Towards Infinite PCSP: A Dichotomy for Monochromatic Cliques

Demian Banakh
Faculty of Mathematics and Computer Science, Jagiellonian University, Kraków, Poland
Doctoral School of Exact and Natural Sciences, Jagiellonian University, Kraków, Poland

Alexey Barsukov
Department of Algebra, Faculty of Mathematics and Physics, Charles University, Prague, Czechia

Tamio-Vesa Nakajima
Faculty of Mathematics and Informatics, Philipps University, Marburg, Germany

Abstract

The logic MMSNP is a well-studied fragment of Existential Second-Order logic that, from a computational perspective, captures finite-domain Constraint Satisfaction Problems (CSPs) modulo polynomial-time reductions. At the same time, MMSNP contains many problems that are expressible as ω-categorical CSPs but not as finite-domain ones.

We initiate the study of *Promise* MMSNP (PMMSNP), a promise analogue of MMSNP. We show that every PMMSNP problem is poly-time equivalent to a (finite-domain) *Promise* CSP (PCSP), thereby extending the classical MMSNP–CSP correspondence to the promise setting. We then investigate the complexity of PMMSNPs arising from forbidding monochromatic cliques, a class encompassing promise graph colouring problems. For this class, we obtain a full complexity classification conditional on the *Rich 2-to-1 Conjecture*, a recently proposed perfect-completeness surrogate of the Unique Games Conjecture.

As a key intermediate step which may be of independent interest, we prove that it is NP-hard, under the Rich 2-to-1 Conjecture, to properly colour a uniform hypergraph even if it is promised to admit a colouring satisfying a certain technical condition called *reconfigurability*. This proof is an extension of the recent work of Braverman, Khot, Lifshitz and Minzer (Adv. Math. 2025). To illustrate the broad applicability of this theorem, we show that it implies most of the linearly-ordered colouring conjecture of Barto, Battistelli, and Berg (STACS 2021).

2012 ACM Subject Classification Theory of computation → Constraint and logic programming; Theory of computation → Problems, reductions and completeness; Theory of computation → Complexity theory and logic; Mathematics of computing → Combinatorics; Mathematics of computing → Graph theory

Keywords and phrases Promise Constraint Satisfaction Problem, MMSNP, Approximation Algorithms, Rainbow Colouring

Digital Object Identifier 10.4230/LIPIcs.LICS.2026.13

Related Version *Full Version*: `https://arxiv.org/abs/2605.09815` [4]

Funding *Demian Banakh*: This research was funded in whole or in part by the National Science Centre, Poland under the Weave-UNISONO call in the Weave programme 2021/03/Y/ST6/00171 and the Polish National Agency for Academic Exchange under Zawacka programme. For the purpose of Open Access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript (AAM) version arising from this submission.
Alexey Barsukov: Funded by the European Union (ERC, POCOCOP, 101071674). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Acknowledgements We would like to thank Andrei Krokhin and Amey Bhangale for useful discussions and the anonymous reviewers for their valuable suggestions. Furthermore, we sincerely thank the organisers of the 2025 CSP World Congress, where this project was initiated.

1 Introduction

Monotone Monadic SNP (MMSNP) is a large fragment of Existential Second-Order Logic (ESO) introduced by Feder and Vardi [34], with the goal of having a P vs. NP-complete dichotomy (whereas the whole logic ESO contains NP-intermediate problems assuming P $\neq$ NP [33, 51]). Fixing a language σ consisting of relational symbols $R_1, \ldots, R_k$, an MMSNP formula is an existential second-order formula of the form

$$\Phi = \exists M_1 \ldots M_c \forall x_1 \ldots x_n \; \phi(x_1, \ldots, x_n),$$

which additionally satisfies the following syntactic restrictions:

Monotone: every σ-atom $R_i(\ldots)$ within ϕ has an odd number of nested negations before it.

Monadic: every relation among $M_1, \ldots, M_c$ is unary i.e. it is a set.

No inequality: the quantifier-free part ϕ does not contain = or $\neq$.

▶ **Example 1.** The sentence

$$\Phi = \exists A, B \; \forall x, y \begin{pmatrix} (A(x) \vee B(x)) \\ \wedge \\ (\neg A(x) \vee \neg B(x)) \\ \wedge \\ (A(x) \wedge A(y)) \Rightarrow \neg E(x, y) \\ \wedge \\ (B(x) \wedge B(y)) \Rightarrow \neg E(x, y) \end{pmatrix}$$

is in MMSNP. In this case, $\sigma = \{E\}$ has one binary relation symbol, so relational σ-structures are directed graphs. For a digraph $\mathbb{G}$, note that $\mathbb{G} \models \Phi$ (read $\mathbb{G}$ models Φ, or Φ is satisfied in $\mathbb{G}$) if and only if we can partition the vertex set G of $\mathbb{G}$ into two sets A, B so that there is no edge between vertices from the same set i.e. when $\mathbb{G}$ is 2-colourable.

As illustrated by Example 1, there is a natural relation between MMSNP satisfiability and *Constraint Satisfaction Problems* (CSPs). For a relational σ-structure $\mathbb{A} = (A; R_1^{\mathbb{A}}, \ldots, R_k^{\mathbb{A}})$, CSP($\mathbb{A}$) is a computational problem, where we are given a finite σ-structure $\mathbb{I} = (I; R_1^{\mathbb{I}}, \ldots, R_k^{\mathbb{I}})$ and must decide the existence of a mapping $h: I \to A$ that preserves every relation $R_i \in \sigma$. Such a mapping is called a *homomorphism* and is denoted by $h: \mathbb{I} \to \mathbb{A}$. Whenever $|A| < \infty$, we are able to construct, similarly to Example 1, a sentence Φ in MMSNP such that $\mathbb{I} \to \mathbb{A}$ if and only if $\mathbb{I} \models \Phi$, for every input $\mathbb{I}$.

Feder and Vardi showed that, although MMSNP is strictly more expressive than finite-domain CSP, for every MMSNP sentence Φ there exists a finite relational structure $\mathbb{A}$ such that MMSNP(Φ) is equivalent to CSP($\mathbb{A}$) under randomised poly-time reductions; they also conjectured a P vs. NP-complete dichotomy for both classes. These reductions were later derandomised by Kun [50].

The conjecture of Feder and Vardi about CSPs inspired an active research programme. It culminated two decades later, when Bulatov and Zhuk independently proved that CSPs admit a P vs. NP-complete dichotomy, using algebraic tools [26, 61, 62]. This automatically implied a similar dichotomy for MMSNPs. Shortly thereafter, Bodirsky, Madelaine, and Mottet [15, 16] showed that the complexity of MMSNP problems also depends on the algebraic properties of corresponding (ω-categorical) CSPs, thus confirming the conjecture of Bodirsky and Pinsker [18] on MMSNP problems.

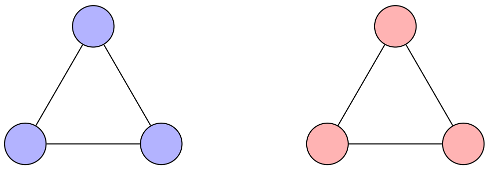

Figure 1 Template $\mathcal{F}_3^{(2)}$ for 2-colouring a graph avoiding monochromatic triangles.

The logic of forbidden patterns

Despite the fact that MMSNP is defined in terms of syntactic restrictions applied to ESO sentences, it has an equivalent "combinatorial" definition called *forbidden pattern problems* [53]. Each such problem is described by a finite family $\mathcal{F}$ of (element-)coloured connected relational structures $(\mathbb{F}, f)$ called *forbidden patterns*, where the mapping $f\colon F \to [c]$ stands for a colouring of the elements of $\mathbb{F}$ with c-many colours, for some $c \in \mathbb{N}$. For a given relational structure $\mathbb{I}$, the question is to decide whether the elements of $\mathbb{I}$ are colourable in such a way that there is no (colour respecting) homomorphism from a forbidden pattern of $\mathcal{F}$. A coloured structure that avoids a homomorphism from a forbidden pattern to itself is called *$\mathcal{F}$-free*. For example, the problem of 2-colouring a graph so as to avoid all monochromatic triangles belongs to MMSNP; a drawing of the family $\mathcal{F}$ for this problem can be found in Figure 1. Every forbidden pattern problem is clearly in MMSNP and every problem in MMSNP is a finite union of forbidden pattern problems [53, Corollary 13]. Taking into account this equivalence and that the problems studied in this paper are much easier to formulate in the framework of forbidden patterns, we will stick to this notation further on. So, we also denote by MMSNP$(\mathcal{F})$ the forbidden pattern problem associated with $\mathcal{F}$.

MMSNP and finite CSP

As mentioned above, Feder and Vardi [34] demonstrated that each MMSNP is equivalent to some finite-domain CSP. In fact, their construction is quite straightforward. They define the CSP template as follows: the domain is the set $[c]$ of colours used in $\mathcal{F}$; every forbidden pattern $(\mathbb{F}, f) \in \mathcal{F}$ yields an $|F|$-ary relation in the template that consists of all colourings $g\colon F \to [c]$ that do not belong to $\mathcal{F}$. The reduction from MMSNP$(\mathcal{F})$ to this CSP then follows in a natural way: given $\mathbb{I}$, we produce $\mathbb{I}'$ – an instance of our CSP – by listing, for each forbidden pattern $(\mathbb{F}, f) \in \mathcal{F}$, all homomorphisms from $\mathbb{F}$ to $\mathbb{I}$.

▶ **Example 2** (No monochromatic triangle)**.** Suppose that there are just two colours: b, r; and that $\mathcal{F}_3^{(2)}$ consists of two copies of $\mathbb{K}_3$ – one is coloured fully in b, the other fully in r (cf. Figure 1). The σ-structure associated with $\mathcal{F}_3^{(2)}$ has domain $\{b, r\}$ and one symmetric ternary relation $\{b, r\}^3 \setminus \{(b, b, b), (r, r, r)\}$; thus the corresponding CSP template is just the ternary *Not All Equal* relation (NAE_3). The reduction from our MMSNP to the corresponding CSP simply applies this constraint to every triangle (and every loop) in the input graph.

The reduction in the opposite direction relies on the *Sparse Incomparability Lemma* which states that, for every finite σ-structure $\mathbb{B}$ and $\ell > 1$, the complexity of CSP$(\mathbb{B})$ remains the same even after restricting the input to structures without cycles of length $< \ell$.

MMSNP and infinite CSP

Some sentences in MMSNP describe problems that are not finite-domain CSPs: for example, the *triangle-free* property can be expressed in MMSNP but there are triangle-free graphs with arbitrarily high chromatic number. (Hence no finite CSP has as Yes instances precisely the set of finite triangle-free graphs.) The whole class of CSPs of ω-categorical structures is too large: it contains even coNP-complete and coNP-intermediate problems [12].[1] On the other hand, many natural families of infinite-domain CSPs exhibit a P vs. NP-complete dichotomy: reducts of $(\mathbb{Q}, <)$ [13], reducts of homogeneous graphs [17], reducts of the random poset [47], reducts of the random tournament [56]. All these classes fall into the scope of the family of *reducts of finitely bounded homogeneous structures*.

