure to local closure.

Lemma 8.13. *For any constraint sets $\mathcal{C}' \subseteq \mathcal{C}$, any variable subset T, if $\mathcal{C}'$ is t-small, then*

$$\mathrm{cl}_T(\mathcal{C}') \subseteq \mathrm{cl}_T^t(\mathcal{C}).$$

Proof. $\mathrm{cl}_T(\mathcal{C}')$ is T-closed by Lemma 5.2. $\mathrm{cl}_T(\mathcal{C}') \subseteq \mathcal{C}' \subseteq \mathcal{C}$ by Lemma 5.3. Since $\mathcal{C}'$ is t-small, so is $\mathrm{cl}_T(\mathcal{C}')$. The result follows by definition of $\mathrm{cl}_T^t(\mathcal{C})$. □

Lemma 8.14. *For any constraint set $\mathcal{C}$, any variable subsets $T \subseteq S$, if $\mathrm{cl}_S^t(\mathcal{C})$ is t-small, then*

$$\mathrm{cl}_T(\mathrm{cl}_S^t(\mathcal{C})) = \mathrm{cl}_T^t(\mathcal{C}). \tag{27}$$

As a result, $\mathrm{cl}_T(\mathrm{cl}_S^t(I)) = \mathrm{cl}_T^t(I)$ if $\mathrm{cl}_S^t(\mathcal{C})$ is t-small.

Proof. $\mathrm{cl}_T^t(\mathcal{C}) \subseteq \mathrm{cl}_S^t(\mathcal{C})$ by Lemma 8.9. Further, $\mathrm{cl}_T^t(\mathcal{C})$ is T-closed by Lemma 5.2, so $\mathrm{cl}_T^t(\mathcal{C}) \subseteq \mathrm{cl}_T(\mathrm{cl}_S^t(\mathcal{C}))$ by definition of $\mathrm{cl}_T(\mathrm{cl}_S^t(\mathcal{C}))$.

The reverse inclusion is Lemma 8.13 (with $\mathcal{C}' \stackrel{\text{def}}{=} \mathrm{cl}_S^t(\mathcal{C})$). □

8.2. **Hierarchy solution.**

Proposition 8.15. *Suppose σ is a nontrivial scheme for dismissible instances and is compatible with closure. Let $d \stackrel{\text{def}}{=} t\gamma$. Then any $(2t, 2\gamma)$-expanding instance I has a level-d hierarchy solution s given by*

$$s(S) \stackrel{\text{def}}{=} \pi_S \circ \sigma \circ \mathrm{cl}_S^t(I) \quad \text{for } d\text{-small } S \subseteq V. \tag{28}$$

Proof. For any d-small subset $S \in \mathcal{S}_d$, its closure $\mathrm{cl}_S^t(\mathcal{C})$ is t-small by Lemma 8.11 (with $r \stackrel{\text{def}}{=} t$), therefore $\mathrm{cl}_S^t(\mathcal{C})$ is dismissible by Lemma 8.12 and $\sigma \circ \mathrm{cl}_S^t(I)$ is well defined.

Since $\mathrm{cl}_\emptyset^t(I)$ is dismissible, Lemma 8.14 and Definition 5.6 imply that $\mathrm{cl}_\emptyset^t(I)$ is the null instance $I_\emptyset \stackrel{\text{def}}{=} (\emptyset, \emptyset)$ that has no variable or constraint. Therefore $s(I_\emptyset) = \pi_\emptyset \circ \sigma \circ \mathrm{cl}_\emptyset^t(I) = \sigma(I_\emptyset) = \mathbb{1}_\mathbb{0}$ and s is nontrivial.

By Lemma 3.3, it remains to show that s is consistent with projection. Indeed, for any $S \in \mathcal{S}_d$, $T \subseteq S$,

$$\pi_{S\to T} \circ s(S) \overset{(28)}{=} \pi_{S\to T} \circ \pi_S \circ \sigma \circ \mathrm{cl}_S^t(I) \overset{(5)}{=} \pi_T \circ \sigma \circ \mathrm{cl}_S^t(I) \overset{(11)}{=} \pi_T \circ \sigma \circ \mathrm{cl}_T \circ \mathrm{cl}_S^t(I)$$
$$\overset{(27)}{=} \pi_T \circ \sigma \circ \mathrm{cl}_T^t(I) \overset{(28)}{=} s(T). \qquad \square$$

A random CSP instance is expanding with high probability. The following result will be proved in Appendix B.2 and is a variant of [KMOW17, Theorem 4.12].

Lemma 8.16. *Let $\lambda \overset{\text{def}}{=} \tau - 1 \geqslant 1$, $0 < \gamma \leqslant \lambda/2$, $\zeta \overset{\text{def}}{=} \Delta^{2/(\lambda-\gamma)}/n$. Except with probability $o_{\zeta;k}(1)$, a random k-CSP instance with n vertices and $m \overset{\text{def}}{=} \Delta n$ constraints is (t,γ)-expanding, where*

$$t = \frac{n}{\Delta^{2/(\lambda-\gamma)}} \cdot \frac{1}{2^{O(k)}}.$$

We are now ready to prove Theorem 1.1.

Theorem 1.1. *Let $\tau \geqslant 2$. For each elementary hierarchy (BW, LP, SDP, and AIP), if a k-CSP is τ-wise neutral for the hierarchy, then except with probability $o_{n;k}(1)$, a random instance of the CSP with n variables and Δn constraints has a hierarchy solution of level $\Omega_k(n/(\Delta^{2/(\tau-1)} \log \Delta))$.*

Proof. Fix $\gamma \overset{\text{def}}{=} 1/\log\Delta, \lambda \overset{\text{def}}{=} \tau - 1 \geqslant 1$. We may assume $\Delta^{2/\lambda} \log \Delta \leqslant n$, for otherwise the level lower bound is trivial, as every hierarchy has a solution of level 0. Since $\Delta^{2/(\lambda-3/\log\Delta)} = \Theta(\Delta^{2/\lambda})$ for $\lambda \geqslant 1$, we have $\zeta \overset{\text{def}}{=} \Delta^{2/(\lambda-3\gamma)}/n = o_n(1)$. Lemma 8.16 implies that except with probability $o_{n;k}(1)$, a random instance I is $(2t, 3\gamma)$-expanding, where $t = n/(\Delta^{2/(\lambda-3\gamma)} 2^{O(k)}) = \Omega_k(n/\Delta^{2/\lambda})$.

Proposition 8.15 further turns any nontrivial scheme σ of a hierarchy for dismissible instances that is compatible with closure into a level $d \overset{\text{def}}{=} t\gamma$ hierarchy solution for I. For each elementary hierarchy except SDP, any τ-wise neutral k-CSP has such a scheme σ; (See also Lemmas 5.10 and 5.11.)

BW: It follows from Proposition 6.2.

LP: This is essentially [BGMT12, Lemma 3.2] or [KMOW17, Theorem 5.12]. See Lemma 10.4 below (with $K = \emptyset$ and $\mathcal{E} = \emptyset$ so that $\underline{\mathrm{cl}} = \mathrm{cl}$) for a self-contained proof.

AIP: It follows from Proposition 7.6 and Theorem 7.9.

The SDP hierarchy requires additional work due to the extra constraint in Eq. (1). The SDP solution can be constructed by [KMOW17, Theorem 1.2]. See Theorem 11.26 and Lemma 11.27 below for a self-contained proof. $\square$

8.3. SDP+AIP hierarchy solution for 4-SAT.

Previous LP and SDP lower bounds for k-SAT, starting with [BGMT12, Sch08], were all based on the k-XOR distribution — the uniform distribution over all assignments of even (or odd) parity. As discussed in the Introduction, such a solution cannot fool the LP+AIP hierarchy, because it strengthens every constraint into k-XOR, and AIP can refute unsatisfiable k-XOR instances.

In this subsection, we give an easy workaround for k-SAT whenever $k \geqslant 4$, by constructing a $(k-2)$-wise uniform distribution supported on all satisfying assignments of a k-SAT constraint (Lemma 8.17). Such a distribution implies the SDP hierarchy solution constructed in Theorem 1.1 induces an LP hierarchy solution that has full support on small closures (Definition 8.19). Together with $(k-1)$-wise neutrality for AIP (Proposition 4.4), k-SAT is $(k-2)$-wise neutral for SDP+AIP. This easily implies k-SAT fools the SDP+AIP hierarchy whenever $k \geqslant 4$.

The main result of this subsection, Theorem 8.21, will be subsumed by Theorem 1.2, whose proof will make up the rest of this paper. We include a proof of Theorem 8.21 because it highlights an interesting distinction between 3-SAT and 4-SAT: the latter is pairwise neutral for SDP+AIP but

the former is not. This means LP+AIP and SDP+AIP lower bounds for 4-SAT follow rather easily from our results so far, while those for 3-SAT need significantly more effort.

Lemma 8.17. *For every integer $k \geqslant 4$, every k-SAT constraint C has a $(k-2)$-wise uniform distribution η_C supported precisely on all satisfying assignments of C. That is,* $\mathrm{supp}(\eta_C) = A_C$.

Proof. Suppose C forbids the assignment $b \in \mathbb{Z}_2^{V(C)}$. Let η_C be the distribution of sampling a random satisfying assignment $a \in A_C$ as follows:

(1) Pick S uniformly at random among all $S \subseteq V(C)$ of size $k-1$. Let $T \stackrel{\text{def}}{=} V(C) \setminus S$.
(2) Pick $a_T \in \mathbb{Z}_2^T$ uniformly at random.
(3) Pick a_S from $\mathbb{Z}_2^S$ as follows:
   (a) If $(a \oplus b)_T \neq 0$, then a_S is uniform from $\mathbb{Z}_2^S$.
   (b) If $(a \oplus b)_T = 0$, then a_S is uniform from $\mathbb{Z}_2^S$ conditioned on $(a \oplus b)_S$ having odd parity.

Because $(a \oplus b)_T \neq 0$ (in Case (a)) or $(a \oplus b)_S \neq 0$ (in Case (b)), assignment a satisfies C with probability 1. Every $a \in A_C$ belongs to the support of η_C, because such assignment a satisfies $(a \oplus b)_v = 1$ for some $v \in V(C)$, and there is a positive probability of choosing $S \not\ni v$ from Step (1), a_T from Step (2), and a_S from Step (3a). Finally, let $\eta_{C|S_0}$ denote the distribution of η_C conditioned on $S = S_0$ in Step (1). Claim 8.18 implies $\eta_{C|S_0}$ is $(k-2)$-wise uniform for every S_0, and therefore so is η_C, being a mixture of such distributions. □

Claim 8.18. *For every $S_0 \subseteq V(C)$ of size $k-1$, the conditional distribution $\eta_{C|S_0}$ is $(k-2)$-wise uniform.*

Proof. Fix any $R \subseteq V(C)$ of size $k-2$. When choosing a_T in Step (2), $a_{T \cap R}$ is uniform. Conditioned on a_T, $a_{S \cap R}$ is uniform because $|S \cap R| \leqslant k-2$, and the distributions of both (a) and (b) are $(k-2)$-wise uniform on S. □

Our next definition concerns hierarchy solutions that have full support on small closures.

Definition 8.19 (Full support)**.** Fix $t \in \mathbb{N}$. A level-d hierarchy solution *has full support* if $\mathrm{supp}(s(S)) \supseteq \pi_S(A_J)$ for every d-small $S \subseteq V$, where $J \stackrel{\text{def}}{=} \mathrm{cl}_S^t(I)$.

Lemma 8.20. *Let s_{LP} and s_{AIP} be level-d solutions for LP and AIP hierarchies respectively. Suppose s_{LP} has full support. If s_{AIP} is constructed by Proposition 8.15 from σ_{AIP}, then $s \stackrel{\text{def}}{=} (s_{\mathrm{LP}}, s_{\mathrm{AIP}})$ is a level-d solution for the LP+AIP hierarchy.*

Proof. Given any d-small $S \subseteq V$, consider $J \stackrel{\text{def}}{=} \mathrm{cl}_S^t(I)$. Then

$$\mathrm{supp} \circ s_{\mathrm{AIP}}(S) \overset{(28)}{=} \mathrm{supp} \circ \pi_S \circ \sigma_{\mathrm{AIP}}(J) \overset{(56)}{\subseteq} \pi_S \circ \mathrm{supp} \circ \sigma_{\mathrm{AIP}}(J) \overset{(55)}{\subseteq} \pi_S \circ A_J \overset{(*)}{\subseteq} \mathrm{supp} \circ s_{\mathrm{LP}}(S),$$

where $(*)$ is due to s_{LP} having full support. □

Theorem 8.21. *For any $k \geqslant 4$, except with probability $o_{n;k}(1)$, a random instance of 4-SAT with n variables and Δn constraints has an SDP+AIP hierarchy solution of level $\Omega_{k,\Delta}(n)$.*

Proof. Let $d'(n, k, \tau, \Delta)$ denote the level guaranteed by Theorem 1.1 (with $\tau \stackrel{\text{def}}{=} k-2$), and let $d \stackrel{\text{def}}{=} d'/2$ be half of that.

Every k-SAT constraint C has a $(k-2)$-wise uniform distribution η_C supported on all satisfying assignments by Lemma 8.17. From this, Theorem 1.1 constructs a level-d SDP solution $(\alpha_S)_S$. The proof of Theorem 1.1 assigns the canonical distribution μ_J ([KMOW17, Definition 5.10], equivalently our Definition 10.2) to the closure $J \stackrel{\text{def}}{=} \mathrm{cl}_S^t(I)$ of every $2d$-small subset $S \subseteq V$. Since η_C is supported on all satisfying assignments, the canonical distribution is supported on all satisfying assignments of J (Lemma 12.3). Therefore the level-$2d$ LP solution $(\mu_R)_R$ induced by $(\alpha_S)_S$ satisfies Definition 8.19 and has full support on small closures.

k-SAT is also $(k-1)$-wise neutral for AIP by Proposition 4.4. Theorem 1.1 constructs a level-$2d$ AIP solution $(w_R)_R$, using Proposition 8.15. Lemma 8.20 implies $(\mu_R, w_R)_R$ is a level-$2d$ LP+AIP hierarchy solution such that $(\mu_R)_R$ is induced by the level-d SDP hierarchy solution $(\alpha_S)_S$. □

9. Augmented closure

A key idea towards the proof of Theorem 1.2 is augmented closure, made precise in this section. Throughout this section, $K \subseteq V$ is a fixed variable subset.

Definition 9.1. Given $S \subseteq V$, the *K-augmented S-closure* is $\underline{\mathrm{cl}}_S \stackrel{\text{def}}{=} \mathrm{cl}_{S\cup K}$. Further, given also $t \in \mathbb{N}$, the *K-augmented t-local S-closure* is $\underline{\mathrm{cl}}^t_S \stackrel{\text{def}}{=} \mathrm{cl}^t_{S\cup K}$.

We now weaken the compatibility requirement for schemes, by restricting the removed S-exterior constraint to be outside a constraint subset $\mathcal{E}$.

Definition 9.2. Let $\mathcal{E} \subseteq \bigcup \boldsymbol{\mathcal{C}}$ be a constraint subset. A scheme σ for a family $\boldsymbol{\mathcal{C}}$ of constraint sets is *compatible with a neutral solution ν outside $\mathcal{E}$* if for any $\mathcal{C} \in \boldsymbol{\mathcal{C}}$, any $S \subseteq V(\mathcal{C})$, any $C \in \overline{R}_S(\mathcal{C}) \setminus \mathcal{E}$,

$$\pi_{V(\mathcal{C})\to S}(\sigma(\mathcal{C})) = \pi_{V(\mathcal{C})\to S}\left(\sigma(\mathcal{C}') \otimes \nu^{V(C)\setminus V(\mathcal{C}')}\right), \tag{29}$$

where $\mathcal{C}' \stackrel{\text{def}}{=} \mathcal{C} \setminus \{C\}$.

Lemma 5.11 upgrades "compatibility with constraint removal" into "compatibility with closure", by iteratively removing S-exterior constraints C. The next lemma shows that as long as the removed constraints C are all outside $\mathcal{E}$, we get a similar compatibility with closure property.

Lemma 9.3. *Let σ be a scheme for $\boldsymbol{\mathcal{C}}$ compatible with ν outside $\mathcal{E}$. Then its extended scheme σ^* for $\boldsymbol{\mathcal{I}}(\boldsymbol{\mathcal{C}})$ is compatible with augmented closure containing $\mathcal{E}$. That is, for every $I = (U, \mathcal{C}) \in \boldsymbol{\mathcal{I}}$ and $S \subseteq U$, if $\mathcal{E} \subseteq \underline{\mathrm{cl}}_S(\mathcal{C})$, then*

$$\pi_{U\to S}(\sigma(I)) = \pi_{V(\underline{\mathrm{cl}}_S(I))\to S}(\sigma(\underline{\mathrm{cl}}_S(I))). \tag{30}$$

Proof. We prove by induction on $\mathcal{C} \in \boldsymbol{\mathcal{C}}$ that for any instance $I = (U, \mathcal{C})$, any $S \subseteq U$, if $\mathcal{E} \subseteq \underline{\mathrm{cl}}_S(\mathcal{C})$, then $\pi_S(\sigma^*(I)) = \pi_S(\sigma^*(\underline{\mathrm{cl}}_S(I)))$.