For CSPs of these structures, Bodirsky and Pinsker conjectured a P vs. NP-complete dichotomy with a similar universal algebraic characterisation as for finite-domain CSPs, see [18]. The name for this class of structures comes from the way they are created. It starts from a *finite* family $\mathcal{F}$ of structures, called the *bounds* and the class $\mathrm{Forb}^{\mathrm{emb}}(\mathcal{F})$ of finite structures that have no substructure from $\mathcal{F}$. The class $\mathrm{Forb}^{\mathrm{emb}}(\mathcal{F})$ is supposed to have the *amalgamation property*: for every $\mathbb{B}_1, \mathbb{B}_2$ in $\mathrm{Forb}^{\mathrm{emb}}(\mathcal{F})$ that have a common substructure $\mathbb{A}$ witnessed by embeddings $e_i\colon \mathbb{A} \hookrightarrow \mathbb{B}_i$ (for $i \in \{1, 2\}$), there is a structure $\mathbb{C}$ in $\mathrm{Forb}^{\mathrm{emb}}(\mathcal{F})$ and embeddings $f_i\colon \mathbb{B}_i \hookrightarrow \mathbb{C}$ (for $i \in \{1, 2\}$) such that $f_1 \circ e_1 = f_2 \circ e_2$. Then, using Fraïssé's Theorem [41, Section 6.1], we inductively construct a countably infinite structure $\mathbb{H}$ (called the *Fraïssé-limit* of $\mathrm{Forb}^{\mathrm{emb}}(\mathcal{F})$), whose finite substructures are exactly those from $\mathrm{Forb}^{\mathrm{emb}}(\mathcal{F})$ and which is *homogeneous*: every isomorphism between finite substructures of $\mathbb{H}$ extends to an automorphism of $\mathbb{H}$. The final structure, called a (first-order) *reduct* of $\mathbb{H}$, is obtained from $\mathbb{H}$ by adding new first-order definable relations and then forgetting some relations. Note that structures whose CSPs capture MMSNP problems are constructed similarly: for instance, the class $\mathrm{Forb}^{\mathrm{emb}}(\mathcal{F}_3^{(2)})$ of coloured graphs without monochromatic triangles (Figure 1) already has the amalgamation property. So, it is not surprising that every MMSNP problem is equivalent to a finite union of CSPs of finitely bounded homogeneous structures [15, 16], and, as it was mentioned, a dichotomy for them is the special case of the conjecture of Bodirsky and Pinsker.

In spite of this natural connection of MMSNP to infinite-domain CSPs, our techniques in this paper will be concerned solely with finite-domain relational structures.

Recolouring and containment

In their paper [34], Feder and Vardi showed that containment for problems in MMSNP is decidable. (By containment, we mean deciding whether all Yes instances of one problem are also Yes instances of the other problem.) For two families $\mathcal{F}_1$ and $\mathcal{F}_2$ of forbidden patterns, the containment (of the corresponding MMSNP problems) is characterised by the existence of a mapping $r\colon [c_1] \to [c_2]$ between the sets of colours that lifts up to a mapping from $\mathrm{Forb}^{\mathrm{hom}}(\mathcal{F}_1)$ to $\mathrm{Forb}^{\mathrm{hom}}(\mathcal{F}_2)$ [52], where $\mathrm{Forb}^{\mathrm{hom}}(\mathcal{F})$ is the set of finite vertex-coloured structures that do not admit a homomorphism from any member of $\mathcal{F}$. Such a mapping is called a *recolouring*. We write $\mathcal{F}_1 \to \mathcal{F}_2$ if a recolouring from $\mathcal{F}_1$ to $\mathcal{F}_2$ exists.

Promise MMSNP

In this paper, we initiate the study of a promise problem variant of MMSNP, inspired by the development of *Promise CSP*. Namely, a *Promise MMSNP* (PMMSNP) over a pair of MMSNP templates $\mathcal{F}, \mathcal{G}$ such that $\mathcal{F} \to \mathcal{G}$ asks to distinguish whether the input structure admits an $\mathcal{F}$-free colouring, or not even a $\mathcal{G}$-free one. We let $\mathrm{PMMSNP}(\mathcal{F}, \mathcal{G})$ denote this problem. Note that the fact that $\mathcal{F} \to \mathcal{G}$ implies that the Yes and No instances of this

[1] However, the question whether there are ω-categorical structures whose CSP is NP-intermediate is still open [11].

problem are disjoint. Taking into account that every MMSNP is a union of ω-categorical CSPs, our research can be considered as the first step towards understanding the complexity of *Promise CSPs* of countably infinite structures.

We prove that every Promise MMSNP is poly-time equivalent to a Promise CSP defined by a suitable pair of finite structures. Recall that for a promise structure $\mathbb{A}$ and a target structure $\mathbb{B}$ such that $\mathbb{A} \to \mathbb{B}$, $\mathrm{PCSP}(\mathbb{A}, \mathbb{B})$ is a problem in which, given a structure $\mathbb{I}$, the goal is to distinguish whether $\mathbb{I} \to \mathbb{A}$ or $\mathbb{I} \not\to \mathbb{B}$. Here as well, the fact that $\mathbb{A} \to \mathbb{B}$ ensures that the Yes and No instances are disjoint.

▶ **Theorem 3.** *If $\mathcal{F} \to \mathcal{G}$, then there exist finite structures $\mathbb{S} \to \mathbb{T}$ such that* $\mathrm{PMMSNP}(\mathcal{F}, \mathcal{G})$ *and* $\mathrm{PCSP}(\mathbb{S}, \mathbb{T})$ *are poly-time equivalent.*

This is a direct analogue of the MMSNP–CSP correspondence of Feder and Vardi in the promise realm. The PCSP complexity landscape remains largely unresolved, despite significant progress, and even simple subclasses, like Boolean PCSPs or Approximate Graph Colouring, are not known to admit a full complexity classification. Therefore the dichotomy for all PMMSNPs is most likely out of reach for now. For that reason, we restrict our attention to certain particular types of families $\mathcal{F}$ and $\mathcal{G}$.

Monochromatic cliques

Consider the naturally arising graph colouring problem of the following form: given a graph $\mathbb{G}$, decide whether there is a c-colouring of $\mathbb{G}$ that avoids monochromatic k-cliques. This problem is readily seen to be an MMSNP, where the forbidden patterns are monochromatic k-cliques, coloured with one of c colours; we denote such a family with $\mathcal{F}_k^{(c)}$. This problem is NP-hard when $c, k \geq 2$ and $\max\{c, k\} > 2$, along the lines of Example 2; and it is in P otherwise.

Such problems also exist in the graph theory under the name of $\mathbb{K}_k$*-free colouring* [28]. The $\mathbb{K}_k$*-free chromatic number* $\chi_k(\mathbb{G})$ of a graph $\mathbb{G}$ is the minimum number c such that $\mathbb{G}$ has a c-colouring avoiding a monochromatic k-clique. The computational complexity of $\mathbb{K}_k$-free colouring has already been studied; for example, in [43], the authors prove that $\mathrm{MMSNP}(\mathcal{F}_3^{(c)})$ is NP-complete, for every $c \geq 2$.

We note that the problem $\mathcal{F}_k^{(c)}$ cannot be expressed as a finite-domain CSP if $k > 2$. Indeed, for any finite graph $\mathbb{H}$, there exists a graph with girth greater than k and chromatic number greater than that of $\mathbb{H}$ (by [32]). Any c-colouring of this graph is $\mathcal{F}_k^{(c)}$-free, but it does not admit a homomorphism to $\mathbb{H}$.

On the other hand, it is very easy to construct the infinite-domain CSP templates for such problems. Clearly, the class of finite $\mathcal{F}_k^{(c)}$-free coloured graphs has the amalgamation property, so there is no need to add any new relations, because the Fraïssé-limit of this class is already a finitely bounded homogeneous structure. The graph $\mathbb{H}_k^{(c)}$ such that $\mathrm{CSP}(\mathbb{H}_k^{(c)}) = \mathrm{MMSNP}(\mathcal{F}_k^{(c)})$ is obtained from c-many disjoint copies of the universal $\mathbb{K}_k$-free graph by connecting every two vertices from distinct copies with an edge. So, the problems that forbid monochromatic cliques are among the tamest MMSNP problems that are not finite-domain CSPs.

In the second part of the paper, we investigate the complexity of the promise version of these graph colouring problems, that is, Promise MMSNPs over a pair of families $\mathcal{F}_k^{(c)}$ and $\mathcal{F}_\ell^{(d)}$. In this case, the following two types of relations will often appear in the corresponding

PCSP templates obtained from Theorem 3: first, for integers $0 \leq k \leq \ell$, the ℓ-ary relation $k\text{-in-}\ell$ consists of all tuples $\overline{a} \in \{0,1\}^\ell$ for which $\sum_i a_i = k$. Also, the *Not All Equal* relation $\mathsf{NAE}_r^{(d)} \subseteq [d]^r$ consists of all non-constant tuples;[2] a tuple $\overline{a}$ is constant if $a_i = a_j$ for all i, j.

The following example illustrates that PMMSNPs over forbidden monochromatic cliques include non-trivial interesting problems in which the promise turns an otherwise NP-complete MMSNP problem into an efficiently solvable one.

▶ **Example 4.** Let us consider the PMMSNP problem with template $(\mathcal{F}_3^{(2)}, \mathcal{F}_4^{(2)})$. In this problem, we are given a graph $G = (V, E)$ and must decide whether it has a 2-colouring that avoids monochromatic triangles, or not even a 2-colouring that avoids monochromatic 4-cliques. We will show that something slightly more strict is possible: if we assume the input graph has a 2-colouring avoiding all monochromatic triangles, we will find a 2-colouring avoiding all monochromatic 4-cliques. (Thus, if our algorithm successfully outputs a 2-colouring avoiding all monochromatic 4-cliques we may answer Yes, otherwise we must answer No.)

To do this, construct a 4-uniform hypergraph whose vertices coincide with the vertices of G, and whose hyperedges are given by the 4-cliques of G. Call this hypergraph $H = (V, E')$.[3] The condition that G has a 2-colouring avoiding all monochromatic triangles implies that there exists a 2-colouring $c : V \to \{0,1\}$ of H that, for every $(x, y, z, t) \in E'$, colours half of x, y, z, t with 0 and the other half with 1. We must output a colouring $d : V \to \{0,1\}$ that leaves no edge of H monochromatic. This is just the problem $\mathrm{PCSP}(2\text{-in-}4, \mathsf{NAE}_4)$, well-known from the PCSP literature, which admits a poly-time solution via the *Affine Integer Programming (AIP) relaxation* [19]. Let us quickly present this algorithm.

By assumption, there exists $c : V \to \{0,1\}$ such that

$$\forall (x, y, z, t) \in E' : \qquad c(x) + c(y) + c(z) + c(t) = 2. \tag{1}$$

Of course, it is NP-hard to find this mapping; however, it is possible to solve the system of equations in (1) over $\mathbb{Z}$, by computing the Smith normal form [42]; suppose $\tilde{c}$ is our solution. Now, consider any edge (x, y, z, t). It is not possible that all of $\tilde{c}(x), \tilde{c}(y), \tilde{c}(z), \tilde{c}(t)$ are ≤ 0 (or else the sum would be $\leq 0 < 2$); nor is it possible that all are > 0 (or else the sum would be at least $4 > 2$). Hence, if we set $v \in V$ to 1 or 0 depending on whether $\tilde{c}(v) \geq 1$ or $\tilde{c}(v) \leq 0$, we can find the desired solution in NAE_4.

Rich 2-to-1 Conjecture

The Promise MMSNPs over $\mathcal{F}_k^{(c)}, \mathcal{F}_\ell^{(d)}$ include the notorious Approximate Graph Colouring (take $k = \ell = 2$) which can be viewed as a PCSP over two cliques. For example, $\mathrm{PCSP}(\mathbb{K}_3, \mathbb{K}_6)$ is equivalent to $\mathrm{PMMSNP}(\mathcal{F}_2^{(3)}, \mathcal{F}_2^{(6)})$. The best known hardness results for the Approximate Graph Colouring include c vs. $\Theta(2^c/\sqrt{c})$ for any $c \geq 4$ by [49], and c vs. $2c - 1$ by [9] for any $c \geq 3$. Therefore, an unconditional complexity classification for our PMMSNPs is currently out of reach.

However, under additional inapproximability assumptions, further progress on the Approximate Graph Colouring problem has been achieved. Assuming the d-to-1 Conjecture of [44], NP-hardness was proved for all $\mathrm{PCSP}(\mathbb{K}_c, \mathbb{K}_d)$ such that $3 \leq c \leq d$ in [30, 40]. Later, [23] suggested a substantially different hardness proof under the recently proposed Rich 2-to-1 Conjecture [24]. Outside of the Approximate Graph Colouring, the Rich 2-to-1

[2] We sometimes omit the superscript (d) if the domain is clear from the context.
[3] Note that triangles in G that belong to no 4-clique do not affect H.

Conjecture was the basis of several hardness results for Boolean PCSPs [5, 21]. These (and other) conjectures provide a way to bypass the limitations of the unconditional sources of hardness, and make progress on various open inapproximability problems in order to better understand their underlying mathematical structure, which could, in turn, lead to new, stronger, unconditional sources of hardness in the future.

Let us briefly introduce the Rich 2-to-1 Conjecture now (for the formal definition, see Section 5). The (perhaps) more famous Unique Games Conjecture of Khot [44] states that it is NP-hard to even approximately solve an *almost satisfiable* instance of Unique Games (equivalently, a CSP instance in which each constraint is a bijection). This conjecture has inspired numerous influential results on hardness of inapproximability of various computational problems [45, 46]. On the other hand, it is trivial to solve a *satisfiable* instance of Unique Games, therefore this conjecture has no use in tackling so called *perfect completeness* approximation problems, that is, problems of finding an approximate solution to *satisfiable* instances. The Rich 2-to-1 Conjecture was proposed in 2021 by [24] to overcome this issue. Namely, it states that it is NP-hard to approximately solve a *satisfiable* instance of Rich 2-to-1 Games (equivalently, a CSP instance in which each constraint is a 2-to-1 function with an extra regularity property imposed). On *almost satisfiable* instances, the authors of [24] proved that this conjecture is equivalent to the Unique Games Conjecture, which makes it a perfect-completeness surrogate of the UGC.

Our contributions

We view the contributions of this paper as twofold. Firstly, a classification (resulting in a dichotomy) of all Promise MMSNP problems over forbidden monochromatic cliques.

▶ **Theorem 5.** *Let c, d, k, ℓ be positive integers such that $\mathcal{F}_k^{(c)} \to \mathcal{F}_\ell^{(d)}$. Then*

- $\ell \geq c(k-1)$ *and* $\mathrm{PMMSNP}(\mathcal{F}_k^{(c)}, \mathcal{F}_\ell^{(d)})$ *is solvable in polynomial time via AIP, or*
- $\ell < c(k-1)$ *and* $\mathrm{PMMSNP}(\mathcal{F}_k^{(c)}, \mathcal{F}_\ell^{(d)})$ *is NP-hard, assuming the Rich 2-to-1 Conjecture.*

As the first step towards Theorem 5, we establish that every Promise MMSNP of interest is poly-time equivalent to a certain finite-domain Promise CSP in which both the promise and the target templates have a single, symmetric relation. The PCSPs of this form are readily seen to be expressible as certain promise hypergraph colouring problems. For concreteness, we deal with relations over a universe $[c]$ that consist of all tuples of length ℓ in which each $a \in [c]$ appears less than k times. Such a relation is denoted with $R_{k:\ell}^{(c)}$. For example, $R_{3:3}^{(2)}$ is the ternary *Not All Equal* relation. Consequently, we show that every promise MMSNP over two monochromatic clique families is poly-time equivalent to a PCSP defined by a pair of relations of this form.

▶ **Lemma 6** (Special case of Theorem 3)**.** *Let c, d, k, ℓ be positive integers. If $\mathcal{F}_k^{(c)} \to \mathcal{F}_\ell^{(d)}$, then the problems* $\mathrm{PMMSNP}(\mathcal{F}_k^{(c)}, \mathcal{F}_\ell^{(d)})$ *and* $\mathrm{PCSP}(R_{k:\ell}^{(c)}, R_{\ell:\ell}^{(d)})$ *are poly-time equivalent.*

Thus, the desired dichotomy (Theorem 5) amounts to classifying all PCSPs of this form.

▶ **Corollary 7.** *Let c, d, k, ℓ be positive integers such that $\mathcal{F}_k^{(c)} \to \mathcal{F}_\ell^{(d)}$. Then*

- $\ell \geq c(k-1)$ *and* $\mathrm{PCSP}(R_{k:\ell}^{(c)}, \mathsf{NAE}_\ell^{(d)})$ *is solvable in polynomial time via AIP, or*
- $\ell < c(k-1)$ *and* $\mathrm{PCSP}(R_{k:\ell}^{(c)}, \mathsf{NAE}_\ell^{(d)})$ *is NP-hard, assuming the Rich 2-to-1 Conjecture.*

In fact, we will demonstrate that the complexity of a $\mathrm{PCSP}(R, \mathsf{NAE}_\ell)$ is governed by certain connectivity properties of the left-hand side relation. To formalize this notion, let us call a relation *reconfigurable* if any of its members can be obtained from any other member by modifying one coordinate at a time without leaving the relation at any point.

In particular, we prove the following hardness result for PCSPs, which is the technical core and the second contribution of this paper. Let $\mathsf{NAE}_r^{(d)}$ denote the r-ary *Not All Equal* relation on d elements.

▶ **Theorem 8.** *Let $\varnothing \neq R \subseteq A^r$ be a symmetric*[4] *reconfigurable*[5] *relation of arity $r \geq 2$. For any d such that $R \to \mathsf{NAE}_r^{(d)}$, the problem $\mathrm{PCSP}(R, \mathsf{NAE}_r^{(d)})$ is NP-hard, assuming the Rich 2-to-1 Conjecture.*

It is an extension of the work of [23]: they proved this theorem in the special case of $R = \{(a,b) \in \{1,2,3\} \mid a \neq b\}$, which coincides with the Approximate Graph Colouring. It is worth noting that their proof in fact covers Theorem 8 for $r = 2$, although the authors do not mention it explicitly. As a side note, a symmetric $R \subseteq A^2$ (where every value in A appears in R) is reconfigurable if and only if the (undirected) graph $(A; R)$ is connected and not bipartite [49]. Thus, the Brakensiek-Guruswami conjecture [19], which postulates that $\mathrm{PCSP}(G, H)$ is NP-hard whenever G is not bipartite and H is loopless, already follows from the case $r = 2$ of Theorem 8, assuming the Rich 2-to-1 Conjecture.

In the language of [2], Theorem 8 states that symmetric reconfigurable relations are *promise-useless* – the extra information about the input instance that they provide does not help to find even a d-colouring that leaves no edge monochromatic (and hence cannot help us find a solution to *any* nontrivial hypergraph colouring problem). This result builds, albeit conditionally, on the hardness results of [29] (where they study PCSP templates of the form (R, NAE_3) such that no tuple in R is injective) and [22] (where they study PCSP templates of the form $(k\text{-in-}\ell \cup S, \mathsf{NAE}_\ell^{(2)})$ for some relation S over $\{0,1\}$).

Besides, the notion of reconfigurability in Theorem 8 hints towards its topological nature – we elaborate on this question in the Conclusions. It is interesting to note that this notion has appeared before in the Fourier-analytical research as well, e.g. Mossel [55] considers a relation *connected* if it is reconfigurable.

To illustrate the power of Theorem 8, we give an example from the world of approximate hypergraph colouring. Given a hypergraph $H = (V, E)$, a *Linearly Ordered (LO)* c-colouring is an assignment $V \to [c]$ to the vertices of H, such that in every hyperedge of H, its maximum colour appears exactly once. Quite some work has been done on the hardness of approximate LO colouring in uniform hypergraphs [8, 48, 57, 58]; for now, let us focus on the case of 3-uniform hypergraphs. For the following, let $(3, c, d)$-approximate LO colouring denote the problem of distinguishing whether a 3-uniform hypergraph is LO c-colourable, or not even LO d-colourable. The state-of-the-art hardness result is the following.

▶ **Theorem 9** (Simple reduction from [48])**.** $(3, c, c+1)$*-approximate LO colouring is NP-hard for $c \geq 2$.*

Theorem 8 completes the picture for $c \geq 3$, under the Rich 2-to-1 conjecture. Since the template corresponding to LO c-colouring a 3-uniform hypergraph with $c \geq 3$ colours is reconfigurable (cf. Figure 2), we deduce:

▶ **Corollary 10.** $(3, c, d)$*-approximate LO colouring is NP-hard for all $3 \leq c \leq d$, assuming the Rich 2-to-1 Conjecture.*

---

[4] A relation is symmetric if it is closed under permutations of coordinates.
[5] In fact, the reconfigurability assumption can be relaxed (see Section 5).

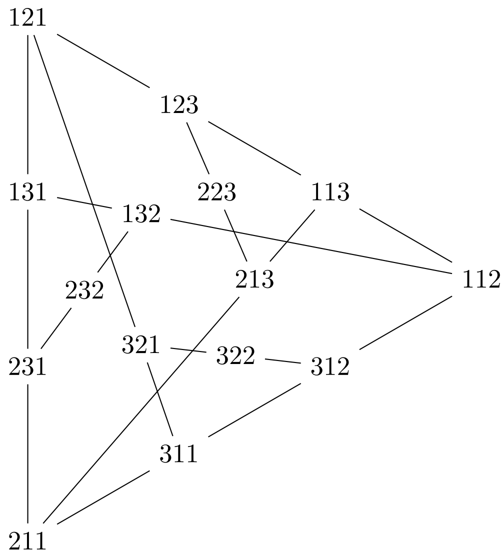

Figure 2 Reconfiguration graph for LO_3. The tuple (a, b, c) will be rendered abc for compactness.