Base Case: $\mathcal{C} = \underline{\mathrm{cl}}_S(\mathcal{C})$. The result follows because $\mathcal{C} = \underline{\mathrm{cl}}_S(\mathcal{C})$, so $I = \underline{\mathrm{cl}}_S(I)$.

Induction Step: $\mathcal{C} \supsetneq \underline{\mathrm{cl}}_S(\mathcal{C})$. Let $S' \stackrel{\text{def}}{=} S \cup K$. By Proposition 5.4 and Eq. (10), $\mathcal{C} \supsetneq R_{S'}(\mathcal{C}) \supseteq \mathrm{cl}_{S'}(\mathcal{C}) = \underline{\mathrm{cl}}_S(\mathcal{C})$. Fix an arbitrary $C \in \overline{R}_{S'}(\mathcal{C})$. Then $C \notin \mathcal{E}$. Let $\mathcal{C}' \stackrel{\text{def}}{=} \mathcal{C} \setminus \{C\}$ and $T \stackrel{\text{def}}{=} S' \cap V(\mathcal{C})$. Then $C \in \overline{R}_T(\mathcal{C}) \setminus \mathcal{E}$. Since σ is compatible with ν outside $\mathcal{E}$,

$$\begin{aligned}
\pi_S(\sigma^*(I)) &\stackrel{(13)}{=} \pi_S\left(\sigma(\mathcal{C}) \otimes \nu^{U\setminus V(\mathcal{C})}\right) \stackrel{(52)}{=} \pi_T(\sigma(\mathcal{C})) \otimes \nu^{S\setminus V(\mathcal{C})} \\
&\stackrel{(29)}{=} \pi_T\left(\sigma(\mathcal{C}') \otimes \nu^{V(C)\setminus V(\mathcal{C}')}\right) \otimes \nu^{S\setminus V(\mathcal{C}')} \\
&\stackrel{(52)}{=} \pi_S\left(\sigma(\mathcal{C}') \otimes \nu^{(S\cup V(C))\setminus V(\mathcal{C}')}\right) \stackrel{(13)}{=} \pi_S(\sigma^*(I')) \stackrel{\text{I.H.}}{=} \pi_S(\sigma^*(\underline{\mathrm{cl}}_S(I'))),
\end{aligned}$$

where $I' = (U, \mathcal{C}')$. The desired result now follows from the fact that $\underline{\mathrm{cl}}_S(\mathcal{C}') = \underline{\mathrm{cl}}_S(\mathcal{C})$ (and hence $\underline{\mathrm{cl}}_S(I') = \underline{\mathrm{cl}}_S(I)$), which holds because (1) $\mathcal{C}' \supseteq R_{S'}(\mathcal{C}) \supseteq \mathrm{cl}_{S'}(\mathcal{C}) = \underline{\mathrm{cl}}_S(\mathcal{C})$ and $\underline{\mathrm{cl}}_S(\mathcal{C})$ is S'-closed, so $\underline{\mathrm{cl}}_S(\mathcal{C}') = \mathrm{cl}_{S'}(\mathcal{C}') \supseteq \underline{\mathrm{cl}}_S(\mathcal{C})$; and (2) of Lemma 5.3. □

Definition 9.4. A constraint solution $\eta : (C \in \mathcal{C}) \to \mathcal{M}^{A_C}$ is *τ-wise ν-neutral outside $\mathcal{E}$* if $\eta(C)$ is τ-wise ν-neutral for every $C \in \mathcal{C} \setminus \mathcal{E}$.

Lemma 9.5. *For any constraint set $\mathcal{C}$, any variable subsets $T \subseteq S$, if $\underline{\mathrm{cl}}^t_S(\mathcal{C})$ is t-small, then*

$$\underline{\mathrm{cl}}_T(\underline{\mathrm{cl}}^t_S(\mathcal{C})) = \underline{\mathrm{cl}}^t_T(\mathcal{C}). \tag{31}$$

As a result, $\underline{\mathrm{cl}}_T(\underline{\mathrm{cl}}^t_S(I)) = \underline{\mathrm{cl}}^t_T(I)$ if $\underline{\mathrm{cl}}^t_S(\mathcal{C})$ is t-small.

Proof. Let $\underline{S} \stackrel{\text{def}}{=} S \cup K, \underline{T} \stackrel{\text{def}}{=} T \cup K$. The desired equality now becomes

$$\mathrm{cl}_{\underline{T}}(\mathrm{cl}^t_{\underline{S}}(\mathcal{C})) = \mathrm{cl}^t_{\underline{T}}(\mathcal{C}),$$

which holds by Lemma 8.14 and our assumption that $\underline{\mathrm{cl}}^t_{\underline{S}}(\mathcal{C}) = \mathrm{cl}^t_{\underline{S}}(\mathcal{C})$ is t-small. □

10. LP scheme

Towards proving Theorem 1.2, we construct LP solutions based on augmented closure defined in Section 9. Throughout this section, we consider a fixed τ-wise uniform CSP $(D, \mathfrak{R})$. Recall that the neutral solution $\nu = \nu_{\mathrm{LP}}$ maps every $v \in V$ to the uniform distribution over $D^{\{v\}}$.

Definition 10.1. Given a constraint set $\mathcal{C}$, a *family η of satisfying distributions for $\mathcal{C}$* maps every $C \in \mathcal{C}$ to a distribution $\eta(C) \in \Delta(A_C)$ of satisfying assignments of C.

Recall the following distribution of assignments on an instance I: ([BGMT12, Section 3.2] and [KMOW17, Definition 5.10])

Definition 10.2 (Canonical distribution)**.** Suppose $I = (V, \mathcal{C})$ is an instance and η a family of satisfying distributions for $\mathcal{C}$. Define the *canonical distribution* μ_I to be the following distribution over satisfying assignments $a \in A_{V(I)}$:

(1) Draw a_v uniformly from $D^{\{v\}}$ independently for isolated variable $v \in V \setminus V(\mathcal{C})$; and
(2) Draw $a_V(C)$ from $\eta(C)$ independently for $C \in \mathcal{C}$, conditioned on agreeing at their common variables in $\overline{B}(\mathcal{C})$.

Given a constraint set $\mathcal{C}$, define $\mu_{\mathcal{C}} \stackrel{\text{def}}{=} \mu_I$, where $I \stackrel{\text{def}}{=} (V(\mathcal{C}), \mathcal{C})$ is the instance consisting of $\mathcal{C}$.

Definition 10.3 (η-satisfiable)**.** An assignment $a \in D^{V(\mathcal{C})}$ *η-satisfies* a constraint set $\mathcal{C}$ if $a_{V(C)} \in \mathrm{supp}(\eta(C))$ for every $C \in \mathcal{C}$. $\mathcal{C}$ is *η-satisfiable* if some assignment a η-satisfies $\mathcal{C}$. An instance I is *η-satisfiable* if its constraint set is.

The canonical distribution μ_I in Definition 10.2 is well defined if (and only if) I is η-satisfiable.

The next lemma generalizes [KMOW17, Theorem 5.12] and [BGMT12, Lemma 3.2], allowing η to be not τ-wise uniform on a constraint subset $\mathcal{E}$, provided the augmented closure contains $\mathcal{E}$.

Lemma 10.4. *Suppose an instance $I = (V, \mathcal{C})$ is η-satisfiable. Suppose η is τ-wise uniform outside $\mathcal{E} \subseteq \mathcal{C}$. For any $S \subseteq V$, if $\mathcal{E} \subseteq \underline{\mathrm{cl}}_S(\mathcal{C})$, then the marginal of μ_I on S equals the marginal of $\mu_{\underline{\mathrm{cl}}_S(I)}$ on S. That is, $\pi_S(\mu_I) = \pi_S\left(\mu_{\underline{\mathrm{cl}}_S(I)}\right)$.*

Proof. $\mathcal{C}' \mapsto \mu_{\mathcal{C}'}$ is a scheme for constraint subsets $\mathcal{C}' \subseteq \mathcal{C}$. Proposition 10.5 below says that this scheme is compatible with the uniform neutral solution outside $\mathcal{E}$.

By Lemma 9.3, the extended scheme $J \mapsto \mu_J$ for subinstances $J \subseteq I$ is compatible with augmented closure containing $\mathcal{E}$. Then Eq. (30) implies the desired result. □

Proposition 10.5. *For any η-satisfiable constraint set $\mathcal{C}$, any $S \subseteq V(\mathcal{C})$, any $C \in \overline{R}_S(\mathcal{C}) \setminus \mathcal{E}$, let $V \stackrel{\text{def}}{=} V(\mathcal{C}), \mathcal{C}' \stackrel{\text{def}}{=} \mathcal{C} \setminus \{C'\}, T \stackrel{\text{def}}{=} V(\mathcal{C}') \cup S$. Then $\pi_T(\mu_{\mathcal{C}}) = \mu_{\mathcal{C}'} \otimes \nu^{S \setminus V(\mathcal{C}')}$.*

Proof. By Definition 10.2, for every $a \in D^V$,

$$\mu_{\mathcal{C}}(a) = \frac{1}{Z_{\mathcal{C}}} \prod_{C' \in \mathcal{C}} \eta(C')\left(a_{V(C')}\right), \quad \text{where } Z_{\mathcal{C}} \stackrel{\text{def}}{=} \sum_{a \in D^V} \prod_{C' \in \mathcal{C}} \eta(C')\left(a_{V(C')}\right),$$

and $Z_{\mathcal{C}} > 0$ because $\mathcal{C}$ is η-satisfiable. Now for every $b \in D^T$,

$$\pi_T(\mu_{\mathcal{C}})(b) = \sum_{\substack{a \in D^V \\ a_T = b}} \mu_{\mathcal{C}}(a) = \frac{1}{Z_{\mathcal{C}}} \prod_{C' \in \mathcal{C}'} \eta(C')\left(b_{V(C')}\right) \cdot \sum_{\substack{a \in D^V \\ a_T = b}} \eta(C)\left(a_{V(C)}\right). \tag{32}$$

Let $U \stackrel{\text{def}}{=} V(C) \cap T = V(C) \cap (S \cup \overline{B}(\mathcal{C}))$. Since $C \notin R_S(\mathcal{C})$, $|U| \leqslant \tau$. Also $C \notin \mathcal{E}$ implies $\eta(C)$ is τ-wise uniform. Thus

$$\sum_{\substack{a \in D^V \\ a_T = b}} \eta(C)\left(a_{V(C)}\right) = \pi_U(\eta(C))(b_U) = \frac{1}{|D^U|}, \tag{33}$$

which is independent of b. Therefore $\pi_T(\mu_{\mathcal{C}})(b)$ is proportional to $\prod_{C' \in \mathcal{C}'} \eta(C')\left(b_{V(C')}\right)$, to which $\mu'(b)$ is also proportional, where $\mu' \stackrel{\text{def}}{=} \mu_{\mathcal{C}'} \otimes \nu^{S \setminus V(\mathcal{C}')}$. Thus $\pi_T(\mu_{\mathcal{C}}) = \mu'$. □

Proposition 10.6. *Let η be a family of satisfying distributions for a constraint set $\mathcal{C}$. Suppose $\mathcal{C}$ is dismissible with respect to some $\tau \in \mathbb{N}$, and $\operatorname{supp}(\eta(C))$ is τ-wise neutral in the BW hierarchy for $C \in \mathcal{C}$. Then $\mathcal{C}$ is η-satisfiable.*

Proof. Strengthen every constraint $C = (S, R)$ in $\mathcal{C}$ to $C' \stackrel{\text{def}}{=} (S, \operatorname{supp}(\eta(C)))$ to get the new constraint set $\mathcal{C}' \stackrel{\text{def}}{=} \{C' \mid C \in \mathcal{C}\}$. Applying Corollary 6.3 to $\mathcal{C}'$, the desired conclusion follows, because any assignment that satisfies $\mathcal{C}'$ η-satisfies $\mathcal{C}$. □

Proposition 10.7. *If η is a family of τ-wise uniform satisfying distributions for a dismissible constraint set $\mathcal{C}$ then $\mathcal{C}$ is η-satisfiable.*

Proof. For every $C \in \mathcal{C}$, since $\eta(C)$ is τ-wise uniform for LP, $\operatorname{supp}(\eta(C))$ is τ-wise neutral for BW. The result now follows from Proposition 10.6. □

11. SDP solution

Towards proving Theorem 1.2, we construct SDP solutions for τ-wise uniform CSP instances. Our SDP solution is heavily inspired by [KMOW17] and reuses many of their ideas. Our presentation is significantly different, more modular, and shorter.

Throughout this section, fix an instance $I = (V, \mathcal{C})$, $\tau \in \mathbb{N}$ and $\eta : (C \in \mathcal{C}) \to \Delta(A_C)$ so that η is τ-wise uniform outside $\mathcal{E} \subseteq \mathcal{C}$. Also fix a constraint size bound $t \in \mathbb{N}$ and a variable subset size bound $d \in \mathbb{N}$. Choose $K \subseteq V$ so that the K-augmented local closure $\underline{\mathrm{cl}}^t_{\emptyset}(\mathcal{C})$ contains $\mathcal{E}$, and therefore $\emptyset \subseteq \underline{\mathrm{cl}}^t_{\emptyset}(\mathcal{C}) \subseteq \underline{\mathrm{cl}}^t_S(\mathcal{C})$ for any $S \subseteq V$ by Lemma 8.9.

Readers interested only in SDP solutions without full support can take $K = \emptyset$ and $\mathcal{E} = \emptyset$. In this special case K-augmented local closures $\underline{\mathrm{cl}}^t_S$ are simply usual local closures cl^t_S. The general case $K \neq \emptyset$ is only needed for SDP solutions with full support.

11.1. Ancestor closure.

Crucial to lower bounds in other hierarchies is the notion of closure that captures constraints (and variables) affecting the local solution on a subset S. Closure alone cannot capture all the constraints affecting the local SDP solution on S. [KMOW17] introduced *witness* for this reason. We now introduce a simplification of witness that we call *ancestor closure* $\underline{\mathrm{cl}}^t_{\mathcal{A}(S)}$.

Definition 11.1. Given a family $\mathcal{A} \subseteq \mathcal{P}(V)$ of subsets of variables, define the augmented local $\mathcal{A}$-closure $\underline{\mathrm{cl}}^t_{\mathcal{A}} \stackrel{\text{def}}{=} \underline{\mathrm{cl}}^t_U$, where $U \stackrel{\text{def}}{=} \bigcup \mathcal{A} \subseteq V$ is the union of the subsets in $\mathcal{A}$.

Definition 11.2 (Embeddable)**.** Given variable subsets $S, T \subseteq V$ of an instance $I = (V, \mathcal{C})$, T is *embeddable into* S if there are $|T|$ vertex-disjoint paths[7] from T to S in $\underline{\mathrm{cl}}^t_{\{S,T\}}(I)$. Write $T \rightarrowtail S$ if T is embeddable into S.

Definition 11.3 (Ancestor)**.** Let $\rightsquigarrow$ be the transitive closure of the binary relation $\rightarrowtail$. Let $\mathcal{A}(S) \stackrel{\text{def}}{=} \{T \mid T \rightsquigarrow S\}$ be the family of ancestors of S for $S \subseteq V$.

[7] A path p in a hypergraph H is a sequence of distinct vertices such that every two consecutive vertices in p both belong to some common hyperedge in H.

$\mathcal{A}(S)$ contains only subsets T no bigger than S, because only such T is embeddable into S. Also note that $S \rightarrowtail S$ for any $S \subseteq V$.

In other words, the ancestors $\mathcal{A}(S)$ of S transitively include $T \subseteq V$ maximally connected to S. And the ancestor closure $\underline{\mathrm{cl}}^t_{\mathcal{A}(S)}$ is the (local augmented) closure of all ancestors of S combined.

Let $\Gamma \stackrel{\text{def}}{=} \mathcal{S}_d/\sim$ be the collection of equivalence classes of d-small variable subsets, where subsets S and T are equivalent if $\mathcal{A}(S) = \mathcal{A}(T)$. For every $\mathcal{Q} \in \Gamma$, define $\mathcal{A}(\mathcal{Q}) \stackrel{\text{def}}{=} \mathcal{A}(S)$ for any $S \in \mathcal{Q}$.

Remark 11.4. Equivalently, given an instance I, define a directed graph G_d on node set $\mathcal{S}_d$ with directed edge set $\rightarrowtail \cap\, \mathcal{S}_d^2$. Then $T \rightsquigarrow S$ if and only if G_d has a path from T to S, and Γ is the collection of strongly connected components of G_d.

11.2. **Orthogonal decomposition.**

Rather than using the global Gram–Schmidt procedures in [BCK15, KMOW17], we instead construct an explicit orthogonal decomposition for the vector space of small juntas.

Definition 11.5. Given $S \subseteq V$, let $Y_S \subseteq \mathbb{R}^{D^V}$ be the vector space of S-juntas, consisting of functions $f : D^V \to \mathbb{R}$ depending only on S (that is, there exists $g : D^S \to \mathbb{R}$ such that $f(a) = g(a_S)$ for $a \in D^V$). Given a family $\mathcal{Q} \subseteq \mathcal{P}(V)$ of subsets of variables, define $Y_{\mathcal{Q}} \stackrel{\text{def}}{=} \sum_{S \in \mathcal{Q}} Y_S$ as the span of S-juntas over $S \in \mathcal{Q}$.