This confirms (except for the case $c = 2$) the conjecture of Barto, Battistelli and Berg [8] on approximate LO colouring of 3-uniform hypergraphs, under the Rich 2-to-1 conjecture. We also remark that there is an essential link between a topological proof of the hardness of $(3, 3, 4)$-approximate LO colouring [36] and our proof here: the reconfigurability is an essential implication of [36].

It is worth noting that $(4, c, d)$-approximate LO colouring was recently shown to be unconditionally NP-hard for all $3 \le c \le d$ by [57].

Related work

We shall mention some previous work in the context of Theorem 8: it consists in various rainbow/balanced hypergraph colouring inapproximability results; we will use the PCSP parlance for clarity.

Observe that $\mathrm{PCSP}(R^{(c)}_{\ell:\ell}, R^{(d)}_{\ell:\ell})$ is just the approximate non-monochromatic hypergraph colouring problem. This problem serves as one of the key sources of hardness for PCSPs in general: see [9], which proves that a PCSP is NP-hard by gadget reduction from this problem if and only if it fails to admit an *Olšák polymorphism*, a condition which can be checked algorithmically; through this condition they were able to prove hardness of the approximate graph colouring $\mathrm{PCSP}(\mathbb{K}_c, \mathbb{K}_{2c-1})$ for $c \ge 3$.

In Table 1, we summarize existing hardness results for several regimes of promise hypergraph colouring that are subsumed, albeit conditionally, by this work. Finally, PCSPs of our form were mentioned in [60] as *Balanced Hypergraph Colouring* problems.

[6] This conjecture, which is, in some sense, a hypergraph generalisation of the "fish-shaped" conjecture of Dinur et al. [30], was proposed in [20].

[7] They prove it only for even c and $d = 2$, but with slight modifications to their proof, it is possible to generalise it to any $c, d \ge 2$.

Table 1 A summary of hypergraph colouring hardness results. These results only hold for choices of parameters where the PCSP template is valid, i.e. only when $\mathbb{A} \to \mathbb{B}$ for PCSP($\mathbb{A}, \mathbb{B}$).

Problem	Parameters	Assumptions	Reference
$\mathrm{PCSP}(R^{(c)}_{\ell:\ell}, R^{(d)}_{\ell:\ell})$	$\ell \geq 3$ and $d \geq c \geq 2$	–	[31, 57]
$\mathrm{PCSP}(R^{(\ell+1)}_{3:\ell}, R^{(d)}_{\ell:\ell})$	$\ell, d \geq 3$	V Label Cover[6]	[20]
$\mathrm{PCSP}(R^{(c)}_{q+c:cq}, R^{(d)}_{cq:cq})$	$q, c, d \geq 2$	–	[38]
$\mathrm{PCSP}(R^{(c)}_{q+d:cq}, R^{(d)}_{cq:cq})$	$q, c, d \geq 2$	–	[39][7]
$\mathrm{PCSP}(R^{(c)}_{k:\ell}, R^{(d)}_{\ell:\ell})$	$2 \leq \ell < c(k-1)$	Rich 2-to-1	This work

2 Preliminaries

Notation

We write $[n]$ for the set $\{1, \ldots, n\}$ and $\mathbb{N}$ for the set of positive integers. Given a set A, any element of A^n is an *n-ary tuple*. A tuple $\overline{x} \in A^n$ can be treated as a function $\overline{x}\colon [n] \to A$. Conversely, we sometimes treat any function from a finite set to A as a tuple, implicitly enumerating the elements of this finite set. The i-th entry of a tuple $\overline{x} \in A^n$ is denoted by x_i.

A (constraint) *language* is a set σ whose every member $R \in \sigma$, called a *relation symbol*, is assigned a positive integer $\mathrm{ar}(R)$, called the *arity* of R. For $\sigma = \{R_1, \ldots, R_n\}$, a (relational) *$\sigma$-structure* is a tuple $\mathbb{A} = (A, R_1^{\mathbb{A}}, \ldots, R_k^{\mathbb{A}})$, where A is a set and $R_i^{\mathbb{A}} \subseteq A^{\mathrm{ar}(R_i)}$, for each $i \in [k]$. The set A is called the *domain* of $\mathbb{A}$ and, for every $i \in [k]$, $R_i^{\mathbb{A}}$ is called the *interpretation of R_i in $\mathbb{A}$* or simply a *relation*. In this paper, all structures are to be assumed finite, unless stated otherwise.

For two σ-structures $\mathbb{A}$ and $\mathbb{B}$, a *homomorphism* from $\mathbb{A}$ to $\mathbb{B}$ is a map $h\colon \mathbb{A} \to \mathbb{B}$ that preserves relations.

MMSNP

Let τ be a fixed language and $\mathcal{F}$ be a finite family of pairs $(\mathbb{F}, f)$, where $\mathbb{F}$ is a finite τ-structure, and $f\colon F \to [c]$ for some $c \in \mathbb{N}$. The set $[c]$ is called the *colour set* of $\mathcal{F}$ and each $(\mathbb{F}, f) \in \mathcal{F}$ is called a *forbidden pattern* of $\mathcal{F}$.

A *c-expansion* (or a *c-colouring*) of a τ-structure $\mathbb{I}$ is a pair $(\mathbb{I}, g)$ where $g\colon I \to [c]$. If the number of colours is not important, we will just write an *expansion* (or a *colouring*). Conversely, the *reduct* of $(\mathbb{I}, g)$ is the τ-structure $\mathbb{I}$.

We say that an expansion $(\mathbb{I}, g)$ is *$\mathcal{F}$-free* if every $(\mathbb{F}, f) \in \mathcal{F}$ and every homomorphism h from $\mathbb{F}$ to $\mathbb{I}$ satisfies $f \neq g \circ h$.

▶ **Definition 11** (MMSNP). *Let τ be a constraint language. For a finite family $\mathcal{F}$ of forbidden patterns over τ, the problem* MMSNP($\mathcal{F}$) *comes in two flavours:*

(decision) *given a τ-structure $\mathbb{I}$, output* Yes *if $\mathbb{I}$ has an $\mathcal{F}$-free expansion, and output* No *otherwise;*

(search) *given a τ-structure $\mathbb{I}$, output an $\mathcal{F}$-free expansion.*

The decision MMSNP is poly-time equivalent to the search MMSNP: it follows from the poly-time equivalence of MMSNP and precoloured MMSNP [16, Corollary 6.10]. For the sake of brevity, we will write $\mathcal{F}$ instead of MMSNP($\mathcal{F}$) to denote the decision problem associated with $\mathcal{F}$.

Let $(\mathbb{I}, g)$ be a c-expansion of $\mathbb{I}$. Given a mapping $r\colon [c] \to [d]$, any d-expansion $(\mathbb{I}, h)$ that satisfies $r \circ g = h$ is called an *r-preimage* of $(\mathbb{I}, g)$.

Let $\mathcal{F}$ and $\mathcal{G}$ be two families of forbidden patterns over the same constraint language with colour sets $[c]$ and $[d]$, respectively. A mapping $r\colon [c] \to [d]$ is a *recolouring* from $\mathcal{F}$ to $\mathcal{G}$, if no forbidden pattern of $\mathcal{G}$ has a $\mathcal{F}$-free r-preimage. We write $\mathcal{F} \to \mathcal{G}$ if there is a recolouring from $\mathcal{F}$ to $\mathcal{G}$.

▶ **Definition 12** (PMMSNP). *Let τ be a constraint language. For finite families of forbidden patterns $\mathcal{F}$ and $\mathcal{G}$ over τ such that $\mathcal{F} \to \mathcal{G}$, the problem* $\mathrm{PMMSNP}(\mathcal{F}, \mathcal{G})$ *comes in two flavours:*

(decision) *given a τ-structure $\mathbb{I}$, output* Yes *if $\mathbb{I}$ has an $\mathcal{F}$-free expansion, and output* No *if $\mathbb{I}$ even does not have a $\mathcal{G}$-free expansion;*

(search) *given a τ-structure $\mathbb{I}$ promised to have an $\mathcal{F}$-free expansion, output a $\mathcal{G}$-free expansion of $\mathbb{I}$.*

CSP

It was already mentioned that every problem in MMSNP is poly-time equivalent to a *Constraint Satisfaction Problem* (CSP) [34].

▶ **Definition 13.** *Let σ be a constraint language. For a σ-structure $\mathbb{B}$, the problem* $\mathrm{CSP}(\mathbb{B})$ *comes in two flavours:*

(decision) *given a σ-structure $\mathbb{I}$, output* Yes *if $\mathbb{I} \to \mathbb{B}$, and output* No *if $\mathbb{I} \not\to \mathbb{B}$;*

(search) *given a σ-structure $\mathbb{I}$, output a homomorphism from $\mathbb{I}$ to $\mathbb{B}$.*

The decision CSP is poly-time equivalent to the search CSP [25, Corollary 4.9].

▶ **Definition 14** (PCSP). *Let σ be a relational language. For any two σ-structures $\mathbb{A}$ and $\mathbb{B}$ such that $\mathbb{A} \to \mathbb{B}$, the problem* $\mathrm{PCSP}(\mathbb{A}, \mathbb{B})$ *comes in two flavours:*

(decision) *given a σ-structure $\mathbb{I}$, output* Yes *if $\mathbb{I} \to \mathbb{A}$, and output* No *if $\mathbb{I} \not\to \mathbb{B}$;*

(search) *given a σ-structure $\mathbb{I}$ promised to admit a homomorphism to $\mathbb{A}$, output a homomorphism from $\mathbb{I}$ to $\mathbb{B}$.*

The search PCSP is at least as hard as the decision PCSP, however the converse is not known. We provide algorithms for the search version, and proofs of hardness for the decision version.

"Sandwiching" is a trivial reduction technique between PCSPs:

▶ **Observation 15** (See [9,19]). *If $\mathbb{A} \to \mathbb{B} \to \mathbb{C} \to \mathbb{D}$, then* $\mathrm{PCSP}(\mathbb{A}, \mathbb{D})$ *reduces to* $\mathrm{PCSP}(\mathbb{B}, \mathbb{C})$ *in poly-time.*

In the case $\mathbb{B} = \mathbb{C}$, the problem $\mathrm{PCSP}(\mathbb{B}, \mathbb{C})$ is just a CSP, namely $\mathrm{CSP}(\mathbb{B}) = \mathrm{CSP}(\mathbb{C})$.

If $R \subseteq A^r$ and $S \subseteq B^r$, we sometimes abuse the notation and write $\mathrm{PCSP}(R, S)$ in place of $\mathrm{PCSP}((A; R), (B; S))$ for convenience.

3 Equivalence with PCSP

In this section, we show that the problem $\mathrm{PMMSNP}(\mathcal{F}, \mathcal{G})$ is poly-time equivalent to some finite-domain PCSP. The reductions generalise the ones of Feder and Vardi [34], where the finite-domain CSP encoded all valid colourings of the reducts of the forbidden structures. However, in the promise setting, the two MMSNP sentences could have different sets of reducts, so applying Feder and Vardi's reduction to each MMSNP sentence would give two finite structures over different constraint languages. We resolve this issue.

▶ **Theorem 3.** *If $\mathcal{F} \to \mathcal{G}$, then there exist finite structures $\mathbb{S} \to \mathbb{T}$ such that* $\mathrm{PMMSNP}(\mathcal{F}, \mathcal{G})$ *and* $\mathrm{PCSP}(\mathbb{S}, \mathbb{T})$ *are poly-time equivalent.*

We first argue that $\mathcal{F}$ and $\mathcal{G}$ can be assumed to be in certain *normal form*. A structure $\mathbb{F}$ is *biconnected* if for every $x \in F$, the substructure of $\mathbb{F}$ induced on $F \setminus \{x\}$ is not isomorphic to the disjoint union of two structures. A family $\mathcal{F}$ is in *normal form* if every $\mathbb{F} \in \mathcal{F}$ is biconnected and if $\mathcal{F}$ is closed under taking quotients, i.e., identifying any two elements in any $\mathbb{F} \in \mathcal{F}$ results in a structure from $\mathcal{F}$. The following result allows us to assume the normal form without loss of generality.

▶ **Lemma 16** (Proposition 3.3 and Lemma 4.4 in [16]). *Every MMSNP problem is a finite disjunction of MMSNP problems in normal form.*

Further in this section, let $\mathcal{F} = \mathcal{F}_1 \vee \cdots \vee \mathcal{F}_m$ and $\mathcal{G} = \mathcal{G}_1 \vee \cdots \vee \mathcal{G}_n$ be two disjunctions of MMSNPs in normal form such that $\mathcal{F} \to \mathcal{G}$. Denote the colour set of each $\mathcal{F}_i$ (resp. $\mathcal{G}_j$) by $[c_i]$ (resp. $[d_j]$) for some $c_i \in \mathbb{N}$ (resp. $d_j \in \mathbb{N}$).

▶ **Observation 17.** *For every $i \in [m]$ there is $j \in [n]$ such that $\mathcal{F}_i \to \mathcal{G}_j$.*

We shall define $\mathbb{S}$ and $\mathbb{T}$ via a generalisation of the construction given in the Introduction. Let τ be the language of both $\mathcal{F}$ and $\mathcal{G}$. For each $j \in [n]$, let $\mathcal{G}_j^*$ denote the set of all reducts of forbidden patterns in $\mathcal{G}_j$; for each $i \in [m]$, define $\mathcal{F}_i^*$ similarly. Denote by σ the language which contains, for every $j \in [n]$ and every $\mathbb{G} \in \mathcal{G}_j^*$, a $|G|$-ary symbol $R_{\mathbb{G}}$. We also expand σ with an auxiliary binary symbol $\sim$. This will be the constraint language of both $\mathbb{S}$ and $\mathbb{T}$.

Fix $j \in [n]$; let $\mathbb{T}_j$ be the σ-structure with domain $[d_j]$, where

- for every $\mathbb{G} \in \mathcal{G}_j^*$, the relation $R_{\mathbb{G}}^{\mathbb{T}_j} := \{\bar{t} \colon G \to [d_j] \mid (\mathbb{G}, \bar{t}) \notin \mathcal{G}_j\}$;
- for every $j' \in [n] \setminus \{j\}$ and every $\mathbb{G} \in \mathcal{G}_{j'}^*$, the relations $R_{\mathbb{G}}^{\mathbb{T}_j} := [d_j]^G$ and $\sim^{\mathbb{T}_j} := [d_j]^2$ are full.

Fix $i \in [m]$; by Observation 17, $\mathcal{F}_i \to \mathcal{G}_j$ for some $j \in [n]$. Let $\mathbb{S}_i$ be the σ-structure with domain $[c_i]$, where

- for every $\mathbb{G} \in \mathcal{G}_j^*$, the relation $R_{\mathbb{G}}^{\mathbb{S}_i} := \{\bar{s} \colon G \to [c_i] \mid (\mathbb{G}, \bar{s}) \text{ is } \mathcal{F}_i\text{-free}\}$;
- for every $j' \in [n] \setminus \{j\}$ and every $\mathbb{G} \in \mathcal{G}_{j'}^*$, the relations $R_{\mathbb{G}}^{\mathbb{S}_i} := [c_i]^G$ and $\sim^{\mathbb{S}_i} := [c_i]^2$ are full.

Finally, we define $\mathbb{S} := \mathbb{S}_1 \sqcup \cdots \sqcup \mathbb{S}_m$ and $\mathbb{T} := \mathbb{T}_1 \sqcup \cdots \sqcup \mathbb{T}_n$, where $\sqcup$ stands for the disjoint union. We will argue that $\mathrm{PMMSNP}(\mathcal{F}, \mathcal{G})$ is poly-time equivalent to $\mathrm{PCSP}(\mathbb{S}, \mathbb{T})$, concluding the proof of Theorem 3.

▷ Claim 18. $\mathrm{PMMSNP}(\mathcal{F}, \mathcal{G})$ poly-time reduces to $\mathrm{PCSP}(\mathbb{S}, \mathbb{T})$.

Proof. Given an instance of $\mathrm{PMMSNP}(\mathcal{F}, \mathcal{G})$ – a τ-structure $\mathbb{X}$ – we construct an instance of $\mathrm{PCSP}(\mathbb{S}, \mathbb{T})$ – a σ-structure which we denote by $\sigma(\mathbb{X})$. Namely, let $\sigma(\mathbb{X})$ be the structure with the same domain as $\mathbb{X}$ in which every relation $R_{\mathbb{G}}^{\sigma(\mathbb{X})} := \mathrm{Hom}(\mathbb{G}, \mathbb{X})$ is defined as the set of all homomorphisms from $\mathbb{G}$ to $\mathbb{X}$. The relation $\sim^{\sigma(\mathbb{X})} := X^2$ is full.

If $\mathbb{X}$ is a Yes instance of $\mathrm{MMSNP}(\mathcal{F})$, then there is $s \colon X \to [c_i]$ for some $i \in [m]$ such that $(\mathbb{X}, s)$ is $\mathcal{F}_i$-free. It follows from the construction of $\sigma(\mathbb{X})$ and $\mathbb{S}_i$ that s is a homomorphism from $\sigma(\mathbb{X})$ to $\mathbb{S}_i$.

Conversely, if $h \colon X \to T$ is a homomorphism from $\sigma(\mathbb{X})$ to $\mathbb{T}$, then the image of h is contained in some $\mathbb{T}_j$, as $\sim^{\sigma(\mathbb{X})} = X^2$. It follows from the construction of $\sigma(\mathbb{X})$ that h colours every copy of every $\mathbb{G} \in \mathcal{G}_j$ in an allowed way, so $(\mathbb{X}, h)$ is $\mathcal{G}$-free. ◁

For the converse reduction, we make use of the Sparse Incomparability Lemma which allows us to assume no short cycles in structures on input to $\mathrm{PCSP}(\mathbb{S},\mathbb{T})$. We say that $\mathbb{S}$ has girth $> k$ if for any choice of $\ell \le k$ relational tuples $\bar{s}_1,\dots,\bar{s}_\ell$ of arities $r_1,\dots,r_\ell$ in $\mathbb{S}$, the total number of elements appearing in the tuples is greater than $\sum_{i=1}^{\ell}(r_i-1)$. That is, k or fewer relational tuples never form a cycle.

▶ **Lemma 19** (Sparse Incomparability Lemma [34, 50])**.** *Let σ be a finite relational language, and $k,t \in \mathbb{N}$. For every σ-structure $\mathbb{I}$ there exists a poly-time constructible σ-structure $\mathbb{I}'$ with girth $> k$ such that for every σ-structure $\mathbb{T}$ of size $< t$ the equivalence $\mathbb{I} \to \mathbb{T} \Leftrightarrow \mathbb{I}' \to \mathbb{T}$ holds. Moreover, $\mathbb{I}' \to \mathbb{I}$.*

Furthermore, we can assume that input structures are *connected*, i.e. not isomorphic to a disjoint union of two structures. To see why this is the case, observe that $\mathbb{A} \sqcup \mathbb{B} \to \mathbb{S}$ if and only if $\mathbb{A} \to \mathbb{S}$ and $\mathbb{B} \to \mathbb{S}$. This implies that we can solve a disconnected instance of $\mathrm{PCSP}(\mathbb{S},\mathbb{T})$ by solving its every connected component. This means that $\mathbb{I} \to \mathbb{S}$ if and only if $\mathbb{I} \to \mathbb{S}_i$, for some $i \in [m]$, and similarly $\mathbb{I} \to \mathbb{T} \Leftrightarrow \mathbb{I} \to \mathbb{T}_j$, for some $j \in [n]$. Since $\sim^{\mathbb{T}_j} = T_j^2$ is full, we can also assume that $\sim^{\mathbb{I}} = I^2$ is full.

▷ Claim 20. $\mathrm{PCSP}(\mathbb{S},\mathbb{T})$ poly-time reduces to $\mathrm{PMMSNP}(\mathcal{F},\mathcal{G})$.

Proof. Given an instance of $\mathrm{PCSP}(\mathbb{S},\mathbb{T})$ – a σ-structure $\mathbb{I}$ – we construct an instance of $\mathrm{PMMSNP}(\mathcal{F},\mathcal{G})$ – a τ-structure which we denote by $\tau(\mathbb{I})$.

We first obtain $\mathbb{I}'$ by applying Lemma 19 with parameters $k > \max\{|G| \mid \mathbb{G} \in \mathcal{F}_1 \cup \dots \cup \mathcal{F}_m \cup \mathcal{G}_1 \cup \dots \cup \mathcal{G}_n\}$ and $t > \max\{|S|,|T|\}$ and also taking the auxiliary relation $\sim^{\mathbb{I}'} = I'^2$ to be full. Denote by $\tau(\mathbb{I})$ the τ-structure with the same domain as $\mathbb{I}'$ which is obtained by replacing every $\bar{t} \in R_{\mathbb{G}}^{\mathbb{I}'}$ with a copy of $\mathbb{G}$ induced on $\bar{t}$, for every $\mathbb{G} \in \mathcal{G}_j^*$. Since all $\mathcal{F}_i$ and $\mathcal{G}_j$ are in normal form and since $\mathbb{I}'$ has large girth, every pair $\mathcal{F}_i \to \mathcal{G}_j$ has the following property: if there is a colouring s of $\tau(\mathbb{I})$ that leaves every copy of $\mathbb{G}$ $\mathcal{F}_i$-free (resp. $\mathcal{G}_j$-free) for each $\mathbb{G} \in \mathcal{G}_j^*$, then $(\tau(\mathbb{I}),s)$ is $\mathcal{F}_i$-free (resp. $\mathcal{G}_j$-free).