The subspace Y_S of S-juntas coincides with the span of indicators $\mathbb{1}_a \in \mathbb{R}^{D^V}$ over partial assignments $a \in D^S$ to S.

Our SDP solution builds upon an orthogonal decomposition for $Y_{\mathcal{S}_d}$. Our decomposition is inspired by Efron–Stein's, which we now recall.

Below when we consider the inner product, orthogonality, and orthogonal projection "under a distribution μ," we mean "in the real Hilbert space $L^2(\mu)$ of (equivalence classes of square integrable) functions with respect to μ."

Definition 11.6 (Efron–Stein [Mos10, Definition 2.10 rephrased])**.** Let ν^V be any product distribution over D^V as in Eq. (7). Given a subset $S \subseteq V$, define Z_S to be the subspace of Y_S orthogonal to all the subspaces of proper subsets of S:

$$Z_S \stackrel{\text{def}}{=} Y_S \cap Y_{\mathcal{P}^*(S)}^{\perp} \quad \text{under } \nu^S,$$

where $\mathcal{P}^*(S) \stackrel{\text{def}}{=} \mathcal{P}(S) \setminus \{S\}$. The subspaces $\{Z_S \mid S \subseteq V\}$ are mutually orthogonal under ν^V and together they span $L^2(\nu^V)$. The Efron–Stein decomposition is the orthogonal decomposition

$$L^2(\nu^V) = \bigoplus_{S \subseteq V} Z_S \qquad \text{under } \nu^V. \tag{34}$$

When $D = \mathbb{Z}_2$ and ν is uniform over D, the Efron–Stein decomposition coincides with the discrete Fourier decomposition for functions on the Boolean domain, and every Z_S is spanned by the Fourier character χ_S.

Definition 11.7. Given a family $\mathcal{A} \subseteq \mathcal{P}(V)$ of subsets of variables, define $\mu_{\mathcal{A}} \stackrel{\text{def}}{=} \mu_J$ to be the canonical distribution on the $\mathcal{A}$-closure $J \stackrel{\text{def}}{=} \underline{\mathrm{cl}}^t_{\mathcal{A}}(I)$ of I. Define the inner product

$$\langle f, g \rangle_{\mathcal{A}} \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{\mu_{\mathcal{A}}}[fg] \quad \text{for } f, g \in Y_{\mathcal{A}}. \tag{35}$$

Intuitively, the augmented local $\mathcal{A}$-closure $\underline{\mathrm{cl}}^t_{\mathcal{A}}(I)$ of I is the smallest subinstance J whose canonical distribution μ_J correctly defines an inner product on $Y_{\mathcal{A}}$, provided $\underline{\mathrm{cl}}^t_{\mathcal{A}}(I)$ is η-satisfiable so that $\mu_{\mathcal{A}}$ is well defined.

We implicitly make the following assumption for the rest of this subsection:

Assumption 11.8 (Small satisfiable closure)**.** $\underline{\mathrm{cl}}^t_{\mathcal{A}(S)}(\mathcal{C})$ is t-small and η-satisfiable for every $S \in \mathcal{S}_{2d}$.

This assumption ensures that all the distributions $\mu_{\mathcal{A}(S)}$ and $\mu_{\{S,T\}}$ considered below are well defined for $S, T \in \mathcal{S}_d$. As such, Eq. (35) is a genuine inner product on $Y_{\mathcal{A}(S)} \subseteq L^2(\mu_{\mathcal{A}(S)})$.

Analogous to Efron–Stein, we decompose $Y_{\mathcal{S}_d}$ into orthogonal subspaces, replacing the inclusion relationship of subsets by that of ancestors.

Definition 11.9. Given $\mathcal{Q} \in \Gamma$, define $\mathcal{A}^*(\mathcal{Q}) \stackrel{\text{def}}{=} \mathcal{A}(\mathcal{Q}) \setminus \mathcal{Q}$ to be the family of proper ancestors of $\mathcal{Q}$. Also define $X_{\mathcal{Q}}$ to be the subspace of $Y_{\mathcal{Q}}$ orthogonal to the subspaces of proper ancestors:

$$X_{\mathcal{Q}} \stackrel{\text{def}}{=} Y_{\mathcal{Q}} \cap Y_{\mathcal{A}^*(\mathcal{Q})}^{\perp} \quad \text{under } \mu_{\mathcal{A}(\mathcal{Q})}. \tag{36}$$

Given $\mathcal{Q} \in \Gamma$, let $\Gamma(\mathcal{Q}) \stackrel{\text{def}}{=} \{\mathcal{R} \in \Gamma \mid \mathcal{A}(\mathcal{R}) \subseteq \mathcal{A}(\mathcal{Q})\}$ be the equivalence classes of ancestors contained in $\mathcal{Q}$'s ancestors. In the next subsection, we will construct the *local* orthogonal decomposition

$$Y_{\mathcal{A}(\mathcal{Q})} = \bigoplus_{\mathcal{R} \in \Gamma(\mathcal{Q})} X_{\mathcal{R}} \qquad \text{under } \mu_{\mathcal{A}(\mathcal{Q})}. \tag{37}$$

We will also construct a *global* orthogonal decomposition over *all* equivalence classes:

$$Y_{\mathcal{S}_d} = \bigoplus_{\mathcal{R} \in \Gamma} X_{\mathcal{R}} \qquad \text{under } (\cdot,\cdot)_d, \tag{38}$$

where $(\cdot,\cdot)_d$ is the bilinear extension of the local inner products $\langle \cdot,\cdot \rangle_{\{S,T\}}$ to $Y_{\mathcal{S}_d}$. That is,

$$(f,g)_d = \langle f,g \rangle_{\{S,T\}} \stackrel{(35)}{=} \mathop{\mathbb{E}}_{\mu_{\{S,T\}}}[fg] \qquad \text{for } f \in Y_S, g \in Y_T, S, T \in \mathcal{S}_d, \tag{39}$$

and $(\cdot,\cdot)_d$ is extended to the span $Y_{\mathcal{S}_d}$ of d-juntas via bilinearity:

$$\left(\sum_{P \in \mathcal{P}} f_P, \sum_{Q \in \mathcal{Q}} g_Q \right)_d = \sum_{P \in \mathcal{P}, Q \in \mathcal{Q}} (f_P, g_Q)_d \qquad \text{for } f_P, g_Q \in Y_{\mathcal{S}_d}, \text{ any finite } \mathcal{P}, \mathcal{Q}. \tag{40}$$

The symmetric bilinear form $(\cdot,\cdot)_d$ is called pseudo-expectation in [KMOW17, Definition 5.15]. We will justify Eq. (38) by showing that subspaces $X_{\mathcal{R}}$ and $X_{\mathcal{Q}}$ are orthogonal under $(\cdot,\cdot)_d$ for distinct $\mathcal{R}, \mathcal{Q} \in \Gamma$. This in turn implies $(\cdot,\cdot)_d$ is positive-semidefinite and hence a semi-inner product.

Remark 11.10. If $I = (V, \emptyset)$ is an empty instance with no constraint, then our subspaces $\{X_{\mathcal{Q}} \mid \mathcal{Q} \in \Gamma\}$ in Eq. (36) coincides with the Efron–Stein subspaces $\{Z_S \mid S \subseteq V\}$ in Definition 11.6. Indeed, since I has no constraint, T is embeddable into S if and only if $T \subseteq S$. Every equivalence class $\mathcal{Q}$ contains exactly one subset S, so that $\mathcal{A}(\mathcal{Q}) = \mathcal{P}(S)$ and $\mathcal{A}^*(\mathcal{Q}) = \mathcal{P}^*(S)$. Assumption 11.8 holds even for $d = |V|$. Also, $\mu_I = \nu^V$ is uniform over D^V, and $(\cdot,\cdot)_d = \langle \cdot,\cdot \rangle_{\nu^V}$. Our orthogonal decomposition of small juntas in Eq. (38) reduces to the Efron–Stein decomposition.

11.3. Proof of orthogonal decomposition.

We continue to implicitly assume Assumption 11.8 in this subsection.

The next lemma ensures $(\cdot,\cdot)_d$ is well defined.

Lemma 11.11. *For any $S \in \mathcal{S}_{2d}$, any $\mathcal{R} \subseteq \mathcal{Q} \stackrel{\text{def}}{=} \mathcal{A}(S)$, any $f, g \in Y_{\mathcal{R}}$,*

$$\langle f,g \rangle_{\mathcal{R}} = \langle f,g \rangle_{\mathcal{Q}}. \tag{41}$$

Proof. Write $\mu(J)$ to mean μ_J for any instance J. Then

$$\langle f,g \rangle_{\mathcal{R}} \stackrel{(35)}{=} \mathop{\mathbb{E}}_{\mu(\underline{\mathrm{cl}}^t_{\mathcal{R}}(I))}[fg] \stackrel{(*)}{=} \mathop{\mathbb{E}}_{\mu(\underline{\mathrm{cl}}_{\mathcal{R}}(\underline{\mathrm{cl}}^t_{\mathcal{Q}}(I)))}[fg] \stackrel{(\star)}{=} \mathop{\mathbb{E}}_{\mu(\underline{\mathrm{cl}}^t_{\mathcal{Q}}(I))}[fg] \stackrel{(35)}{=} \langle f,g \rangle_{\mathcal{Q}},$$

where $(*)$ is Lemma 9.5 (with $S \stackrel{\text{def}}{=} \bigcup \mathcal{Q}, T \stackrel{\text{def}}{=} \bigcup \mathcal{R}$) and $(\star)$ is Lemma 10.4 (with $S \stackrel{\text{def}}{=} \bigcup \mathcal{R}, I \stackrel{\text{def}}{=} \underline{\mathrm{cl}}^t_{\mathcal{Q}}(I)$, together with the assumption that $\mathcal{E} \subseteq \underline{\mathrm{cl}}^t_S(\mathcal{C})$). □

Lemma 11.12. *There is a bilinear form $(\cdot,\cdot)_d$ satisfying (39) and (40).*

Proof. Construct $(\cdot,\cdot)_d$ using the Efron–Stein decomposition $Y_{\mathcal{S}_d} = \bigoplus_{S\in\mathcal{S}_d} Z_S$ for the span of d-juntas. For $f, g \in Y_{\mathcal{S}_d}$, using Eq. (34), expand $f = \sum_{S\in\mathcal{S}_d} f_S$ and $g = \sum_{T\in\mathcal{S}_d} g_T$ (where $f_S \in Z_S, g_T \in Z_T$), and define

$$(f,g)_d \stackrel{\text{def}}{=} \sum_{S,T\in\mathcal{S}_d} \langle f_S, g_T\rangle_{\{S,T\}}. \tag{42}$$

Bilinearity Eq. (40) follows because the definition of $(\cdot,\cdot)_d$ in Eq. (42) is bilinear.

For any $S \in \mathcal{S}_{2d}$, any $f, g \in Y_{\mathcal{A}(S)}$, using Eq. (34), expand $f = \sum_{R\in\mathcal{A}(S)} f_R$ and $g = \sum_{T\in\mathcal{A}(S)} g_T$ (where $f_R \in Z_R, g_T \in Z_T$), and

$$(f,g)_d \stackrel{(42)}{=} \sum_{R,T\in\mathcal{A}(S)} \langle f_R, g_T\rangle_{\{R,T\}} \stackrel{(41)}{=} \sum_{R,T\in\mathcal{A}(S)} \langle f_R, g_T\rangle_{\mathcal{A}(S)} \stackrel{(*)}{=} \langle f, g\rangle_{\mathcal{A}(S)}, \tag{43}$$

where $(*)$ uses bilinearity of $\mathcal{A}(S)$'s inner product. Eqs. (41) and (43) also imply Eq. (39). □

Lemma 11.13. *For any $\mathcal{Q} \in \Gamma$,*

$$\sum_{\mathcal{R}\in\Gamma(\mathcal{Q})} X_{\mathcal{R}} = Y_{\mathcal{A}(\mathcal{Q})}. \tag{44}$$

Proof. For any $\mathcal{R} \in \Gamma(\mathcal{Q})$, we have $X_{\mathcal{R}} \subseteq Y_{\mathcal{R}} \subseteq Y_{\mathcal{A}(\mathcal{Q})}$. Taking the span of $X_{\mathcal{R}}$ over all $\mathcal{R} \in \Gamma(\mathcal{Q})$, we get the "$\subseteq$" inclusion of Eq. (44).

We now prove the "$\supseteq$" inclusion by induction, assuming the result holds for every proper ancestor component of $\mathcal{Q}$. Note that $\{\emptyset\} \in \Gamma(\mathcal{Q})$ for every $\mathcal{Q} \in \Gamma$ because $\emptyset \rightarrowtail S$ for every $S \subseteq V$, and $T \rightarrowtail \emptyset$ implies $T = \emptyset$. Let $\Gamma^*(\mathcal{Q}) \stackrel{\text{def}}{=} \Gamma(\mathcal{Q}) \setminus \{\mathcal{Q}\}$ be the proper ancestor components of $\mathcal{Q}$.

Base Case: $\mathcal{Q} = \{\emptyset\}$. Then $\mathcal{A}(\mathcal{Q}) = \{\emptyset\}$, $\mathcal{A}^*(\mathcal{Q}) = \emptyset$, and $Y_{\mathcal{A}^*(\mathcal{Q})} = \{0\}$. Thus $X_\emptyset = Y_\emptyset \cap Y_{\mathcal{A}^*(\mathcal{Q})}^\perp = Y_\emptyset \cap \{0\}^\perp = Y_\emptyset$.

Induction Step: $\mathcal{Q} \neq \{\emptyset\}$. We have

$$Y_{\mathcal{A}(\mathcal{Q})} = Y_{\mathcal{Q}} + Y_{\mathcal{A}^*(\mathcal{Q})} \stackrel{(*)}{=} X_{\mathcal{Q}} + Y_{\mathcal{A}^*(\mathcal{Q})} = X_{\mathcal{Q}} + \sum_{S\in\mathcal{A}^*(\mathcal{Q})} Y_S \stackrel{(\star)}{\subseteq} X_{\mathcal{Q}} + \sum_{\mathcal{S}\in\Gamma^*(\mathcal{Q})} Y_{\mathcal{A}(\mathcal{S})},$$

where $(*)$ uses the orthogonal decomposition $Y_{\mathcal{Q}} = X_{\mathcal{Q}} \oplus Y_{\mathcal{A}^*(\mathcal{Q})}$ under $\mu_{\mathcal{A}(\mathcal{Q})}$, and $(\star)$ uses the fact that every $S \in \mathcal{A}^*(\mathcal{Q})$ belongs to a unique $\mathcal{S} \in \Gamma^*(\mathcal{Q})$. By Induction Hypothesis for $\mathcal{S}$, and using $\Gamma(\mathcal{S}) \subseteq \Gamma^*(\mathcal{Q})$,

$$Y_{\mathcal{A}(\mathcal{Q})} \subseteq X_{\mathcal{Q}} + \sum_{\mathcal{S}\in\Gamma^*(\mathcal{Q})} \sum_{\mathcal{R}\in\Gamma(\mathcal{S})} X_{\mathcal{R}} \subseteq X_{\mathcal{Q}} + \sum_{\mathcal{R}\in\Gamma^*(\mathcal{Q})} X_{\mathcal{R}} = \sum_{\mathcal{R}\in\Gamma(\mathcal{Q})} X_{\mathcal{R}}.$$ □

Corollary 11.14. $\sum_{\mathcal{R}\in\Gamma} X_{\mathcal{R}} = Y_{\mathcal{S}_d}$.

Proof. $Y_{\mathcal{S}_d} = \sum_{\mathcal{Q}\in\Gamma} Y_{\mathcal{Q}} \stackrel{(44)}{=} \sum_{\mathcal{Q}\in\Gamma} \sum_{\mathcal{R}\in\Gamma(\mathcal{Q})} X_{\mathcal{R}} = \sum_{\mathcal{R}\in\Gamma} X_{\mathcal{R}}.$ □

Recall that given vertex subsets R, S, T of a hypergraph H, R is an (S,T)-separator if every path in H from S to T contains some vertex in R.

Lemma 11.15 (Implicit in [KMOW17, Lemma 6.14])**.** *Let S, T be subsets of variables of an instance $J = (V, \mathcal{C})$, and R be an (S,T)-separator in J. Consider picking a random assignment b from μ_J. Then b_S and b_T are conditionally independent given b_R.*

Proof. Let V_S be the set of vertices in $J \setminus R$ reachable from S. Define V_T similarly. V_T is disjoint from V_S since R is a separator.

Further, no constraint $C \in \mathcal{C}$ contains both $u \in V_S$ and $v \in V_T$, for otherwise $J \setminus R$ has a path from S to T (via u and v), contradicting the assumption that R is an (S,T)-separator.

Define $\mathcal{C}_T \stackrel{\text{def}}{=} \{C \in \mathcal{C} \mid C \ni v \text{ for some } v \in V_T\}$. Then $V_T \cup \mathcal{C}_T$ contains no vertex in V_S.

Let μ_R be the marginal distribution of b_R (and likewise for μ_S, μ_T), and $\mu_{S|R}$ be the conditional distribution of b_S given b_R (and likewise for $\mu_{T|RS}, \mu_{ST|R}$). Then $\mu_{ST|R} = \mu_{S|R}\mu_{T|RS}$.