If $\mathbb{I} \to \mathbb{S}_i$, then also $\mathbb{I}' \to \mathbb{I} \to \mathbb{S}_i$ by Lemma 19. If h is a homomorphism from $\mathbb{I}'$ to $\mathbb{S}_i$, then it colours all the copies of structures from $\mathcal{G}_j^*$ in $\tau(\mathbb{I})$ such that each of them is $\mathcal{F}_i$-free, where j is witnessing the containment $\mathcal{F}_i \to \mathcal{G}_j$. This implies that $(\tau(\mathbb{I}),h)$ is $\mathcal{F}_i$-free, and $\tau(\mathbb{I})$ is a Yes instance of $\mathrm{MMSNP}(\mathcal{F})$.

Conversely, if $\tau(\mathbb{I})$ is a Yes instance of $\mathrm{MMSNP}(\mathcal{G})$, then $\tau(\mathbb{I})$ has a $\mathcal{G}_j$-free expansion for some $j \in [n]$, witnessed by $s\colon I' \to [d_j]$. Since every copy of $\mathbb{G}$ in $\tau(\mathbb{I})$ is $\mathcal{G}_j$-free, for each $\mathbb{G} \in \mathcal{G}_j^*$, the same mapping s is a homomorphism from $\mathbb{I}'$ to $\mathbb{T}_j$. By Lemma 19, we then have $\mathbb{I} \to \mathbb{T}_j \to \mathbb{T}$. ◁

This concludes the proof of PMMSNP–PCSP correspondence.

4 Forbidding monochromatic cliques

In this section we provide a complete complexity classification for Promise MMSNPs defined by forbidding monochromatic cliques. Recall that $\mathcal{F}_k^{(c)}$ denotes the family of all k-cliques such that every two vertices in the same clique have the same colour out of $\{1,\dots,c\}$.

▶ **Theorem 5.** *Let c,d,k,ℓ be positive integers such that $\mathcal{F}_k^{(c)} \to \mathcal{F}_\ell^{(d)}$. Then*

- *$\ell \ge c(k-1)$ and $\mathrm{PMMSNP}(\mathcal{F}_k^{(c)},\mathcal{F}_\ell^{(d)})$ is solvable in polynomial time via AIP, or*
- *$\ell < c(k-1)$ and $\mathrm{PMMSNP}(\mathcal{F}_k^{(c)},\mathcal{F}_\ell^{(d)})$ is NP-hard, assuming the Rich 2-to-1 Conjecture.*

Theorem 3 asserts that we can study Promise CSPs defined by suitable relations instead of Promise MMSNPs. Concretely, for $c, k, \ell \geq 1$, let $R^{(c)}_{k:\ell} \subseteq [c]^\ell$ be the relation consisting of all tuples $(a_1, \ldots, a_\ell)$ such that $\forall a \in [c] : \ |\{i : a_i = a\}| < k$. Observe that $R^{(d)}_{\ell:\ell}$ is isomorphic to $\mathsf{NAE}^{(d)}_\ell$.

Theorem 3 applied to $(\mathcal{F}^{(c)}_k, \mathcal{F}^{(d)}_\ell)$ provides the structures $\mathbb{S}$ and $\mathbb{T}$. Following the construction in Section 3 (in fact, even the reasoning similar to Example 4 suffices), the reader may verify that these structures correspond to $R^{(c)}_{k:\ell}$ and $R^{(d)}_{\ell:\ell}$ respectively – the extra $\sim$ relation is full in both structures, and can be safely discarded without affecting the complexity of $\mathrm{PCSP}(\mathbb{S}, \mathbb{T})$.

▶ **Lemma 6** (Special case of Theorem 3). *Let c, d, k, ℓ be positive integers. If $\mathcal{F}^{(c)}_k \to \mathcal{F}^{(d)}_\ell$, then the problems $\mathrm{PMMSNP}(\mathcal{F}^{(c)}_k, \mathcal{F}^{(d)}_\ell)$ and $\mathrm{PCSP}(R^{(c)}_{k:\ell}, R^{(d)}_{\ell:\ell})$ are poly-time equivalent.*

The rest of this section is organized as follows. First, we want to characterise the containment $\mathcal{F}^{(c)}_k \to \mathcal{F}^{(d)}_\ell$ in terms of the relation between the numbers c, d, k, and ℓ. Next, the proof of the tractability part of the dichotomy is presented. Finally, we prove the hardness part, which is incidentally the most technical part: it requires a basic introduction to Fourier analysis and related notions.

Understanding containment

▶ **Lemma 21.** *Let c, d, k, ℓ be positive integers. Then $\mathcal{F}^{(c)}_k \to \mathcal{F}^{(d)}_\ell$ if and only if $k \leq \lceil \ell / \lceil c/d \rceil \rceil$.*

We call this inequality the *containment condition* for c, d, k, ℓ.

Proof. Suppose that $k \leq \lceil \ell / \lceil c/d \rceil \rceil$. Fix any mapping $r\colon [c] \to [d]$ such that $|r^{-1}(i)| \in \{\lfloor c/d \rfloor, \lceil c/d \rceil\}$ for every $i \in [d]$. Every r-preimage of every monochromatic $\mathbb{K}_\ell$ contains at least $\lceil \ell / \lceil c/d \rceil \rceil$ vertices of the same colour, so it contains a monochromatic copy of $\mathbb{K}_k$, which implies that r is a recolouring.

For the opposite direction, let $r\colon [c] \to [d]$ be an arbitrary recolouring from $\mathcal{F}^{(c)}_k$ to $\mathcal{F}^{(d)}_\ell$. By the pigeonhole principle, there must exist $i \in [d]$ such that $|r^{-1}(i)| \geq \lceil c/d \rceil$. Every r-preimage of $(\mathbb{K}_\ell, v \mapsto i) \in \mathcal{F}^{(d)}_\ell$ contains

- at least $\lceil \ell / \lceil c/d \rceil \rceil$ vertices of the same colour, and
- a monochromatic copy of $\mathbb{K}_k$, since r is a recolouring.

Consequently $k \leq \lceil \ell / \lceil c/d \rceil \rceil$. ◀

In particular, observe that if $c, d, k, \ell > 0$ satisfy the containment condition, then $k \leq \ell$.

Tractability

We shall prove the first part of Theorem 5.

▶ **Theorem 22.** *Let c, d, k, ℓ be positive integers satisfying the containment condition such that $\ell \geq c(k-1)$. Then $\mathrm{PMMSNP}(\mathcal{F}^{(c)}_k, \mathcal{F}^{(d)}_\ell)$ is in P.*

Proof. By Lemma 6, it suffices to prove that $\mathrm{PCSP}(R^{(c)}_{k:\ell}, R^{(d)}_{\ell:\ell})$ is in P. If $\ell > c(k-1)$, then the left-hand side relation is empty, so the problem is trivially tractable. Similarly if $d = 1$, then the right-hand side relation is empty, so the problem is yet again trivial. Thus assume that $\ell = c(k-1)$ and $d \geq 2$. We will argue that

$$R^{(c)}_{k:\ell} \to (k-1)\text{-in-}\ell \to \mathsf{NAE}^{(2)}_\ell \to R^{(d)}_{\ell:\ell}.$$

Then, tractability follows by Observation 15 from the tractability of

$$\mathrm{PCSP}\left((k-1)\text{-in-}\ell, \mathsf{NAE}_\ell^{(2)}\right),$$

which, in turn, follows from the techniques of [19], namely AIP. Let us therefore show the required homomorphisms.

$\boldsymbol{R_{k:\ell}^{(c)} \to (k-1)}$**-in-**$\boldsymbol{\ell}$**.** Since $\ell = c(k-1)$, it follows that each tuple of $R_{k:\ell}^{(c)}$ consists of exactly $k-1$ appearances of each of the c colours in our domain. Mapping one fixed colour to 1 and the rest to 0 gives rise to a homomorphism to $(k-1)$-in-ℓ.

$\boldsymbol{\mathsf{NAE}_\ell^{(2)} \to R_{\ell:\ell}^{(d)}}$**.** Mapping 0 and 1 to any two distinct colours in the domain of $R_{\ell:\ell}^{(d)}$ suffices. (Such colours exist as $d \geq 2$.) ◀

Hardness

For the second part of Theorem 5, we prove the following.

▶ **Theorem 23.** *Let c, d, k, ℓ be positive integers satisfying the containment condition such that $\ell < c(k-1)$. Then $\mathrm{PMMSNP}(\mathcal{F}_k^{(c)}, \mathcal{F}_\ell^{(d)})$ is NP-hard, assuming the Rich 2-to-1 Conjecture.*

Braverman et al. [23] recently presented a yet different proof of conditional hardness for $\mathrm{PCSP}(\mathbb{K}_3, \mathbb{K}_\ell)$ using their analysis of the *invariance principle* for multi-slices, and reducing from the Rich 2-to-1 Conjecture.[8] We follow their approach to derive hardness for the PCSPs of the following form:

▶ **Theorem 8.** *Let $\varnothing \neq R \subseteq A^r$ be a symmetric[9] reconfigurable[10] relation of arity $r \geq 2$. For any d such that $R \to \mathsf{NAE}_r^{(d)}$, the problem $\mathrm{PCSP}(R, \mathsf{NAE}_r^{(d)})$ is NP-hard, assuming the Rich 2-to-1 Conjecture.*

The proof of this result is rather technical, so we will first show how it implies Theorem 23, and then devote the remaining part of this paper to proving Theorem 8.

Proof of Theorem 23. Fix c, d, k, ℓ such that $\ell < c(k-1)$. First, note that by the containment condition

$$k \leq \left\lceil \frac{\ell}{\lceil \frac{c}{d} \rceil} \right\rceil \leq \ell,$$

hence $\ell < c(k-1) \leq c(\ell-1)$. This means that $c, d, k, \ell \geq 2$.

By Lemma 6, it suffices to show that $\mathrm{PCSP}(R_{k:\ell}^{(c)}, R_{\ell:\ell}^{(d)})$ is NP-hard, assuming the Rich 2-to-1 Conjecture. We are going to apply Theorem 8.

Let $R := R_{k:\ell}^{(c)}$ which is a symmetric relation of arity ℓ. Since $\lceil \ell/c \rceil \leq 1 + \ell/c < k$, it is non-empty.

Recall that R consists of all tuples of length ℓ, using only elements in $[c]$, where every element appears less than k times. If we start from some arbitrary tuple of R, and replace (one of) the most frequent colours with (one of) the least frequent colours, we can reach any tuple where each colour appears either $\lfloor \ell/c \rfloor$ or $\lceil \ell/c \rceil$ times. This implies that every tuple in R can be reconfigured to the uniquely chosen "almost balanced" tuple in R.

Therefore $R \neq \varnothing$ is a symmetric reconfigurable relation. On the other hand, $R_{\ell:\ell}^{(d)} = \mathsf{NAE}_\ell^{(d)}$. We conclude the proof by applying Theorem 8. ◀

---

[8] Earlier, Guruswami and Sandeep [40] proved this under the weaker d-to-1 Conjecture. However, we were not able to extend their techniques to our setting.

[9] A relation is symmetric if it is closed under permutations of coordinates.

[10] In fact, the reconfigurability assumption can be relaxed (see Section 5).

5 PCSP hardness

This section is devoted to the proof of Theorem 8. As mentioned earlier, it is a special case of a more general hardness result that we prove. More precisely, the *reconfigurability* shall be relaxed to *connectedness* of Braverman, Khot, Lifshitz, and Minzer [23]; we call it *BKLM-connectedness* for clarity.[11]

▶ **Definition 24** (Based on [23]). *Given a relation $R \subseteq A^r$ and $1 \leq i < r$, we define $\mathbb{G}_i^R$ to be the bipartite graph whose left side consists of all prefixes of tuples in R of length i, whose right side consists of all suffixes of tuples in R of length $r - i$, and where a prefix and a suffix are connected if they form a tuple in R together.*

A relation R is BKLM-connected *if $\mathbb{G}_i^R$ is connected for every $1 \leq i < r$.*

▶ **Theorem 25.** *Let $\emptyset \neq R \subseteq A^r$ be a symmetric BKLM-connected relation of arity $r \geq 2$. For any d such that $R \to \mathsf{NAE}_r^{(d)}$, the problem $\mathrm{PCSP}(R, \mathsf{NAE}_r^{(d)})$ is NP-hard, assuming the Rich 2-to-1 Conjecture.*

Let us first argue that Theorem 8 indeed follows from Theorem 25. This amounts to proving that symmetric reconfigurable relations are BKLM-connected.

▶ **Proposition 26.** *Let $R \subseteq A^r$ be a symmetric reconfigurable relation. Then R is BKLM-connected.*

Proof. Fix $1 \leq i < r$, and consider the bipartite graph $\mathbb{G}_i^R$. Since this graph has no isolated vertices by construction, it is connected if and only if its *line graph*[12] is connected.

Consider now the *reconfiguration graph* of R: the vertices of this graph are the tuples of R; two tuples are connected whenever they differ in precisely one coordinate. Clearly, the reconfiguration graph is a subgraph of the line graph of $\mathbb{G}_i^R$, and they have the same vertex sets. Since the former is connected by the assumption, the latter must be connected as well, which concludes the argument. ◀

The proof of Theorem 25 follows the strategy of [23] through which its authors proved hardness of Approximate Graph Colouring, assuming the Rich 2-to-1 Conjecture. Its nature is heavily analytical; in particular, it relies on the *invariance principle* on *multislice* domains developed in the same paper.

The rest of this section is organized as follows. First we define the Rich 2-to-1 Conjecture, the source of hardness for our PCSPs. Then we introduce the necessary notions from the Fourier analysis. Finally, we present the proof of Theorem 25.

Rich 2-to-1 Conjecture

We will have to concede the CSP formulation of the Rich 2-to-1 Conjecture, used in the Introduction, in favour of the more commonly used *Label Cover* one. For the following, a map $\pi : [2n] \to [n]$ is called *2-to-1* if $|\pi^{-1}(i)| = 2$ for each $i \in [n]$.

---

[11] To be precise, their definition of connectedness is less restrictive, but the semantics of connectivity remains the same.

[12] The line graph of $\mathbb{G}_i^R$ has as vertices the edges of $\mathbb{G}_i^R$ (i.e. tuples of R), and two edges are connected whenever one of their endpoints coincides (i.e. two tuples are connected if they agree on the prefix or on the suffix).

▶ **Definition 27** (Rich 2-to-1 Label Cover). *A Rich 2-to-1 Label Cover instance is a tuple $(U \cup V, E, [2n], [n], \{\pi_e\}_{e\in E})$, where $(U \cup V, E)$ form a bi-regular bipartite graph, and each $\pi_e : [2n] \to [n]$ is a 2-to-1 map. Furthermore, for every fixed $u \in U$, if we sample v from the neighbourhood of u uniformly at random, then the function π_{uv} is uniformly distributed among the set of 2-to-1 maps $[2n] \to [n]$ (this property is called* richness*).*[13]

A labelling *for this instance is an assignment c which sends U to $[2n]$ and V to $[n]$. The value of the labelling is the probability that $c(v) = \pi_{uv}(c(u))$ for (u, v) sampled uniformly at random from E. The value of the instance is the maximum value of any labelling.*

▶ **Conjecture 28** (Rich 2-to-1 conjecture [24]). *For every $\varepsilon > 0$ there exists a large enough n such that it is NP-hard to decide whether a Rich 2-to-1 Label Cover instance $(U \cup V, E, [2n], [n], \{\pi_e\}_{e\in E})$ has value 1, or value $< \varepsilon$.*

For a tuple $\overline{x} \in A^n$ and a map $\pi\colon [2n] \to [n]$, the pullback of $\overline{x}$ along π is the tuple $\overline{y} \in A^{2n}$ such that $y_j = x_{\pi(j)}$ for all $j \in [2n]$. We will usually denote this tuple by $\pi^{-1}(\overline{x})$, abusing the notation.

Conversely, for a function f on $2n$ variables and a map $\pi\colon [2n] \to [n]$, we denote by f^{π} the n-ary function g defined by pulling back its input and passing to f, that is,

$$g(x_1, \dots, x_n) = f(x_{\pi(1)}, \dots, x_{\pi(2n)}).$$

In this case, we also write $f \xrightarrow{\pi} g$. This operation is known as a *minor* in PCSP literature.

Informal overview of Theorem 25

Let us preface this overview by emphasizing that our proof is largely based on the proof of [23, Theorem 6.7]; in particular, the reduction [23, Section 6.2], the completeness proof [23, Section 6.3], and a part of the soundness proof [23, Section 6.4] are followed verbatim, with a few simplifications. Due to this, and the lack of space, we only present a proof sketch here, deferring some technical details to the full version of the paper.

▶ **Theorem 25.** *Let $\varnothing \neq R \subseteq A^r$ be a symmetric BKLM-connected relation of arity $r \geq 2$. For any d such that $R \to \mathsf{NAE}_r^{(d)}$, the problem $\mathrm{PCSP}(R, \mathsf{NAE}_r^{(d)})$ is NP-hard, assuming the Rich 2-to-1 Conjecture.*

We will assume that each $a \in A$ can be extended to a tuple $(a, a_2, \dots, a_r) \in R$; otherwise such a can be removed from A without affecting the complexity of the PCSP in question. Throughout the proof, r, d, and $|A|$ are treated as constants.

In order to prove Theorem 25, we shall reduce from the Rich 2-to-1 Conjecture (see Conjecture 28), transforming Rich 2-to-1 Label Cover instances into instances of $\mathrm{PCSP}(R, \mathsf{NAE}_r)$, i.e., r-uniform hypergraphs. Fix the value of n obtained from Conjecture 28 for a small $\varepsilon > 0$ that we shall define later. Let $\Phi = (U \cup V, E, [2n], [n], \{\pi_e\}_{e\in E})$ be a Rich 2-to-1 Label Cover instance. Let us recall the standard Long Code transformation (see e.g. [9]): each vertex $u \in U$ is replaced with a copy of A^{2n} (called *the cloud of u*), each $v \in V$ is replaced with a copy of A^n (the cloud of v); every edge $(u, v) \in E$ commands identification of all pairs of tuples $\overline{x} \in A^{2n}$ and $\overline{y} \in A^n$ in the clouds of u and v such that $\overline{x} = \pi_{uv}^{-1}(\overline{y})$; additionally, some hyperedges are added, which we discuss later.

Our transformation differs from the above only in the use of a "shorter" code: instead of the whole Cartesian power of A, we are content with a certain subset of it, called *multislice*.

[13] In other words, if one fixes any $u \in U$ and considers the maps π_e associated with the edges incident to u, every 2-to-1 map appears equally often.

▶ **Definition 29** (Multislice). *For any $n \in \mathbb{N}$ and $\overline{\#} \in \{1,\dots,n\}^A$ such that $\sum_a \#_a = n$, let $\mathcal{S}\overline{\#}$ be the subset of A^n consisting of all tuples in which each $a \in A$ appears exactly $\#_a$ times. By $2\overline{\#}$ we denote the tuple $(2\#_a)_a \in \{0,\dots,2n\}^A$. If $\#_a \geq \alpha \cdot n$ for all $a \in A$, then we call the tuple $\overline{\#}$ α-*balanced.

The Long Code reduction framework further stipulates that, for any $v_1,\dots,v_r \in V$ with a common neighbour $u \in U$ and any $\overline{y}^{(1)},\dots,\overline{y}^{(r)}$ from the respective clouds, a hyperedge $(\overline{y}^{(1)},\dots,\overline{y}^{(r)})$ is added whenever $(\pi^{-1}_{uv_1}(y^{(1)})_i,\dots,\pi^{-1}_{uv_r}(y^{(r)})_i) \in R$ for all $i \in [2n]$. Since our domain is stripped down to a union of multislices, not all of these hyperedges exist – we only add some of them. To be precise, let Ω be a probability distribution over R. We define the relation μ_n^Ω as the following set[14]

$$\left\{(\overline{x}^{(1)},\dots,\overline{x}^{(r)}) \in \prod_{i=1}^{r} A^{2n} \;:\; \forall \overline{a} \in R. \left|\left\{j : (x_j^{(1)},\dots,x_j^{(r)}) = \overline{a}\right\}\right| = 2n \cdot \Omega(\overline{a})\right\}.$$

Intuitively, one can think of all $r \times 2n$ matrices in which the frequency of each column $\overline{a}$ is proportional to $\Omega(\overline{a})$. Indeed, the uniform distribution over the set μ_n^Ω is designed to "simulate" Ω^{2n}.

Observe that μ_n^Ω is, in fact, a relation over a product of r multislices. Now, we only include those hyperedges which satisfy $(\pi^{-1}_{uv_1}(y^{(1)}),\dots,\pi^{-1}_{uv_r}(y^{(r)})) \in \mu_n^\Omega$.

The reduction is summarized concisely in the following paragraph. We will make use of *marginal distributions*: Given a distribution Ω over $A_1 \times \dots \times A_r$, its *marginal distribution* to $i \in [r]$ is a distribution Ω_i over A_i defined as

$$\Omega_i(a) = \sum_{a_1 \in A_1,\dots,a_r \in A_r} \Omega(a_1,\dots,a_{i-1},a,a_{i+1},\dots,a_r)$$

for each $a \in A_i$.