To sample from $\mu_{T|RS}$, it suffices to sample according to the subinstance $V_T \cup \mathcal{C}_T$ given b_{RS}. Since $V_T \cup \mathcal{C}_T$ contains no vertex in V_S, $\mu_{T|RS}$ depends only on b_R but not on $b_{S\setminus R}$. Therefore $\mu_{T|RS} = \mu_{T|R}$ and hence $\mu_{ST|R} = \mu_{S|R}\mu_{T|R}$. □

For $\mathcal{Q} \in \Gamma$, $S \in \mathcal{Q}$, define X_S to be the subspace of Y_S orthogonal to those of proper ancestors:

$$X_S \stackrel{\text{def}}{=} Y_S \cap Y_{\mathcal{A}^*(\mathcal{Q})}^{\perp} \quad \text{under } \mu_{\mathcal{A}(\mathcal{Q})}.$$

The next proposition is the main result of this subsection. It shows that if T is not embeddable into S, then $X_T \perp Y_S$. In order to prove this statement by induction, we also need to simultaneously prove an auxiliary statement.

Proposition 11.16. *For any integer c such that $0 \leqslant c \leqslant d$, the following statements hold:*

$A(c)$: If $T \in \mathcal{S}_c$ is not embeddable into $S \in \mathcal{S}_d$, then $X_T \perp Y_S$ under $(\cdot,\cdot)_d$.
$B(c)$: If $R \in \mathcal{S}_c, T \in \mathcal{S}_d$ and $|R| < |T|$, then $Y_R \perp X_T$ under $(\cdot,\cdot)_d$.

Proof. We prove by induction on c that $B(c-1) \Longrightarrow A(c)$ and $A(c) \Longrightarrow B(c)$.

$B(c-1) \Longrightarrow A(c)$:

Base Case: $c = |T| = 0$. Statement $A(0)$ holds vacuously, because $T = \emptyset$ is embeddable into any S.

Induction Step: $c = |T| > 0$. If T is not embeddable into S, Hypergraph Menger (Theorem A.13) implies $\underline{\mathrm{cl}}^t_{\{S,T\}}(I)$ has an (S,T)-separator R of size strictly less than $|T|$. For any $f \in X_T, g \in Y_S$,

$$(f,g)_d \overset{(39)}{=} \mathop{\mathbb{E}}_{\mu_{\{S,T\}}}[fg] \overset{(*)}{=} \mathop{\mathbb{E}}_{\mu_R}\left[\mathop{\mathbb{E}}_{\mu_{T|R}}[f] \cdot \mathop{\mathbb{E}}_{\mu_{S|R}}[g]\right] = \left\langle \mathop{\mathbb{E}}_{\mu_{T|R}} f, \mathop{\mathbb{E}}_{\mu_{S|R}} g \right\rangle_{\{R\}}, \tag{45}$$

where $(*)$ is Lemma 11.15, and $\mathbb{E}_{\mu_{T|R}} : Y_T \to Y_R$ denotes the conditional expectation operator, defined as

$$\left(\mathop{\mathbb{E}}_{\mu_{T|R}}[f]\right)(a) \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{b \sim \mu_{\{T,R\}}}[f(b_T) \mid b_R = a] \quad \text{for } a \in D^R, f \in Y_T, T \subseteq V, R \subseteq V.$$

The conditional expectation operator $\mathbb{E}_{\mu_{T|R}}$ is well known to coincide with the orthogonal projection operator under $\mu_{\{T,R\}}$ from Y_T to Y_R; see e.g. [Dur19, Theorem 4.1.15]. By Statement $B(c-1)$ and Eq. (39), since $R \in \mathcal{S}_{c-1}$ and $|R| < |T|$, we have $Y_R \perp X_T$ under $\mu_{\{T,R\}}$, so $\mathbb{E}_{\mu_{T|R}} f = 0$ after projecting from X_T to Y_R, and Eq. (45) vanishes. Since $f \in X_T$ and $g \in Y_S$ are arbitrary, $X_T \perp Y_S$ under $(\cdot,\cdot)_d$.

$A(c) \Longrightarrow B(c)$:

Claim 11.17. *$X_Q \perp X_T$ under $(\cdot,\cdot)_d$ for any $Q \in \mathcal{S}_c$ and any $T \in \mathcal{S}_d$ such that $|Q| < |T|$.*

Proof. If Q is embeddable into T, then $X_T \perp Y_{\mathcal{A}^*(\mathcal{T})}$ under $\mu_{\mathcal{A}(T)}$ where $\mathcal{T} \in \Gamma$ and $T \in \mathcal{T}$, which implies $X_T \perp X_Q$ under $(\cdot,\cdot)_d$ by Eq. (39) since $Q \in \mathcal{A}^*(\mathcal{T})$ and $X_Q \subseteq Y_Q \subseteq Y_{\mathcal{A}^*(\mathcal{T})}$.

If Q is not embeddable into T, by Statement $A(c)$, $X_Q \perp Y_T \supseteq X_T$ under $(\cdot,\cdot)_d$. □

R belongs to a unique component $\mathcal{R} \in \Gamma$. By Lemma 11.13,

$$Y_R \subseteq Y_{\mathcal{A}(\mathcal{R})} \overset{(44)}{=} \sum_{\mathcal{Q} \in \Gamma(\mathcal{R})} X_{\mathcal{Q}} = \sum_{\mathcal{Q} \in \Gamma(\mathcal{R})} \sum_{Q \in \mathcal{Q}} X_Q = \sum_{Q \in \mathcal{A}(R)} X_Q. \tag{46}$$

For $Q \in \mathcal{A}(R)$, $|Q| \leqslant |R| < |T|$, and the previous Claim implies

$$X_T \perp \left(\sum_{Q \in \mathcal{A}(R)} X_Q \right) \overset{(46)}{\supseteq} Y_R \quad \text{under } (\cdot,\cdot)_d. \qquad \square$$

Corollary 11.18. *$X_{\mathcal{Q}} \perp X_{\mathcal{R}}$ under $(\cdot,\cdot)_d$ for distinct $\mathcal{Q}, \mathcal{R} \in \Gamma$.*

Proof. Since $\mathcal{Q}$ and $\mathcal{R}$ are distinct, assume without loss of generality every $T \in \mathcal{R}$ is not embeddable into any $S \in \mathcal{Q}$. For any $S \in \mathcal{Q}, T \in \mathcal{R}$, $X_S \perp X_T$ under $(\cdot,\cdot)_d$ by Proposition 11.16 (Statement $A(d)$). The desired result follows by taking the span over $S \in \mathcal{Q}$ and $T \in \mathcal{R}$. $\square$

Our orthogonal decomposition Eq. (38) now follows from Corollaries 11.14 and 11.18. The local orthogonal decomposition Eq. (37) also follows from Lemma 11.13 and Corollary 11.18.

11.4. **SDP solution.**

By Eqs. (38), (41) and (43), the vector space $Y_{\mathcal{S}_d}$ equipped with the bilinear form $(\cdot,\cdot)_d$ equals the direct sum $\mathcal{X} \overset{\text{def}}{=} \bigoplus_{\mathcal{Q} \in \Gamma} X_{\mathcal{Q}}$ of inner product spaces, with the inner product

$$\langle F, G \rangle_{\mathcal{X}} \overset{\text{def}}{=} \sum_{\mathcal{Q} \in \Gamma} \langle F_{\mathcal{Q}}, G_{\mathcal{Q}} \rangle_{\mathcal{Q}} \overset{(35)}{=} \sum_{\mathcal{Q} \in \Gamma} \mathop{\mathbb{E}}_{\mu_{\mathcal{Q}}} [F_{\mathcal{Q}} G_{\mathcal{Q}}], \tag{47}$$

whenever $F = \bigoplus_{\mathcal{Q} \in \Gamma} F_{\mathcal{Q}}$ and $G = \bigoplus_{\mathcal{Q} \in \Gamma} G_{\mathcal{Q}}$, where $F_{\mathcal{Q}}, G_{\mathcal{Q}} \in X_{\mathcal{Q}}$.

Under Assumption 11.8, $X_{\mathcal{Q}}$ is a genuine inner product space for every $\mathcal{Q} \in \Gamma$ (after identifying functions whose difference have zero norm), and hence so is $\mathcal{X}$.

Theorem 11.19. *Under Assumption 11.8, for any $S, T \in \mathcal{S}_d$, any $f \in Y_S$, $g \in Y_T$,*

$$\langle f, g \rangle_{\mathcal{X}} = \mathop{\mathbb{E}}_{\mu_{\{S,T\}}} [fg].$$

Proof. Let $\mathcal{Q}$ be the equivalence class of S, and similarly $\mathcal{R}$ be that of T. By Eq. (37), expand $f = \sum_{\mathcal{A} \in \Gamma(\mathcal{Q})} f_{\mathcal{A}}$ and $g = \sum_{\mathcal{B} \in \Gamma(\mathcal{R})} g_{\mathcal{B}}$, where $f_{\mathcal{A}} \in X_{\mathcal{A}}$ and $g_{\mathcal{B}} \in X_{\mathcal{B}}$. Then

$$\begin{aligned} \langle f, g \rangle_{\mathcal{X}} &\overset{(47)}{=} \sum_{\mathcal{A} \in \Gamma(\mathcal{Q}) \cap \Gamma(\mathcal{R})} \langle f_{\mathcal{A}}, g_{\mathcal{A}} \rangle_{\mathcal{A}} \overset{(41),(43)}{=} \sum_{\mathcal{A} \in \Gamma(\mathcal{Q}) \cap \Gamma(\mathcal{R})} (f_{\mathcal{A}}, g_{\mathcal{A}})_d \\ &\overset{(38)}{=} \sum_{\substack{\mathcal{A} \in \Gamma(\mathcal{Q}) \\ \mathcal{B} \in \Gamma(\mathcal{R})}} (f_{\mathcal{A}}, g_{\mathcal{B}})_d \overset{(40)}{=} (f, g)_d \overset{(39)}{=} \mathop{\mathbb{E}}_{\mu_{\{S,T\}}} [fg]. \end{aligned}$$ $\square$

11.5. **Small ancestor closure.**

In this subsection, we justify Assumption 11.8 by Corollary 11.25. The next lemma ([KMOW17, Lemma 6.12] rephrased) says that if $\mathcal{C}'$ is obtained by adding constraints to $\mathcal{C}$, then the sparsity can only increase by at most the number of new boundary variables in $\mathcal{C}'$.

Lemma 11.20. *Let $\mathcal{C} \subseteq \mathcal{C}'$ be constraint sets, so that $\mathcal{C}$ is S-closed, $\mathcal{C}'$ is S'-closed, and $S' \subseteq S \cup V(\mathcal{C}')$. Then $\operatorname{sp}(\mathcal{C}', S') \leqslant \operatorname{sp}(\mathcal{C}, S) + |S' \cap B(\mathcal{C}') \setminus (S \cup \overline{B}(\mathcal{C}))|$.*

Proof. Expand $\operatorname{sp}(\mathcal{C}, S)$ and $\operatorname{sp}(\mathcal{C}', S')$ using Eq. (26). Every $C \in \mathcal{C}$ or $v \in S \cup \overline{B}(\mathcal{C})$ contributes no more to $\operatorname{sp}(\mathcal{C}', S')$ than to $\operatorname{sp}(\mathcal{C}, S)$, because it belongs to no fewer incidences in $J(\mathcal{C}', S')$ than in $J(\mathcal{C}, S)$. Since $S' \subseteq S \cup V(\mathcal{C}')$, only new boundary variables $v \in S' \cap B(\mathcal{C}') \setminus (S \cup \overline{B}(\mathcal{C}))$ contribute more to $\operatorname{sp}(\mathcal{C}', S')$ than to $\operatorname{sp}(\mathcal{C}, S)$, in which case $\operatorname{sp}_{\mathcal{C}'}(v) = 1$ by Remark 8.2. $\square$

Given $\mathcal{Q} \subseteq \mathcal{P}(V)$, shorthand $\mathcal{C}_{\mathcal{Q}} \overset{\text{def}}{=} \operatorname{cl}^t_{\mathcal{Q}}(\mathcal{C})$ and $\operatorname{sp}(\mathcal{C}, \mathcal{Q}) \overset{\text{def}}{=} \operatorname{sp}(\mathcal{C}, \bigcup \mathcal{Q})$. Also write $\mathcal{C}_S \overset{\text{def}}{=} \operatorname{cl}^t_S(\mathcal{C})$ for $S \subseteq V$.

Lemma 11.21. *For any $(2t, 3\gamma)$-expanding constraint set $\mathcal{C}$, any $t\gamma$-small $S \subseteq V$, we have $\mathcal{C}_{\mathcal{A}(S)}$ is t-small and $\mathrm{sp}(\mathcal{C}_{\mathcal{A}(S)}, \mathcal{A}(S)) \leqslant 2|S|$.*

Proof. Denote by $G_d[\mathcal{Q}]$ the directed subgraph of G_d induced by $\mathcal{Q}$ (see Remark 11.4). Call a subfamily $\mathcal{Q} \subseteq \mathcal{A}(S)$ *S-connected* if every $R \in \mathcal{Q}$ has a path p to S in $G_d[\mathcal{Q}]$ (i.e. $p \subseteq G_d[\mathcal{Q}]$). Note that $\mathcal{A}(S)$ is itself S-connected. For any S-connected $\mathcal{Q}$ containing S, we prove by induction on $\mathcal{Q}$ that $\mathcal{C}_\mathcal{Q}$ is t-small and $\mathrm{sp}(\mathcal{C}_\mathcal{Q}, \mathcal{Q}) \leqslant 2|S| \leqslant 2t\gamma$.

Base Case: $\mathcal{Q} = \{S\}$. Then $\mathcal{C}_\mathcal{Q} = \mathrm{cl}^t_S(\mathcal{C})$. $\mathrm{sp}(\mathcal{C}_\mathcal{Q}, S) \leqslant 2|S|$ by Remark 8.3, and $\mathcal{C}_\mathcal{Q}$ is t-small by Lemma 8.11 and $t\gamma$-smallness of S.

Induction Step: $\mathcal{Q} \supsetneq \{S\}$. Fix any $T \in \mathcal{Q}$ whose distance to S in $G_d[\mathcal{Q}]$ is largest. Then $\mathcal{R} \stackrel{\text{def}}{=} \mathcal{Q} \setminus \{T\}$ is also S-connected, because every node in $\mathcal{R}$ has a path to S in $G_d[\mathcal{Q}]$ without going through T. T also has a path to S in $G_d[\mathcal{Q}]$, and the node R after T on this path belongs to $\mathcal{R}$. Therefore $T \rightarrowtail R \in \mathcal{R}$.

By Induction Hypothesis and Lemma 11.22 (with its $R \stackrel{\text{def}}{=} \bigcup \mathcal{R}$), $\mathcal{C}_\mathcal{Q}$ is t-small. By Lemma 11.23 and Induction Hypothesis, $\mathrm{sp}(\mathcal{C}_\mathcal{Q}, \mathcal{Q}) \leqslant \mathrm{sp}(\mathcal{C}_\mathcal{R}, \mathcal{R}) \leqslant 2|S|$. □

The next two lemmas are based on [KMOW17, Claims 6.15 and 6.17].

Lemma 11.22. *Suppose $\mathcal{C}$ is a $(2t, 3\gamma)$-expanding constraint set, $R \subseteq V$ and $T \in \mathcal{S}_{t\gamma}$. Let $Q \stackrel{\text{def}}{=} R \cup T$. If $\mathcal{C}_R$ is t-small and $\mathrm{sp}(\mathcal{C}_R, R) \leqslant 2t\gamma$, then $\mathcal{C}_Q$ is also t-small.*

Proof. Start with $\mathcal{C}'_0 \stackrel{\text{def}}{=} \mathcal{C}_R$. Keep adding Q-closed t-small $\mathcal{C}_i \subseteq \mathcal{C}$ to $\mathcal{C}'_{i-1}$ to get $\mathcal{C}'_i \stackrel{\text{def}}{=} \mathcal{C}'_{i-1} \cup \mathcal{C}_i$, until $\mathcal{C}'_r = \mathcal{C}_Q$. We prove by induction on i that $\mathcal{C}'_i$ is t-small.

Base Case: $i = 0$. Then $\mathcal{C}'_0 = \mathcal{C}_R$ is t-small by assumption.

Induction Step: $i > 0$. Let $R_i \stackrel{\text{def}}{=} R \cup (T \cap V(\mathcal{C}'_i))$. $\mathcal{C}'_i$ is R_i-closed because $\mathcal{C}'_i$ is Q-closed and $Q \cap V(\mathcal{C}'_i) \subseteq R_i$. By Lemma 11.20, $\mathrm{sp}(\mathcal{C}'_i, R_i) \leqslant \mathrm{sp}(\mathcal{C}_R, R) + |T \cap B(\mathcal{C}'_i) \setminus (R \cup \overline{B}(\mathcal{C}_R))| \leqslant 3t\gamma$. Further, $\mathcal{C}'_{i-1}$ is t-small by Induction Hypothesis, so $\mathcal{C}'_i$ is $2t$-small. Thus $\mathcal{C}'_i$ is t-small because $\mathcal{C}$ is $(2t, 3\gamma)$-expanding. □

Lemma 11.23. *Under the assumptions of the previous lemma, if T is embeddable into R, then $\mathrm{sp}(\mathcal{C}_Q, Q) \leqslant \mathrm{sp}(\mathcal{C}_R, R)$.*

Proof. Let $R' \stackrel{\text{def}}{=} R \cup \overline{B}(\mathcal{C}_R)$ be the old variables and $T' \stackrel{\text{def}}{=} T \setminus R'$ be the new variables in T. Lemma 11.20 implies $\mathrm{sp}(\mathcal{C}_Q, Q) \leqslant \mathrm{sp}(\mathcal{C}_R, R) + |T'|$. Since $T \rightarrowtail R$, there are vertex-disjoint paths p_v from every $v \in T$ to R in $\mathrm{cl}^t_Q(I)$.

Treat a path p from T' to R as a sequence $(v_1, C_1, v_2, C_2, \ldots, v_s)$ of distinct vertices interleaved with constraints, so that consecutive variables v_i and v_{i+1} both belong to the constraint C_i between them, for every $1 \leqslant i < s$. The first variable v_1 is new and the last v_s is old, so p contains a new element y whose successor z is old (i.e. $z \in R \cup \overline{B}(\mathcal{C}_R) \cup \mathcal{C}_R$, but not for y). The unordered pair (y, z) is a new constraint-variable incidence in $J(\mathcal{C}_Q, Q)$ not in $J(\mathcal{C}_R, R)$. Further, exactly one of y or z is a variable.

Because the paths $\{p_v \mid v \in T'\}$ are vertex-disjoint (i.e. variable-disjoint), the above pairs $\{(y_v, z_v) \mid v \in T'\}$ from different paths are distinct new constraint-variable incidences to old constraints or old vertices. These $|T'|$ new incidences cancel out the $|T'|$ sparsity increase in Lemma 11.20, so $\mathrm{sp}(\mathcal{C}_Q, Q) \leqslant \mathrm{sp}(\mathcal{C}_R, R)$. □

Lemma 11.24. *For any t, any families of variable subsets $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{P}(V)$, any constraint set $\mathcal{C}$, $\mathrm{cl}^t_\mathcal{A}(\mathcal{C}) \subseteq \mathrm{cl}^t_\mathcal{B}(\mathcal{C})$.*

Proof. Apply Lemma 8.9 with its $T \stackrel{\text{def}}{=} \bigcup \mathcal{A}$ and $S \stackrel{\text{def}}{=} \bigcup \mathcal{B}$. □

Corollary 11.25. *For any $(2t, 3\gamma)$-expanding constraint set $\mathcal{C}$, any $S \subseteq V$ such that $S \cup K$ is $t\gamma$-small, the K-augmented ancestor closure $\underline{\mathrm{cl}}^t_{\mathcal{A}(S)}(\mathcal{C})$ is t-small.*

Proof. Note that $K \in \mathcal{A}(K)$. Also $\mathcal{A}(S) \subseteq \mathcal{A}(T)$ for any $S \subseteq T \subseteq V$ because $S \rightarrowtail T$ via paths with no edges. These two facts imply $\{K\} \cup \mathcal{A}(S) \subseteq \mathcal{A}(K) \cup \mathcal{A}(S) \subseteq \mathcal{A}(K \cup S)$. The desired result now follows by Lemma 11.21 (with its $S \stackrel{\text{def}}{=} K \cup S$) and Lemma 11.24 (with its $\mathcal{A} \stackrel{\text{def}}{=} \{K\} \cup \mathcal{A}(S)$ and $\mathcal{B} \stackrel{\text{def}}{=} \mathcal{A}(K \cup S)$). □

11.6. **Wrapping up.**

Our next theorem generalizes [KMOW17, Theorem 1.2] (which concerns $K = \emptyset$ and $\mathcal{E} = \emptyset$):

Theorem 11.26. *Suppose an instance $I = (V, \mathcal{C})$ of a τ-wise uniform CSP is $(2t, 3\gamma)$-expanding. Fix $K \subseteq V$ and $\mathcal{E} \subseteq \mathcal{C}$ so that η is τ-wise uniform outside $\mathcal{E}$. Choose $d \in \mathbb{N}$ so that Assumption 11.8 holds. Define*

$$\mu_R \stackrel{\text{def}}{=} \pi_R \circ \mu \circ \underline{\text{cl}}^t_R(I) \qquad \text{for } R \in \mathcal{S}_{2d}; \text{ and} \tag{48}$$

$$\alpha_S(a_S) \stackrel{\text{def}}{=} \mathbb{1}_{a_S} \qquad \text{for } S \in \mathcal{S}_d, a_S \in A_S. \tag{49}$$

Then $(\alpha_S)_S$ is a level-d SDP hierarchy solution inducing the level-$2d$ LP hierarchy solution $(\mu_R)_R$.

Proof. μ_R is well defined because Assumption 11.8 implies $\underline{\text{cl}}^t_R(I)$ is η-satisfiable for $R \in \mathcal{S}_{2d}$. We now show that $(\mu_R)_R$ is consistent with projection. For any $R \in \mathcal{S}_{2d}$, $U \subseteq R$,

$$\begin{aligned}\pi_{R\to U} \circ \mu_R &\overset{(48)}{=} \pi_{R\to U} \circ \pi_R \circ \mu \circ \underline{\text{cl}}^t_R(I) \overset{(5)}{=} \pi_U \circ \mu \circ \underline{\text{cl}}^t_R(I) \\ &\overset{(*)}{=} \pi_U \circ \mu \circ \underline{\text{cl}}_U \circ \underline{\text{cl}}^t_R(I) \overset{(31)}{=} \pi_U \circ \mu \circ \underline{\text{cl}}^t_U(I) \overset{(48)}{=} \mu_U,\end{aligned}$$

where $(*)$ is Lemma 10.4.

For any $S, T \in \mathcal{S}_d$, any $a_S \in A_S$, $a_T \in A_T$, Theorem 11.19 implies Eq. (1) because

$$\langle \alpha_S(a_S), \alpha_T(a_T) \rangle_{\mathcal{X}} \overset{(49)}{=} \langle \mathbb{1}_{a_S}, \mathbb{1}_{a_T} \rangle_{\mathcal{X}} = \mathop{\mathbb{E}}_{\mu_{\{S,T\}}} [\mathbb{1}_{a_S} \mathbb{1}_{a_T}] = \mathop{\mathbb{P}}_{b \sim \mu_{S \cup T}} [b_S = a_S, b_T = a_T].$$

Finally, $(\alpha_S)_S$ is consistent with projection, because for any $S \in \mathcal{S}_d$, $T \subseteq S$, $b \in D^T$,

$$\pi_{S\to T}(\alpha_S)(b) \overset{(2)}{=} \sum_{\substack{a \in D^S \\ a_T = b}} \alpha_S(a) \overset{(49)}{=} \sum_{\substack{a \in D^S \\ a_T = b}} \mathbb{1}_a = \mathbb{1}_b \overset{(49)}{=} \alpha_T(b).$$ □

Lemma 11.27. *When $K = \emptyset$ and $\mathcal{E} = \emptyset$, Assumption 11.8 holds for $d \stackrel{\text{def}}{=} t\gamma/2$ and any $(2t, 3\gamma)$-expanding constraint set $\mathcal{C}$.*

Proof. Lemma 11.21 ($\underline{\text{cl}}^t_{\mathcal{A}(S)}(\mathcal{C})$ is t-small), Lemma 8.12 ($\underline{\text{cl}}^t_{\mathcal{A}(S)}(\mathcal{C})$ is dismissible) and Proposition 10.7 ($\underline{\text{cl}}^t_{\mathcal{A}(S)}(\mathcal{C})$ is η-satisfiable) imply Assumption 11.8. □

12. New SDP solution with full support

We now construct a new SDP solution for a τ-wise uniform CSP, completing the proof of Theorem 1.2. Our SDP solution induces an LP solution that has full support on *all satisfying assignments* of every small closure.

Definition 12.1 ($\mathcal{E}$-planted solution)**.** Given constraint sets $\mathcal{E} \subseteq \mathcal{C}$, define the constraint solution $\eta^{\mathcal{E}} : (C \in \mathcal{C}) \to \Delta(A_C)$ as follows: $\eta^{\mathcal{E}}(C)$ is the uniform distribution over A_C for $C \in \mathcal{E}$, and $\eta^{\mathcal{E}}(C)$ is a τ-wise uniform distribution supported on A_C for $C \in \mathcal{C} \setminus \mathcal{E}$.

Definition 12.2. A relaxed assignment $\alpha : D^V \to \mathcal{M}$ *has full support* on $A \subseteq D^V$ if $\text{supp}(\alpha) = A$. A scheme σ for a family $\boldsymbol{\mathcal{C}}$ of constraints *has full support* if $\text{supp}(\sigma(\mathcal{C})) = A_{\mathcal{C}}$ for $\mathcal{C} \in \boldsymbol{\mathcal{C}}$.

Lemma 12.3. *Let $\eta : (C \in \mathcal{C}) \to \Delta(A_C)$ be a family of satisfying assignments for a constraint set $\mathcal{C}$. Suppose $\eta(C)$ has full support on A_C for every $C \in \mathcal{C}$. Then the canonical distribution μ_I derived from η has full support on A_J for any instance $J = (U, \mathcal{C})$.*

Proof. For every $C \in \mathcal{C}$, $\eta(C)$ has full support on A_C, so $\eta(C)(b) > 0$ for every $b \in A_C$. Now an assignment $\alpha \in A_J$ satisfies all constraints in $\mathcal{C}$, so the canonical distribution $\mu_{\mathcal{C}}$ has probability mass $\mu_{\mathcal{C}}(a) = \frac{1}{Z}\prod_{C\in\mathcal{C}} \eta(C)\left(a_{V(C)}\right) > 0$, where $Z > 0$ is the normalization factor to make $\mu_{\mathcal{C}}$ sum to 1. This implies $\operatorname{supp}(\mu_J) \supseteq A_J$.

On the other hand, if an assignment $a \in D^{V(J)} \setminus A_J$ violates some constraint $C \in \mathcal{C}$, then $a_{V(C)} \notin A_C$ and $\eta(C)\left(a_{V(C)}\right) = 0$, so $\mu_{\mathcal{C}}(a) = \frac{1}{Z}\prod_{C\in\mathcal{C}} \eta(C)\left(a_{V(C)}\right) = 0$. Thus $\operatorname{supp}(\mu_J) \subseteq A_J$. □

Lemma 12.4. *Suppose an instance $I = (V, \mathcal{C})$ of a τ-wise uniform CSP is $(2t, 3\gamma)$-expanding. For any $t\gamma/2$-small $K \subseteq V$, let $\mathcal{E} \stackrel{\text{def}}{=} \mathrm{cl}^t_K(\mathcal{C})$ and $\eta^{\mathcal{E}}$ be the $\mathcal{E}$-planted solution. Then the K-augmented closure $\underline{\mathrm{cl}}^t_{\mathcal{A}(S)}(\mathcal{C})$ is t-small and $\eta^{\mathcal{E}}$-satisfiable for every $t\gamma/2$-small $S \subseteq V$.*

Proof. Corollary 11.25 implies $\underline{\mathrm{cl}}^t_{\mathcal{A}(S)}(\mathcal{C})$ is t-small because $(K \cup S)$ is t-small. By Lemma 8.12, $\underline{\mathrm{cl}}^t_{\mathcal{A}(S)}(\mathcal{C})$ is dismissible with respect to τ. Since the CSP is τ-wise uniform, every $C \in \mathcal{C}$ has a τ-wise uniform distribution η^C supported on its satisfying assignments A_C, so $\operatorname{supp}(\eta^{\mathcal{E}}(C)) \supseteq \operatorname{supp}(\eta^C)$ is τ-wise neutral for BW. By Proposition 10.6 (with $\mathcal{C} \stackrel{\text{def}}{=} \underline{\mathrm{cl}}^t_{\mathcal{A}(S)}(\mathcal{C}), \eta \stackrel{\text{def}}{=} \eta^{\mathcal{E}}$), $\underline{\mathrm{cl}}^t_{\mathcal{A}(S)}(\mathcal{C})$ is $\eta^{\mathcal{E}}$-satisfiable. □

Given an instance I, denote by Σ^{SDP}_d the set of its SDP hierarchy solutions of level d. More precisely, Σ^{SDP}_d is the set of $((\alpha_S)_S, (\mu_R)_R)$, where $(\alpha_S)_S$ is an SDP hierarchy solution for I of level d inducing the LP hierarchy solution $(\mu_R)_R$ of level $2d$.

Lemma 12.5. *Let I be an instance and $d \in \mathbb{N}$. Let ρ be a distribution over a finite set Ω. Suppose for every $\mathcal{E} \in \Omega$, there is $((\alpha^{\mathcal{E}}_S)_S, (\mu^{\mathcal{E}}_R)_R) \in \Sigma^{\mathrm{SDP}}_d$. Then $((\alpha_S)_S, (\mu_R)_R) \in \Sigma^{\mathrm{SDP}}_d$, where*

$$\mu_R \stackrel{\text{def}}{=} \mathop{\mathbb{E}}_{\mathcal{E}\sim\rho} \mu^{\mathcal{E}}_R \qquad \text{for } 2d\text{-small } R \subseteq V\text{; and} \tag{50}$$

$$\alpha_S \stackrel{\text{def}}{=} \bigoplus_{\mathcal{E}\in\Omega} \sqrt{\rho(\mathcal{E})} \cdot \alpha^{\mathcal{E}}_S \qquad \text{for } d\text{-small } S \subseteq V. \tag{51}$$

Proof. For $S, T \in \mathcal{S}_d$, $a_S \in A_S, a_T \in A_T$,

$$\begin{aligned}\langle \alpha_S(a_S), \alpha_T(a_T)\rangle_\rho &\overset{(51)}{=} \sum_{\mathcal{E}\in\Omega} \left\langle \sqrt{\rho(\mathcal{E})}\alpha^{\mathcal{E}}_S(a_S), \sqrt{\rho(\mathcal{E})}\alpha^{\mathcal{E}}_T(a_T)\right\rangle_{\mathcal{X}^{\mathcal{E}}} \overset{(*)}{=} \mathop{\mathbb{E}}_{\mathcal{E}\sim\rho} \left\langle \alpha^{\mathcal{E}}_S(a_S), \alpha^{\mathcal{E}}_T(a_T)\right\rangle_{\mathcal{X}^{\mathcal{E}}} \\ &\overset{(\star)}{=} \mathop{\mathbb{E}}_{\mathcal{E}\sim\rho} \mathop{\mathbb{E}}_{b\sim\mu^{\mathcal{E}}_{\{S,T\}}} [b_S = a_S, b_T = a_T] \overset{(50)}{=} \mathop{\mathbb{E}}_{b\sim\mu_{\{S,T\}}} [b_S = a_S, b_T = a_T],\end{aligned}$$

where $(*)$ uses bilinearity of $\langle\cdot,\cdot\rangle_{\mathcal{X}^{\mathcal{E}}}$, and $(\star)$ is because $(\alpha^{\mathcal{E}}_S)_S$ induces $(\mu^{\mathcal{E}}_R)_R$ for $\mathcal{E} \in \Omega$. Therefore $(\mu_R)_R$ and $(\alpha_S)_S$ satisfy Eq. (1).

$(\alpha_S)_S$ is consistent with projection, because for any $T \subseteq S \subseteq V$,

$$\pi_T(\alpha_S) \overset{(51)}{=} \pi_T\left(\bigoplus_{\mathcal{E}\in\Omega} \sqrt{\rho(\mathcal{E})} \cdot \alpha^{\mathcal{E}}_S\right) \overset{(58)}{=} \bigoplus_{\mathcal{E}\in\Omega} \sqrt{\rho(\mathcal{E})} \cdot \pi_T\left(\alpha^{\mathcal{E}}_S\right) \overset{(\dagger)}{=} \bigoplus_{\mathcal{E}\in\Omega} \sqrt{\rho(\mathcal{E})} \cdot \alpha^{\mathcal{E}}_T \overset{(51)}{=} \alpha_T,$$

where $(\dagger)$ is due to $(\alpha^{\mathcal{E}}_S)_S$ being consistent with projection for $\mathcal{E} \in \Omega$. □

Theorem 12.6. *Suppose an instance $I = (V, \mathcal{C})$ of a τ-wise uniform CSP is $(2t, 3\gamma)$-expanding. Let $d \stackrel{\text{def}}{=} t\gamma/4$. Then there exists $((\alpha_S)_S, (\mu_R)_R) \in \Sigma^{\mathrm{SDP}}_d$ such that $(\mu_R)_R$ has full support.*

Proof. Let $\Omega \stackrel{\text{def}}{=} \{\mathrm{cl}^t_K(\mathcal{C}) \mid K \in \mathcal{S}_{2d}\}$ be the family of closures of $2d$-small variable subsets. For every $K \in \mathcal{S}_{2d}, \mathcal{E} \stackrel{\text{def}}{=} \mathrm{cl}^t_K(\mathcal{C}) \in \Omega$, the constraint solution $\eta^{\mathcal{E}}$ is τ-wise uniform outside $\mathcal{E}$, and

the K-augmented closure $\underline{\mathrm{cl}}^t_\emptyset(\mathcal{C}) = \mathrm{cl}^t_K(\mathcal{C})$ contains (and in fact equals) $\mathcal{E}$. Theorem 11.26 yields $((\alpha^{\mathcal{E}}_S)_S, (\mu^{\mathcal{E}}_R)_R) \in \Sigma^{\mathrm{SDP}}_d$, where Assumption 11.8 is justified by Lemma 12.4.

Lemma 12.5 further constructs $((\alpha_S)_S, (\mu_R)_R) \in \Sigma^{\mathrm{SDP}}_d$, with ρ being any distribution with full support on Ω, such as the uniform distribution over Ω.

For $2d$-small $R \subseteq V$, consider the K-augmented closure with $K \stackrel{\mathrm{def}}{=} R$. Let $J \stackrel{\mathrm{def}}{=} \mathrm{cl}^t_R(I) = \underline{\mathrm{cl}}^t_R(I), \mathcal{E} \stackrel{\mathrm{def}}{=} \mathrm{cl}^t_R(\mathcal{C}) = \underline{\mathrm{cl}}^t_R(\mathcal{C})$ and $\mu^{\mathcal{E}}_J$ be the canonical distribution on J derived from $\eta^{\mathcal{E}}$. Then

$$\mathrm{supp}(\mu_R) \stackrel{(50)}{=} \bigcup_{\mathcal{E}' \in \Omega} \mathrm{supp}\left(\mu^{\mathcal{E}'}_R\right) \supseteq \mathrm{supp}\left(\mu^{\mathcal{E}}_R\right) \stackrel{(48)}{=} \mathrm{supp}\left(\pi_R\left(\mu^{\mathcal{E}}_J\right)\right) \stackrel{(57)}{=} \pi_R\left(\mathrm{supp}\left(\mu^{\mathcal{E}}_J\right)\right) = \pi_R(A_J),$$

where the last equality is Lemma 12.3 (with its $\mathcal{C} \stackrel{\mathrm{def}}{=} \mathcal{E}, \eta \stackrel{\mathrm{def}}{=} \eta^{\mathcal{E}}$). Therefore $(\mu_R)_R$ satisfies Definition 8.19 and has full support on small closures. □

We now prove Theorem 1.2.

Theorem 1.2. *Let $\tau \geqslant 2$. If a k-CSP is τ-wise neutral for both SDP and AIP (separately), then except with probability $o_{n;k}(1)$, a random instance of the CSP with n variables and Δn constraints has an SDP+AIP hierarchy solution of level $\Omega_k(n/(\Delta^{2/(\tau-1)} \log \Delta))$.*

Proof. Fix $\gamma \stackrel{\mathrm{def}}{=} 1/\log\Delta, \lambda \stackrel{\mathrm{def}}{=} \tau - 1$. We may assume $\Delta^{2/\lambda} \log \Delta \leqslant n$, for otherwise the level lower bound is trivial, as every hierarchy has a solution of level 0. Since $\Delta^{2/(\lambda - 3/\log\Delta)} = \Theta(\Delta^{2/\lambda})$ for $\lambda \geqslant 1$, we have $\zeta \stackrel{\mathrm{def}}{=} \Delta^{2/(\lambda-3\gamma)}/n = o_n(1)$. Lemma 8.16 implies that except with probability $o_{n;k}(1)$, a random instance I is $(2t, 3\gamma)$-expanding, where $t = n/(\Delta^{2/(\lambda-3\gamma)} 2^{O(k)}) = \Omega_k(n/\Delta^{2/\lambda})$.

Theorem 12.6 yields a level-d SDP solution $(\alpha_S)_S$ inducing the level-$2d$ LP solution $(\mu_R)_R$ such that $(\mu_R)_R$ has full support, where $d \stackrel{\mathrm{def}}{=} t\gamma/4$. Theorem 7.9 and Proposition 8.15 constructs a level-$2d$ AIP solution $(w_R)_R$. Lemma 8.20 implies that $(\mu_R, w_R)_R$ is a level-$2d$ solution for the LP+AIP hierarchy. □

Recent works constructed explicit (non-random) instances fooling the SDP hierarchy. A random CSP instance has two sources of randomness (Definition B.1):

(A) Randomness of the underlying hypergraph
(B) Randomness of the relation (set of satisfying assignments) of each constraint

To fool a hierarchy, randomness (A) ensures expansion, crucial for hierarchy solution. As explained in Remark 8.5, when $\tau = k - 1$ (e.g. k-SAT), hierarchy solutions can be constructed as long as an instance has a factor graph that is a boundary expander. Randomness (B) ensures every assignment is far from satisfying all constraints. Hopkins and Lin [HL22], building on [DFHT21], constructed explicit 3-XOR instances fooling linear-level SDP using high dimensional expanders known as quantum Tanner codes [LZ22]. They further showed that every assignment to their instances violates a fraction of 3-XOR constraints bounded away from zero. Their analysis crucially exploits the linear structure over $\mathbb{F}_2$ that appears naturally for 3-XOR but not for 3-SAT. It is not clear how to generalize their analysis to construct explicit 3-SAT instances fooling linear-level LP+AIP and SDP+AIP. This seems to call for constructing high dimensional expanders from non-linear codes.

Acknowledgements. We thank Thana Somsirivattana and Mehrshad Taziki for participating in part of this research project and their observations. We thank Jeff Xu and Sirawit Pongnakintr for helpful discussions. We thank anonymous STOC reviewers for their valuable feedback and suggestions.

Appendix A. Omitted proofs

Lemma A.1. *For any constraint sets $\mathcal{C}' \subseteq \mathcal{C}$, any $C \in \mathcal{C}'$,*

$$B(\mathcal{C}) \cap V(C) \subseteq B(\mathcal{C}') \cap V(C).$$

Proof. Every $v \in B(\mathcal{C}) \cap V(C)$ belongs to $C \in \mathcal{C}'$ but not to any $C' \in \mathcal{C} \setminus \{C\}$, in particular no $C' \in \mathcal{C}' \setminus \{C\} \subseteq \mathcal{C} \setminus \{C\}$ contains v. Since $C \in \mathcal{C}'$, v is contained in a unique constraint in $\mathcal{C}'$. □

Proposition A.2. *For any $T \subsetneq S$, $\pi_T(\chi_S) \equiv 0$.*

Proof. For any $a \in \{0,1\}^T$,

$$\pi_T(\chi_S)(a) = \sum_{b \in D^{S \setminus T}} \chi_S(a \cup b) = \chi_T(a) \sum_{b \in D^{S \setminus T}} \chi_{S \setminus T}(b) = 0,$$

as half of the assignments $b \in \{0,1\}^{S \setminus T}$ have even parity and the other half have odd parity. □

Lemma A.3. *Suppose $\mathcal{M}$ is a commutative semiring. For any disjoint S and T, $\alpha_S \in \mathcal{M}^{D^S}$, $\alpha_T \in \mathcal{M}^{D^T}$, $R \subseteq S \cup T$,*

$$\pi_{S \cup T \to R}(\alpha_S \otimes \alpha_T) = \pi_{S \to S \cap R}(\alpha_S) \otimes \pi_{T \to T \cap R}(\alpha_T). \tag{52}$$

Proof. For every $a \in D^R$,

$$\begin{aligned}
\pi_{S \cup T \to R}(\alpha_S \otimes \alpha_T)(a) &\overset{(2)}{=} \sum_{\substack{b \in S, c \in T \\ (b \cup c)_R = a}} (\alpha_S \otimes \alpha_T)(b \cup c) \overset{(6)}{=} \sum_{\substack{b \in S, c \in T \\ (b \cup c)_R = a}} \alpha_S(b) \alpha_T(c) \\
&\overset{(*)}{=} \sum_{\substack{b \in S \\ b_{R \cap S} = a_{R \cap S}}} \alpha_S(b) \sum_{\substack{c \in T \\ c_{R \cap T} = a_{R \cap T}}} \alpha_T(c) \\
&\overset{(2)}{=} \pi_{S \to S \cap R}(\alpha_S)(a_{R \cap S}) \pi_{T \to T \cap R}(\alpha_T)(a_{R \cap T}) \\
&\overset{(6)}{=} (\pi_{S \to S \cap R}(\alpha_S) \otimes \pi_{T \to T \cap R}(\alpha_T))(a),
\end{aligned}$$

where $(*)$ follows from the distributivity of multiplication over addition in $\mathcal{M}$. □

Lemma A.4. *For any variable subsets $U_1, U_2, R_1, R_2 \subseteq V$ such that $U_i \subseteq R_i$, any $b \in D^V$,*

$$\mathbb{1}^{R_1}_{b_{U_1}} \tilde{\otimes} \mathbb{1}^{R_2}_{b_{U_2}} = \mathbb{1}^{R}_{b_U}, \qquad \text{where } R \overset{\text{def}}{=} R_1 \cup R_2 \text{ and } U \overset{\text{def}}{=} U_1 \cup U_2. \tag{53}$$

Proof. For $a \in D^R$, if $a_U \neq b_U$, then $a_{U_i} \neq b_{U_i}$ for some i, and $\mathbb{1}^{R_i}_{b_{U_i}}(a_{R_i}) = 0$, so $\left(\widetilde{\bigotimes}_i \mathbb{1}^{R_i}_{b_{U_i}}\right)(a) = 0 = \mathbb{1}^R_b(a)$. If $a_U = b_U$, then $\mathbb{1}^{R_i}_{b_{U_i}}(a_{R_i}) = 1$ for every i, and $\left(\widetilde{\bigotimes}_i \mathbb{1}^{R_i}_{b_{U_i}}\right)(a) = 1 = \mathbb{1}^R_b(a)$. □

Lemma A.5. *For any constraint set $\mathcal{C}$, any $\mathcal{C}' \subseteq \overline{R}_S(\mathcal{C})$,*

$$\mathrm{cl}_S(\mathcal{C} \setminus \mathcal{C}') = \mathrm{cl}_S(\mathcal{C}). \tag{54}$$

Proof. $\mathrm{cl}_S(\mathcal{C} \setminus \mathcal{C}') \subseteq \mathrm{cl}_S(\mathcal{C})$ by Lemma 5.3.

Also, $\mathcal{C} \setminus \mathcal{C}' \supseteq R_S(\mathcal{C}) \supseteq \mathrm{cl}_S(R_S(\mathcal{C})) \overset{(*)}{=} R^*_S(R_S(\mathcal{C})) \overset{(10)}{=} R^*_S(\mathcal{C}) = \mathrm{cl}_S(\mathcal{C})$, where $(*)$ is Proposition 5.4. Since $\mathrm{cl}_S(\mathcal{C})$ is S-closed, we have $\mathrm{cl}_S(\mathcal{C}) \subseteq \mathrm{cl}_S(\mathcal{C} \setminus \mathcal{C}')$. □

Lemma A.6. *For any subsets $T \subseteq S$, any subsets $B \subseteq A \subseteq D^S$,*

$$\pi_T(B) \subseteq \pi_T(A). \tag{55}$$

Proof. If $c \in \pi_T(B)$, then $b_T = c$ for some $b \in B$, and hence $b \in A$ and $c \in \pi_T(A)$. □

Lemma A.7. *For any $T \subseteq S$, any relaxed assignment $\alpha : D^S \to \mathcal{M}$,*

$$\mathrm{supp}(\pi_T(\alpha)) \subseteq \pi_T(\mathrm{supp}(\alpha)). \tag{56}$$

Proof. If $b \in \mathrm{supp}(\pi_T(\alpha))$, then $0 \neq \pi_T(\alpha)(b) = \sum\{\alpha(a) \mid a \in D^S, a_T = b\}$, so $\alpha(a) \neq 0$ for some $a \in D^S, a_T = b$. This implies $a \in \mathrm{supp}(\alpha)$ and $b \in \pi_T(\mathrm{supp}(\alpha))$. □

Lemma A.8. *For any $T \subseteq S$, any nonnegative-real-valued $\alpha : D^S \to \mathbb{R}_+$,*

$$\operatorname{supp}(\pi_T(\alpha)) = \pi_T(\operatorname{supp}(\alpha)). \tag{57}$$

Proof. For any $b \in D^T$, $b \in \pi_T(\operatorname{supp}(\alpha))$ if and only if $\alpha(a) > 0$ for some $a \in D^S$ such that $a_T = b$, which is equivalent to $\pi_T(\alpha)(b) > 0$, that is, $b \in \operatorname{supp}(\pi_T(\alpha))$. □

Lemma A.9. *For any subsets $T \subseteq S$ and family $\{\alpha^{\mathcal{E}} : D^S \to \mathcal{M} \mid \mathcal{E} \in \Omega\}$ of relaxed assignments,*

$$\pi_T\left(\bigoplus_{\mathcal{E}\in\Omega} \alpha^{\mathcal{E}}\right) = \bigoplus_{\mathcal{E}\in\Omega} \pi_T\left(\alpha^{\mathcal{E}}\right). \tag{58}$$

Proof. For any $b \in D^T$,

$$\pi_T\left(\bigoplus_{\mathcal{E}\in\Omega} \alpha^{\mathcal{E}}\right)(b) \overset{(2)}{=} \sum_{\substack{a\in D^S \\ a_T = b}} \left(\bigoplus_{\mathcal{E}\in\Omega} \alpha^{\mathcal{E}}\right)(a) \overset{(*)}{=} \bigoplus_{\mathcal{E}\in\Omega} \left(\sum_{\substack{a\in D^S \\ a_T = b}} \alpha^{\mathcal{E}}(a)\right) \overset{(2)}{=} \bigoplus_{\mathcal{E}\in\Omega} \pi_T\left(\alpha^{\mathcal{E}}\right)(b),$$

where $(*)$ is by definition of the direct sum construction, that is, coordinate-wise addition. □

Proposition A.10. *Let μ be a $(k-1)$-wise uniform distribution over $\{0,1\}^k$. Then $\mu = \lambda\mu_{even} + (1-\lambda)\mu_{odd}$ for some $0 \leqslant \lambda \leqslant 1$, where μ_{even} (μ_{odd}) is the uniform distribution over $\{0,1\}^k$ conditioned on even (odd) parity.*

Proof. Let χ_S be the parity function on S for $S \subseteq [k]$. $(k-1)$-wise uniformity is equivalent to the linear constraints that $\mathbb{E}_\mu[\chi_S] = 0$ for $\emptyset \subsetneq S \subsetneq [k]$. μ also satisfies the linear equality $\mathbb{E}_\mu[\chi_\emptyset] = 1$ that holds for any distribution. Since the parity functions $\{\chi_S \mid S \subseteq [k]\}$ form the Fourier basis of $\{f : \{0,1\}^k \to \mathbb{R}\}$, there is only one degree of freedom in μ, namely the choice of $\alpha \overset{\text{def}}{=} \mathbb{E}_\mu[\chi_{[k]}]$. We have $|\alpha| = |\mathbb{E}_\mu[\chi_{[k]}]| \leqslant \mathbb{E}_\mu[|\chi_{[k]}|] = 1$. Of the two extremes, $\alpha = 1$ is achieved by μ_{even} and $\alpha = -1$ by μ_{odd}. Therefore $\lambda \overset{\text{def}}{=} (1+\alpha)/2$. □

Proposition A.11. *For any $a \in \{0,1\}^3$, let $R \overset{\text{def}}{=} \{0,1\}^3 \setminus \{a\}$ be the set of satisfying assignments to a 3-SAT constraint forbidding only a. If a distribution $\mu \in \Delta(R)$ over R is pairwise uniform, then μ is the uniform distribution over $\{0,1\}^3$ conditioned on having opposite parity to a.*

Proof. The previous Proposition implies $\mu = \lambda\mu_{\text{even}} + (1-\lambda)\mu_{\text{odd}}$ for some $0 \leqslant \lambda \leqslant 1$. If a has even parity, then $\lambda = 0$, because $a \notin \operatorname{supp}(\mu)$. Likewise if a has odd parity, then $\lambda = 1$. □

Let us recall (vertex) Menger's theorem for graphs, and then prove its counterpart for hypergraphs.

Theorem A.12 (Menger, e.g. [G02])**.** *For any vertex subsets S and T in a multigraph G, the minimum size of an (S,T)-separator equals the maximum number of vertex-disjoint paths between S and T.*

Theorem A.13 (Hypergraph Menger)**.** *For any vertex subsets S and T in a hypergraph H, the minimum size of an (S,T)-seperator equals the maximum number of vertex-disjoint paths between S and T.*

Proof. Define a multigraph G having the same vertex set as H, and G has an edge $(u,v) \in V^2$ if some hyperedge in H contains both u and v. Apply Theorem A.12 to G. Note that for every sequence of vertices, it is a path in G if and only if it is a path in H. □

Appendix B. Random instance

Recall the natural distribution of random CSP instances, e.g. [BGMT12, KMOW17]:

Definition B.1 (Random instance)**.** Fix a k-CSP $(D, \mathfrak{R})$, a finite set V of variables, and $m \in \mathbb{N}$.

(1) A random constraint $C = (S, R)$ is chosen by picking a sequence $S \in V^k$ of k distinct variables uniformly, and R uniformly from $\mathfrak{R}$.
(2) A random instance consists of choosing with replacement m random constraints.

Equivalently, a random instance consists of first picking a k-uniform hypergraph whose hyperedges are chosen independently, and then choosing the relation R of each hyperedge (i.e. constraint) independently and uniformly at random from $\mathfrak{R}$.

B.1. Unsatisfiability.

As is well known, most instances of a predicate CSP $(D, \mathfrak{R})$ at high constraint density are unsatisfiable, if the predicate is violated by some assignment. This result generalizes to non-predicate CSPs, with a twist. Previous works on predicate CSPs (e.g. [Tul09, Lemma A.1]) exploited the symmetry of $\mathfrak{R}$ to argue that in a random instance, every assignment satisfies not many more constraints than a random assignment. The symmetry argument breaks down for a non-predicate CSP, as one can see from the following example.

A k-CSP $(D, \mathfrak{R})$ is *trivially satisfiable by a single value*[8] if there is some value $a \in D$ whose constant assignment $v \in [k] \mapsto a$ satisfies every relation in $\mathfrak{R}$. We also call it *trivially satisfiable* for short. Random instances of such a CSP are always satisfiable for a trivial reason, regardless of how unlikely a random assignment satisfies the relations in $\mathfrak{R}$. An extreme example is $\mathfrak{R} \stackrel{\text{def}}{=} \{R\}, R \stackrel{\text{def}}{=} \{b\}, b \stackrel{\text{def}}{=} v \in [k] \mapsto a$ for some value $a \in D$, while D and k are large. Fortunately, trivial satisfiability turns out to be the only obstacle to unsatisfiability of random instances.

Lemma B.2. *Suppose a k-CSP $(D, \mathfrak{R})$ is not trivially satisfiable. Fix any assignment $b : V \to D$ from a variable set of size $n \geqslant 2|D|k$. Then b violates a random constraint over V with probability at least $1/4^{|D|^k}$.*

Proof. Some value $a \in D$ must be assigned to at least $1/|D|$ fraction of the variables by b. Let s be the number of variables getting the value a under b. Since the CSP is not trivially satisfiable, the constant assignment $v \in [k] \mapsto a$ violates some relation $R \in \mathfrak{R}$. A random constraint is violated if its scope is contained in $b^{-1}(a)$ and its relation is R. A random constraint has its scope contained $b^{-1}(a)$ with probability $p \stackrel{\text{def}}{=} \binom{s}{k}/\binom{n}{k} \geqslant (1/2|D|)^k$ when $n \geqslant 2|D|k$. A random constraint has its relation being R with probability $q \stackrel{\text{def}}{=} 1/|\mathfrak{R}| \geqslant 1/2^{|D|^k}$. Therefore b violates a random constraint with probability at least $pq \geqslant 1/(2|D|^k 2^{|D|^k}) \geqslant 1/4^{|D|^k}$. □

The rest of the proof is the standard Chernoff and union bounds.

Lemma B.3. *If a k-CSP $(D, \mathfrak{R})$ is not trivially satisfiable. As long as $\Delta \geqslant \exp(O(|D|^k))$, except with probability $o_{n;|D|,k}(1)$ over a random instance of the CSP with n variables and $m \stackrel{\text{def}}{=} \Delta n$ constraints, every assignment violates at least $1/8^{|D|^k}$ fraction of the constraints.*

Proof. Let p be the minimum probability over all assignments $b : V \to D$ that a random constraint is violated by b. The previous Lemma implies $p \geqslant 1/4^{|D|^k}$ when $n \geqslant 2|D|k$.

Now fix any assignment $b : V \to D$. In a random instance, violation of the m constraints are m independent events, each with probability $p_b \geqslant p$. Chernoff inequality implies b violates at most $1/8^{|D|^k} \leqslant p/2$ fraction of constraints with probability at most $\exp(-mp/8)$. Taking a union bound over all assignments, some assignment violates fewer than $p/2$ fraction of constraints with probability

[8] Such CSPs are called *reflexive* in [AD22].

at most $\exp(n\log|D| - mp/8)$, which is at most $\exp(-n\log|D|)$ whenever $\Delta \geqslant 16(\log|D|)/p = \exp(O(|D|^k))$. □

B.2. **Expansion.**

[KMOW17] proved the following result showing expansion of random instances:

Lemma B.4 ([KMOW17, Theorem 4.12][9])**.** *Let $\lambda \stackrel{\text{def}}{=} \tau - 1 \geqslant 1$. Fix $0 < \gamma \leqslant .99\lambda$ and $0 < \beta < 1/2$. Then except with probability β, a random instance I with $m \stackrel{\text{def}}{=} \Delta n$ constraints is (t,γ)-expanding, provided*

$$t = \frac{n}{\Delta^{2/(\lambda-\gamma)}} \cdot \frac{1}{k}\left(\frac{\beta}{2^k}\right)^{O(1/\lambda)}.$$

We do not use their result as is, because their size bound t depends on the failure probability β. Had we used Lemma B.4 to prove Theorem 1.1, we could only conclude that random 3-SAT has a hierarchy solution of level $\Omega_{\Delta,\beta}(n)$ except with probability β. The bound $\Omega_{\Delta,\beta}(n)$ is linear in n when β is constant but not when $\beta = o_n(1)$. On the other hand, our Theorem 1.1 that uses Lemma 8.16 instead gives linear level lower bound even for subconstant failure probability, as long as Δ is constant. Our proof of Lemma 8.16 below follows [KMOW17, Theorem 4.12], improved with a well known trick from [CS88, Lemma 1].

We now restate Lemma 8.16. Recall that $o_{\zeta;k}(1)$ represents a function $\varepsilon_k(\zeta)$ that, for every fixed choice of the parameter k, $\varepsilon_k(\zeta) \to 0$ as $\zeta \to 0$ from above.

Lemma B.5. *Let $\lambda \stackrel{\text{def}}{=} \tau - 1 \geqslant 1$, $0 < \gamma \leqslant \lambda/2$, $\zeta \stackrel{\text{def}}{=} \Delta^{2/(\lambda-\gamma)}/n$. Except with probability $o_{\zeta;k}(1)$, a random k-CSP instance with n vertices and $m \stackrel{\text{def}}{=} \Delta n$ constraints is (t,γ)-expanding, where*

$$t = \frac{n}{\Delta^{2/(\lambda-\gamma)}} \cdot \frac{1}{2^{O(k)}}.$$

Proof. By Definition 8.1, the hypergraph not (t,γ)-expanding implies the following: Some constraint subset $\mathcal{C}'$ of size u and some variable subset S of size s has at least q incidences, where $q \stackrel{\text{def}}{=} s + \frac{\tau+1-\gamma}{2}u$ and $1 \leqslant u \leqslant t$. Call this latter event $\mathcal{E}$. We will bound its probability.

For fixed u and s, there are $\binom{m}{u}$ choices of the constraint subset $\mathcal{C}'$, and $\binom{n}{s}$ choices of the variables S. Out of the ku variable occurrences in $\mathcal{C}'$, there are $\binom{ku}{q}$ choices for the q potential incidences.

Consider the event $\mathcal{E}'$ that these q variable occurrences all belong to S. Suppose, instead of *without* replacement, we choose the k variables of a constraint *with* replacement, the probability of $\mathcal{E}'$ can only increase. This is because a random constraint of k distinct variables has w incidences to S with probability $\binom{s}{w}/\binom{n}{w}$, while a constraint of k variables chosen with replacement has w incidences to S with probability $(\frac{s}{n})^w \geqslant \binom{s}{w}/\binom{n}{w}$. The probability of $\mathcal{E}'$ is $(\frac{s}{n})^q$ when choosing all variables of $\mathcal{C}'$ independently with replacement.

Therefore, for fixed u and s, their contribution to the probability of $\mathcal{E}$ is at most

$$\binom{m}{u}\binom{n}{s}\binom{ku}{q}\left(\frac{s}{n}\right)^q \leqslant \left(\frac{em}{u}\right)^u \left(\frac{en}{s}\right)^s 2^{ku}\left(\frac{s}{n}\right)^q = \left(e^{1+\frac{s}{u}}2^k\frac{s}{u}\right)^u \Delta^u \left(\frac{s}{n}\right)^{\frac{\lambda-\gamma}{2}u} \tag{59}$$

where the inequality uses $\binom{a}{b} \leqslant (\frac{ea}{b})^b$ and $\binom{a}{b} \leqslant 2^a$, and the equality uses the definitions of m and q. Let $A \stackrel{\text{def}}{=} \exp(1+\frac{s}{u})2^k\frac{s}{u}$ and $\delta \stackrel{\text{def}}{=} (\lambda-\gamma)/2$. Then

$$(59) = \left(A\Delta\left(\frac{s}{n}\right)^\delta\right)^u \leqslant \left(A^{1/\delta}\Delta^{1/\delta}\frac{ku}{n}\right)^{\delta u} = (B\zeta ku)^{\delta u}$$

[9] τ in our restatement equals $\tau+1$ in [KMOW17, Theorem 4.12], because our CSP is τ-wise neutral and theirs is $(\tau-1)$-wise uniform.

where the inequality uses $s \leqslant ku$, and we let $B \stackrel{\text{def}}{=} A^{1/\delta}$ in the last equality. Summing over all choices of $1 \leqslant s \leqslant ku$ and $1 \leqslant u \leqslant t$,

$$\mathbb{P}[\mathcal{E}] \leqslant \sum_{1 \leqslant u \leqslant t} ku(B\zeta ku)^{\delta u}.$$

Following [CS88, Lemma 1], split this sum according to whether u is at most $U \stackrel{\text{def}}{=} 1/(\sqrt{\zeta}kB)$.

Case $u \leqslant U$: We may assume $\zeta < 1$ because the Lemma concerns ζ approaching zero. Then

$$\sum_{1 \leqslant u \leqslant U} ku(B\zeta ku)^{\delta u} \leqslant \sum_{1 \leqslant u \leqslant U} ku\zeta^{\delta u/2} = O(k\zeta^{\delta/2}) = o_{\zeta;k}(1),$$

using $\delta \geqslant 1/2$.

Case $u > U$**:** By choosing $t \stackrel{\text{def}}{=} 1/(2\zeta Bk)$,

$$\begin{aligned}\sum_{U < u \leqslant t} ku(B\zeta ku)^{\delta u} &\leqslant \sum_{U < u \leqslant t} kt2^{-\delta u} \leqslant \frac{1}{2\zeta B}\frac{2^{-\delta U}}{1-2^{-\delta}} \lesssim_k \frac{2^{-\delta U}}{\zeta} \\ &= \frac{1}{\zeta 2^{\delta/\left(\sqrt{\zeta}kB\right)}} = o_{\zeta;k}(1).\end{aligned}$$

□

Appendix C. Reduction

Since [Tul09], it has been folkloric that hardness reductions preserve hierarchy lower bounds (Lemma C.1); see also [AO18] for a similar conclusion for related proof systems. In this section, we make explicit reduction-based SDP+AIP lower bounds for graph coloring in Corollaries C.7 and C.8, which are simple corollaries of Theorem 1.2 (to an expert in SDP hierarchy lower bound). As mentioned in the abstract of [Tul09], Corollary C.8 is stronger than the NP-hardness results of graph coloring known even under the Unique Games Conjecture and its variants. This section helps compare [CŽ24] with what may follow from Theorem 1.2. We did not include this section in our STOC submission. Chronologically, the results of this section appear after [CŽ24] first appeared.

Consider a hardness reduction B from a CSP $(D, \mathfrak{R})$ to another $(D', \mathfrak{R}')$. Then B maps instances I of $(D, \mathfrak{R})$ to instances I' of $(D', \mathfrak{R}')$. B also maps satisfying assignments $a \in A_I$ of I to those $a' \in A_{I'}$ of I'. Suppose every variable in a' depends on at most t variables in a. More precisely, B is *t-local* if for every variable v' in I', there is a subset $W_{v'}$ of at most t variables in I and a function $g_{v'} : A_{W_{v'}} \to D^{v'}$ mapping a satisfying assignment on $W_{v'}$ to an assignment to v'.

For $S' \subseteq V(I')$, define $W_{S'} \stackrel{\text{def}}{=} \bigcup_{v' \in S'} W_{v'}$. Extend $\{g_{v'}\}$ into functions $g_{S'} : A_{W_{S'}} \to D^{S'}$ for $S' \subseteq V(I')$, where

$$g_{S'}(a)(v') \stackrel{\text{def}}{=} g_{v'}\left(a_{W_{v'}}\right) \qquad \text{for } v' \in S', a \in A_{W_{S'}}. \tag{60}$$

B preserves local satisfiability if for every constraint C' in I' and local assignment $a \in A_{W_{V(C')}}$, $g_{V(C')}(a)$ satisfies C'.

The following lemma appeared frequently in the SDP hierarchy lower bound literature, albeit as special cases and not in this general form. For completeness, we include its proof. (The lemma also applies to general CSP, not just k-CSP.)

Lemma C.1 (Folklore, see e.g. [Tul09])**.** *Suppose a t-local reduction B maps an instance I of $(D, \mathfrak{R})$ to an instance I' of $(D', \mathfrak{R}')$ and preserves local satisfiability. Then for each elementary hierarchy, any level-dt hierarchy solution on I yields a level-d hierarchy solution on I'.*

Proof. Given a level-dt hierarchy solution s on I, define a dependent function $s' : \left(S' \in \binom{V(I')}{\leqslant d}\right) \to \mathcal{M}^{D^{S'}}$, where

$$s'(S')(a') \stackrel{\text{def}}{=} \sum_{a \in A_{W_{S'}}, g_{S'}(a)=a'} s(W_{S'})(a) \qquad \text{for } d\text{-small } S' \subseteq V(I'), a' \in D^{S'}. \tag{61}$$

$s(W_{S'})$ is defined for d-small S' because $|W_{S'}| \leqslant dt$.

Further, $s'(S')$ is in fact supported on $A_{S'}$. Indeed, every a in the sum in Eq. (61) belongs to $A_{W_{S'}}$, in particular $a_{V(C')} \in A_{W_{V(C')}}$ for constraints C' contained in S', so $g_{S'}(a)_{V(C')} \stackrel{(63)}{=} g_{V(C')}\left(a_{W_{V(C')}}\right)$ satisfies C' because B preserves local satisfiability. Therefore $a' = g_{S'}(a)$ satisfies every constraint C' contained in S' whenever a has non-zero contribution to Eq. (61).

s' inherits consistency with projection (Definition 3.1) from s, because for d-small subset $S' \subseteq V(I')$, $T' \subseteq S'$ and $a' \in A_{T'}$,

$$\begin{aligned} s'(T')(a') &\stackrel{(61)}{=} \sum_{a \in A_{W_{T'}}, g_{T'}(a)=a'} s(W_{T'})(a) \stackrel{(*)}{=} \sum_{b \in A_{W_{S'}}, b_{W_{T'}}=a, g_{T'}(a)=a'} s(W_{S'})(b) \\ &\stackrel{(63)}{=} \sum_{b \in A_{W_{S'}}, g_{S'}(b)=b', b'_{T'}=a'} s(W_{S'})(b) \stackrel{(61)}{=} \sum_{b' \in A_{S'}, b'_{T'}=a'} s'(S')(b') \stackrel{(2)}{=} \pi_{T'}(s'(S'))(a'). \end{aligned}$$

In the above, $(*)$ is because s is a hierarchy solution consistent with projection.

Like s, s' also satisfies the nontriviality condition, because $s'(\emptyset)(\mathbb{0}) \stackrel{(61)}{=} s(\emptyset)(\mathbb{0}) = 1$. If the hierarchy is BW, LP, or AIP, then s' is a hierarchy solution by Lemma 3.3. For the SDP hierarchy, we need to verify Eq. (1). Indeed, given an SDP hierarchy solution $s = ((\alpha_S)_S, (\mu_R)_R)$ of I, let $s' = ((\alpha'_{S'})_{S'}, (\mu'_{R'})_{R'})$ be given by Eq. (61). Then for d-small $S', T' \subseteq V(I')$, $a'_{S'} \in A_{S'}, a'_{T'} \in A_{T'}$,

$$\begin{aligned} \langle \alpha'_{S'}(a'_{S'}), \alpha'_{T'}(a'_{T'}) \rangle_{\mathcal{X}} &\stackrel{(61)}{=} \sum_{\substack{(a_S, a_T) \in (A_{W_{S'}}, A_{W_{T'}}) \\ (g_{S'}(a_S), g_{T'}(a_T)) = (a'_{S'}, a'_{T'})}} \langle \alpha_{W_{S'}}(a_S), \alpha_{W_{T'}}(a_T) \rangle_{\mathcal{X}} \\ &\stackrel{(1)}{=} \sum_{\substack{(a_S, a_T) \in (A_{W_{S'}}, A_{W_{T'}}) \\ (g_{S'}(a_S), g_{T'}(a_T)) = (a'_{S'}, a'_{T'})}} \mathbb{P}_{b \sim \mu_{W_{S'} \cup W_{T'}}} [b_{W_{S'}} = a_S, b_{W_{T'}} = a_T] \\ &\stackrel{(\dagger)}{=} \sum_{\substack{a_{S \cup T} \in A_{W_{S' \cup T'}} \\ g_{S' \cup T'}(a_{S \cup T}) = a'_{S' \cup T'}}} \mathbb{P}_{b \sim \mu_{S \cup T}} [b = a_S \cup a_T], \\ &\stackrel{(61)}{=} \mathbb{P}_{b' \sim \mu'_{S' \cup T'}} [b' = a'_{S'} \cup a'_{T'}] \stackrel{(\#)}{=} \mathbb{P}_{b' \sim \mu'_{S' \cup T'}} [b'_{S'} = a'_{S'}, b'_{T'} = a'_{T'}]. \end{aligned} \tag{62}$$

For $(\dagger)$, we write $a_{S \cup T} \stackrel{\text{def}}{=} a_S \cup a_T$ (we may assume $(a_S)_{W_{S'} \cap W_{T'}} = (a_T)_{W_{S'} \cap W_{T'}}$, for otherwise (a_S, a_T) does not contribute). $a'_{S' \cup T'} \stackrel{\text{def}}{=} g_{S' \cup T'}(a_S \cup a_T)$, which coincides with $g_{S'}(a_S) \cup g_{T'}(a_T)$ by Eq. (63), so $a'_{S' \cup T'} = a'_{S'} \cup a'_{T'}$ (which also explains $(\#)$). □

Claim C.2. *For any $T' \subseteq S' \subseteq V(I')$, any $b \in A_{W_{S'}}$,*

$$g_{T'}\left(b_{W_{T'}}\right) = g_{S'}(b)_{T'}, \tag{63}$$

Proof. For $v' \in T'$, $\quad g_{T'}\left(b_{W_{T'}}\right)(v') \stackrel{(60)}{=} g_{v'}(b_{W_{v'}}) \stackrel{(60)}{=} g_{S'}(b)(v') = (g_{S'}(b))_{T'}(v')$. □

Lemma C.3. *Suppose a t-local reduction B maps an instance I of $(D, \mathfrak{R})$ to an instance I' of $(D', \mathfrak{R}')$ and preserves local satisfiability. Then for each combined hierarchy, any level-dt hierarchy solution on I yields a level-d hierarchy solution on I'.*

Proof. Starting with a solution $s = (s_1, s_2)$ to I in the combined hierarchy, define $s' = (s'_1, s'_2)$ from s by Eq. (61). Then each s'_i is derived from s_i by Eq. (61). Lemma C.1 implies each s'_i is a hierarchy solution to I'.

Now consider either the LP+AIP or the BW+AIP hierarchy. Solutions are of the form $s = (s_1, s_2)$, where $s_i : (S \in \mathcal{S}_d) \to \mathcal{M}_i^{A_S}$. For LP+AIP or BW+AIP, $\mathcal{M}_1 \in \{\mathbb{B}, \mathbb{R}_+\}$. Then $x \in \mathcal{M}_1 \setminus \{0\}$ implies $x + y \neq 0$ for $y \in \mathcal{M}_1$. It remains to verify $\operatorname{supp}(s'_1(S')) \supseteq \operatorname{supp}(s'_2(S'))$ for d-small $S' \subseteq V(I')$. Eq. (61) implies $\operatorname{supp}(s'_i(S')) \subseteq g_{S'}(\operatorname{supp}(s_i(W_{S'})))$ for each i. s'_1 satisfies the stronger property $\operatorname{supp}(s'_1(S')) = g_{S'}(\operatorname{supp}(s_1(W_{S'})))$, thanks to the aforementioned property of $\mathcal{M}_1$. Thus

$$\operatorname{supp}(s'_2(S')) \subseteq g_{S'}(\operatorname{supp}(s_2(W_{S'}))) \overset{(*)}{\subseteq} g_{S'}(\operatorname{supp}(s_1(W_{S'}))) = \operatorname{supp}(s'_1(S')),$$

where $(*)$ is due to $\operatorname{supp}(s_2(W_{S'})) \subseteq \operatorname{supp}(s_1(W_{S'}))$, which holds because s is a solution to the combined hierarchy.

Finally consider the SDP+AIP hierarchy. Solutions are of the form $s = (\alpha, r)$, where $r = (\mu, w)$ is a level-$2d$ LP+AIP solution, and the level-d SDP vector solution α induces μ. Since s' is derived from s by Eq. (61), (α', μ') is derived from (α, μ) by Eq. (61), so α' induces μ' by Eq. (62). □

Lemma C.4. *Fix $k \geqslant 3$ and any domain D whose size is a prime. There is a k-CSP with domain D such that except with probability $o_{n;k}(1)$, its random instance I with n variables and $O(|D|^k n)$ constraints has an SDP+AIP hierarchy solution of level $\Omega_k(n)$, but every assignment to I satisfies at most $1/|D|^{\Omega(k)}$ fraction of constraints of I.*

Proof. Apply Theorem 1.2 and [Tul09, Lemma A.1] to the k-CSP with the following predicate Q: $Q \subseteq D^k$ is the union of the Hadamard predicate of the Max-k-CSP in [Tul09, Fact 2.4], which supports a pairwise uniform distribution, and any Hamming ball of radius 2, which supports a pairwise neutral AIP assignment (Proposition 4.5). Q contains only $O((|D|k)^2)/|D|^k = 1/|D|^{\Omega(k)}$ fraction of all $|D|^k$ assignments to a k-ary constraint. □

A version of Lemma C.4 for SDP only appeared in [Tul09, Corollary 6.2], and was the starting point of the following SDP hierarchy lower bounds. A hierarchy lower bound for C-vs-K graph coloring means a hierarchy solution for C-coloring on a graph G, where G is not K-colorable.

Theorem C.5 ([Tul09, Theorem 6.5])**.** *There is a level-$\Omega_C(N)$ SDP hierarchy lower bound for C-vs-$2^{\Omega(C)}$ graph coloring on an N-vertex graph, for every large enough constant C.*

Theorem C.6 ([Tul09, Theorem 6.7])**.** *There is a level-d SDP hierarchy lower bound for C-vs-K graph coloring on an N-vertex graph, where*

$$C = 2^{O(\log N/\log\log N)} \qquad K = N/2^{O(\log N/\log\log N)} \qquad d = 2^{\Omega(\log N/\log\log N)}.$$

Apply Lemmas C.3 and C.4 to the reductions in these two theorems, we strengthen the results to SDP+AIP, with no loss in parameters. We can apply Lemma C.3 because Tulsiani implicitly verified that these reductions are t-local for an appropriate t and preserve local satisfiability.

Corollary C.7. *There is a level-$\Omega_C(N)$ SDP+AIP hierarchy lower bound for C-vs-$2^{\Omega(C)}$ graph coloring on an N-vertex graph, for every large enough constant C.*

Corollary C.8. *There is a level-d SDP+AIP hierarchy lower bound for C-vs-K graph coloring on an N-vertex graph, where*

$$C = 2^{O(\log N/\log\log N)} \qquad K = N/2^{O(\log N/\log\log N)} \qquad d = 2^{\Omega(\log N/\log\log N)}.$$

[KOWŽ23] showed that C-vs-K graph coloring is NP-hard for every constant $C \geqslant 4$ and $K = \binom{C}{\lfloor C/2 \rfloor}$. It is possible to prove linear-level SDP+AIP lower bound for the same choices of C and K, sharpening Corollary C.7. We omit the details.

For comparison, [CŽ24] obtained the following lower bound for graph coloring:

Theorem C.9 ([CŽ24])**.** *For any constant $K \geqslant 3$, there is a lower bound to the basic SDP+AIP relaxation for 3-vs-K graph coloring.*

References

[AD08] Albert Atserias and Víctor Dalmau. A Combinatorial Characterization of Resolution Width. *Journal of Computer and System Sciences*, 74(3):323–334, 2008. Computational Complexity 2003.

[AD22] Albert Atserias and Víctor Dalmau. Promise Constraint Satisfaction and Width. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1129–1153, 2022.

[AGH17] Per Austrin, Venkatesan Guruswami, and Johan Håstad. $(2 + \varepsilon)$-Sat Is NP-hard. *SIAM Journal on Computing*, 46(5):1554–1573, 2017.

[AH11] Per Austrin and Johan Håstad. Randomly Supported Independence and Resistance. *SIAM Journal on Computing*, 40(1):1–27, 2011.

[AM09] Per Austrin and Elchanan Mossel. Approximation Resistant Predicates from Pairwise Independence. *Computational Complexity*, 18:249–272, 2009.

[AO18] Albert Atserias and Joanna Ochremiak. Proof Complexity Meets Algebra. *ACM Trans. Comput. Logic*, 20(1), December 2018.

[AOW15] Sarah R. Allen, Ryan O'Donnell, and David Witmer. How to Refute a Random CSP. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 689–708, 2015.

[BBKO21] Libor Barto, Jakub Bulín, Andrei Krokhin, and Jakub Opršal. Algebraic Approach to Promise Constraint Satisfaction. *Journal of the ACM*, 68(4), 2021.

[BCK15] Boaz Barak, Siu On Chan, and Pravesh K. Kothari. Sum of Squares Lower Bounds from Pairwise Independence. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, page 97–106. ACM, 2015.

[BG19] Joshua Brakensiek and Venkatesan Guruswami. An Algorithmic Blend of LPs and Ring Equations for Promise CSPs. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, page 436–455. SIAM, 2019.

[BG21] Joshua Brakensiek and Venkatesan Guruswami. Promise Constraint Satisfaction: Algebraic Structure and a Symmetric Boolean Dichotomy. *SIAM Journal on Computing*, 50(6):1663–1700, 2021.

[BGMT12] Siavosh Benabbas, Konstantinos Georgiou, Avner Magen, and Madhur Tulsiani. SDP Gaps from Pairwise Independence. *Theory of Computing*, 8(12):269–289, 2012.

[BGWŽ20] Joshua Brakensiek, Venkatesan Guruswami, Marcin Wrochna, and Stanislav Živný. The Power of the Combined Basic Linear Programming and Affine Relaxation for Promise Constraint Satisfaction Problems. *SIAM Journal on Computing*, 49(6):1232–1248, 2020.

[BK09] Libor Barto and Marcin Kozik. Constraint Satisfaction Problems of Bounded Width. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 595–603, 2009.

[BSW01] Eli Ben-Sasson and Avi Wigderson. Short Proofs Are Narrow—Resolution Made Simple. *Journal of the ACM*, 48(2):149–169, 2001.

[Bul17] Andrei A. Bulatov. A Dichotomy Theorem for Nonuniform CSPs. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science*, pages 319–330, 2017.

[Bul20] Andrei A. Bulatov. A Dichotomy Theorem for Nonuniform CSPs Simplified. *CoRR*, abs/2007.09099, 2020.

[Cha16] Siu On Chan. Approximation Resistance from Pairwise-Independent Subgroups. *Journal of the ACM*, 63(3), 2016.

[CM13] Siu On Chan and Michael Molloy. A Dichotomy Theorem for the Resolution Complexity of Random Constraint Satisfaction Problems. *SIAM Journal on Computing*, 42(1):27–60, 2013.

[CS88] Vašek Chvátal and Endre Szemerédi. Many Hard Examples for Resolution. *Journal of the ACM*, 35(4):759–768, 1988.

[CŽ22] Lorenzo Ciardo and Stanislav Živný. CLAP: A New Algorithm for Promise CSPs. In *Proceedings of the 2022 Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1057–1068, 2022.

[CŽ23a] Lorenzo Ciardo and Stanislav Živný. Approximate Graph Colouring and Crystals. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2256–2267, 2023.

[CŽ23b] Lorenzo Ciardo and Stanislav Živný. Approximate Graph Colouring and the Hollow Shadow. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, page 623–631. ACM, 2023.

[CŽ23c] Lorenzo Ciardo and Stanislav Živný. Hierarchies of Minion Tests for PCSPs through Tensors. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 568–580, 2023.

[CŽ24] Lorenzo Ciardo and Stanislav Živný. Semidefinite Programming and Linear Equations vs. Homomorphism Problems. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, 2024. To appear.

[DFHT21] Irit Dinur, Yuval Filmus, Prahladh Harsha, and Madhur Tulsiani. Explicit SoS Lower Bounds from High-Dimensional Expanders. In *12th Innovations in Theoretical Computer Science Conference, 2021*, volume 185, pages 38:1–38:16, 2021.

[DO23] Victor Dalmau and Jakub Opršal. Local Consistency as a Reduction Between Constraint Satisfaction Problems. *CoRR*, abs/2301.05084, 2023.

[dT23] Tommaso d'Orsi and Luca Trevisan. A Ihara-Bass Formula for Non-Boolean Matrices and Strong Refutations of Random CSPs. In *38th Computational Complexity Conference*, volume 264, pages 27:1–27:16, 2023.

[Dur19] Rick Durrett. *Probability: Theory and Examples*. 2019. Version 5.

[Fei02] Uriel Feige. Relations between Average Case Complexity and Approximation Complexity. In *Proceedings of the Thiry-Fourth Annual ACM Symposium on Theory of Computing*, page 534–543. ACM, 2002.

[FKO06] Uriel Feige, Jeong Han Kim, and Eran Ofek. Witnesses for Non-Satisfiability of Dense Random 3CNF Formulas. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 497–508, 2006.

[G02] Frank Göring. A Proof of Menger's Theorem by Contraction. *Discussiones mathematicae: Graph theory*, 22, 2002(1):111–112, 2002. Faculty of Mathematics, Computer Science and Econometrics, University of Zielona Góra.

[GJ76] M. R. Garey and D. S. Johnson. The Complexity of Near-Optimal Graph Coloring. *Journal of the ACM*, 23(1):43–49, Jan 1976.

[Gri01] Dima Grigoriev. Linear Lower Bound on Degrees of Positivstellensatz Calculus Proofs for the Parity. *Theoretical Computer Science*, 259(1):613–622, May 2001.

[GS12] Venkatesan Guruswami and Ali Kemal Sinop. Faster SDP Hierarchy Solvers for Local Rounding Algorithms. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 197–206, 2012.

[GS20] Venkatesan Guruswami and Sai Sandeep. d-To-1 Hardness of Coloring 3-Colorable Graphs with $O(1)$ Colors. In *47th International Colloquium on Automata, Languages, and Programming*, pages 62:1–62:12, 2020.

[HL22] Max Hopkins and Ting-Chun Lin. Explicit Lower Bounds Against $\Omega(n)$-Rounds of Sum-of-Squares . In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science*, pages 662–673, 2022.

[KMOW17] Pravesh K. Kothari, Ryuhei Mori, Ryan O'Donnell, and David Witmer. Sum of Squares Lower Bounds for Refuting Any CSP. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, page 132–145. ACM, 2017.

[KO22] Andrei Krokhin and Jakub Opršal. An Invitation to the Promise Constraint Satisfaction Problem. *ACM SIGLOG News*, 9(3):30–59, Aug 2022.

[KOWŽ23] Andrei Krokhin, Jakub Opršal, Marcin Wrochna, and Stanislav Živný. Topology and Adjunction in Promise Constraint Satisfaction. *SIAM Journal on Computing*, 52(1):38–79, 2023.

[Lau09] Monique Laurent. *Sums of Squares, Moment Matrices and Optimization Over Polynomials*, pages 157–270. Springer, 2009.

[LY94] Carsten Lund and Mihalis Yannakakis. On the Hardness of Approximating Minimization Problems. *Journal of the ACM*, 41(5):960–981, Sep 1994.

[LZ22] Anthony Leverrier and Gilles Zémor. Quantum Tanner Codes. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science*, pages 872–883, 2022.

[Mos10] Elchanan Mossel. Gaussian Bounds for Noise Correlation of Functions. *Geometric and Functional Analysis*, 2010.

[MW16] Ryuhei Mori and David Witmer. Lower Bounds for CSP Refutation by SDP Hierarchies. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, volume 60, pages 41:1–41:30, 2016.

[OC22] Adam Ó Conghaile. Cohomology in Constraint Satisfaction and Structure Isomorphism. In *47th International Symposium on Mathematical Foundations of Computer Science, 2022*, volume 241, pages 75:1–75:16, 2022.

[O'D17] Ryan O'Donnell. SOS Is Not Obviously Automatizable, Even Approximately. In *8th Innovations in Theoretical Computer Science Conference*, volume 67, pages 59:1–59:10, 2017.

[RSW16] Ilya Razenshteyn, Zhao Song, and David P. Woodruff. Weighted Low Rank Approximations with Provable Guarantees. In *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*, page 250–263. ACM, 2016.

[RW17] Prasad Raghavendra and Benjamin Weitz. On the Bit Complexity of Sum-of-Squares Proofs. In *44th International Colloquium on Automata, Languages, and Programming*, volume 80, pages 80:1–80:13, 2017.

[Sch86] Alexander Schrijver. *Theory of Linear and Integer Programming*. Wiley-Interscience series in discrete mathematics and optimization. 1986.

[Sch08] Grant Schoenebeck. Linear Level Lasserre Lower Bounds for Certain k-CSPs. In *Proceedings of the 2008 49th Annual IEEE Symposium on Foundations of Computer Science*, page 593–602, 2008.

[Tul09] Madhur Tulsiani. CSP Gaps and Reductions in the Lasserre Hierarchy. In *Proceedings of the Forty-First Annual ACM Symposium on Theory of Computing*, page 303–312. ACM, 2009.

[TW13] Madhur Tulsiani and Pratik Worah. LS+ Lower Bounds from Pairwise Independence. In *2013 IEEE Conference on Computational Complexity*, pages 121–132, 2013.

[Zhu20] Dmitriy Zhuk. A Proof of the CSP Dichotomy Conjecture. *Journal of the ACM*, 67(5), 2020.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il