Reduction

Let $\Phi = (U \cup V, E, [2n], [n], \{\pi_e\}_{e\in E})$ be a Rich 2-to-1 Label Cover instance. The reduction produces an instance of PCSP$(R, \mathsf{NAE}_r^{(d)})$, that is, an r-uniform hypergraph H. Let Ω be the uniform distribution over R; to simplify the presentation of the proof, we shall assume that $n \cdot \Omega(\overline{a}) \in \mathbb{N}$ for every $\overline{a} \in R$ (this is not necessarily the case, but it suffices to tweak the distribution Ω just slightly – we defer the technical details to the full version of the paper).

Denote by $\rho \coloneqq \Omega_i$ the marginal distribution of Ω to any coordinate $i \in [r]$ (they are all equal as Ω is uniform over a symmetric relation). Put $\overline{\#} = (n \cdot \rho(a))_a \in \mathbb{N}^A$ and note that $\overline{\#}$ is α-balanced for $\alpha \coloneqq \min_{\overline{a}\in R} \Omega(\overline{a}) = 1/|R| > 0$.

We replace each vertex $v \in V$ with a copy of the multislice $\mathcal{S}\overline{\#}$, and represent its elements as pairs $(v, \overline{y})$ where $\overline{y} \in \mathcal{S}\overline{\#}$. All these pairs form the set of vertices $V(H)$. One could also add multislices arising from the vertices $u \in U$ to $V(H)$, but they will be rendered redundant by the following paragraph anyway.

We define the relation $\sim$ on $V(H)$ by $(v,\overline{y}) \sim (v',\overline{y}')$ if there is a common neighbour $u \in U$ of v and v', and a tuple $\overline{x} \in \mathcal{S}2\overline{\#}$ such that $\overline{x} = \pi^{-1}_{u,v}(\overline{y}) = \pi^{-1}_{u,v'}(\overline{y}')$. We extend $\sim$ to an equivalence relation on $V(H)$, which, abusing notation, we also denote by $\sim$. In the hypergraph H, we identify all vertices in the same equivalence class of $\sim$, which can be done in logarithmic space [59].

[14] Of course, this requires that $2n \cdot \Omega(\overline{a}) \in \mathbb{Z}$ for any $\overline{a} \in R$.

For every $v_1, \dots, v_r \in V$ with a common neighbour $u \in U$ and every $(\overline{y}^{(1)}, \dots, \overline{y}^{(r)})$ such that $(\pi_{uv_1}^{-1}(y^{(1)}), \dots, \pi_{uv_r}^{-1}(y^{(r)})) \in \mu_n^{\Omega}$, add to H the hyperedge over

$$\left(v_1, \overline{y}^{(1)}\right), \dots, \left(v_r, \overline{y}^{(r)}\right).$$

▷ Claim 30 (Completeness). If Φ (Label Cover instance) has a perfect labelling, then $H \to R$.

Proof. Let $s\colon U \to [2n]$ and $s'\colon V \to [n]$ be labellings that satisfy all the constraints of Φ. We define the assignment $h : V(H) \to A$ as $h(v, \overline{y}) = y_{s'(v)}$.

We will argue that h is a homomorphism from H to R. Indeed, consider a hyperedge

$$(v_1, \overline{y}^{(1)}), \dots, (v_r, \overline{y}^{(r)}).$$

witnessed by a common neighbour $u \in U$. It suffices to prove that

$$\left(h(v_1, \overline{y}^{(1)}), \dots, h(v_r, \overline{y}^{(r)})\right) = \left(y^{(1)}_{s'(v_1)}, \dots, y^{(r)}_{s'(v_r)}\right) \in R.$$

Thanks to the equivalence constraints $\sim$, the right-hand side tuple is exactly

$$\left(\pi_{uv_1}^{-1}(y^{(1)})_{s(u)}, \dots, \pi_{uv_r}^{-1}(y^{(r)})_{s(u)}\right)$$

which belongs to R by the definition of μ_n^{Ω}. ◁

Soundness

This part is much more complicated than completeness. To spell it out, given a $\mathsf{NAE}_r^{(d)}$-colouring of the hypergraph, we ought to construct a labelling of the original Label Cover instance with a superconstant value. Historically, Fourier analysis has been central in constructing such a labelling. Let us review some standard probabilistic notation commonly used in Fourier-analytic arguments.

For a function $f : X \to \mathbb{R}$ and a distribution Ω over X, we denote by $\mathbb{E}_{\Omega}[f]$ the expectation of $f(x)$ when x is sampled from Ω (which we denote by $x \sim \Omega$). When we write $\mathbb{E}_X[f]$ or $\mathbb{E}[f]$, the uniform distribution over X (the domain of f) is assumed implicitly. By Ω^n we denote the n-fold tensor power of Ω, that is, the distribution over X^n such that the measure $\Omega^n(x_1, \dots, x_n) = \prod_{j=1}^n \Omega(x_j)$.

Cartesian powers and multislices are connected via the following lifting operator, which is a Markov operator of a certain *coupling* defined in [23]. Given $\overline{\#} \in \mathbb{N}^A$, there exists an operator T that maps any function $f\colon \mathcal{S}\overline{\#} \to \mathbb{R}$ to a function $Tf\colon A^{\sum_a \#_a} \to \mathbb{R}$ that satisfies

$$\mathop{\mathbb{E}}_{\mathcal{S}\overline{\#}}[f] = \mathop{\mathbb{E}}_{\rho^n}[Tf]$$

where $n = \sum_a \#_a$ and ρ is a distribution over A such that the measure of each $a \in A$ is proportional to $\#_a$.

The soundness proof amounts to analysing the properties of functions satisfying, in expectation, $f(x_1)f(x_2)\cdots f(x_r) \approx 0$: note that the equality holds for any hyperedge $(x_1, \dots, x_r)$ whenever f is an indicator function of any fixed colour in a NAE_r-colouring. The invariance principle for multislices by [23] argues that these lifting operators preserve products, up to a negligible error. To be precise, here and later we will denote negligible errors by $o(1)$ – a real number that can be pushed arbitrarily close to 0 by taking n large enough (the semantics of n will always be clear from the context).

▶ **Theorem 31** (Corollary of [23, Theorem 1.11]). *Let $\overline{\#}^{(1)}, \ldots, \overline{\#}^{(r)} \in \mathbb{N}^A$ such that $\sum_a \#_a^{(i)} = n$ for each $i \in [r]$. Let Ω be a distribution over A^r; denote $\mu := \mu_n^\Omega$. If the support of Ω is a symmetric BKLM-connected relation, then any $\left\{ f_i : \mathcal{S}\overline{\#}^{(i)} \to [0,1] \right\}_{i=1}^r$ satisfy*

$$\left| \mathop{\mathbb{E}}_{(\overline{x}^{(1)},\ldots,\overline{x}^{(r)})\sim\mu} \left[\prod_{i=1}^r f_i(\overline{x}^{(i)}) \right] - \mathop{\mathbb{E}}_{(\overline{x}^{(1)},\ldots,\overline{x}^{(r)})\sim\Omega^n} \left[\prod_{i=1}^r T_i f_i(\overline{x}^{(i)}) \right] \right| = o(1)$$

where T_i is the lifting operator for $\overline{\#}^{(i)}$.

Roughly speaking, if f is a NAE_r-colouring of the hypergraph, then Tf may be thought of as an *almost* NAE_r-colouring of a hypergraph that we would obtain by applying the full Long Code transformation to the Label Cover instance.

Mossel [55] proved the powerful noise inequalities, from which we conclude these *almost* NAE_r-colourings necessarily possess a coordinate with significant *low-degree influence.* For a function $g\colon A^n \to \mathbb{R}$ and a distribution ρ over A, the degree-k influence of $j \in [n]$, denoted by $\mathbf{Inf}_j^{\le k}[g;\rho]$, is a number in $[0,1]$ which (intuitively) measures how important the j-th input is to the output of g. The precise definition of this notion will not be important to us. We state a specialization of Mossel's result, tailored to our purposes:

▶ **Theorem 32** (Corollary of [55, Proposition 6.4]). *Let Ω be a distribution over A^r whose support is a symmetric BKLM-connected relation of arity $r \ge 2$ and size ≥ 2.*

For all $\delta > 0$ there exist $\tau, k > 0$ such that if $g : A^n \to [0,1]$ with $\mathbb{E}_{\overline{x}\sim\Omega_i^n}[g(\overline{x})] \ge \delta$ for each $i \in [r]$, and

$$\mathop{\mathbb{E}}_{(\overline{x}^{(1)},\ldots,\overline{x}^{(r)})\sim\Omega^n} \left[\prod_{i=1}^r g(\overline{x}^{(i)}) \right] = o(1)$$

then

$$\max_{i\in[r]} \max_{j\in[n]} \mathbf{Inf}_j^{\le k}[g;\Omega_i] > \tau.$$

For completeness, its proof is provided in the Appendix.

The coordinates with significant low-degree influence give rise to a labelling of the original Label Cover instance via the following result. We will say that $\beta > 0$ is $\overline{\#}$*-valid* if $\beta \cdot \sum_a \#_a \in \mathbb{N}$.

▶ **Lemma 33** (See [23, Section 6.4]). *Let $\tau, k, \alpha > 0$ and $\overline{\#} \in \mathbb{N}^A$ be α-balanced with $\sum_a \#_a$ sufficiently large. For any sufficiently small $\overline{\#}$-valid $\beta > 0$, we can assign to every function $g : \mathcal{S}\overline{\#} \to [0,1]$ a coordinate $d_g \in [n]$ in such a way that any $f : \mathcal{S}2\overline{\#} \to [0,1]$ satisfies*

$$\max_{j\in[2n]} \mathbf{Inf}_j^{\le k}[Tf;\rho] > \tau \implies \mathop{\Pr}_{\text{2-to-1 } \pi} \left[\pi(j) = d_{f^\pi} \right] \ge C = C(\alpha,\beta,\tau,k)$$

where T and ρ are the lifting operator and the distribution associated with $\overline{\#}$.

Roughly speaking, the conclusion says that, if one assigns j to the vertex u, and $d_{f^{\pi_{uv}}}$ to each neighbour v of u, then a constant fraction of the edges incident to u will be satisfied.

We provide a concise summary of the soundness proof below.

▷ Claim 34 (Soundness). If $H \to \mathsf{NAE}_r^{(d)}$, then there exists a labelling of Φ (Label Cover instance) with a value $> \varepsilon$.

Proof. Since $H \to \mathsf{NAE}_r^{(d)}$, there is a subset of $V(H)$ of fractional size at least $1/d$ that does not induce any hyperedge in H; let $f\colon V(H) \to \{0,1\}$ be its indicator function. In other words $\mathbb{E}[f] \geq 1/d$, and for every hyperedge $(\overline{y}^{(1)}, \dots, \overline{y}^{(r)})$ we have $f(\overline{y}^{(1)}) \cdot \ldots \cdot f(\overline{y}^{(r)}) = 0$.

For each $v \in V$, let $f_v\colon \mathcal{S}\overline{\#} \to \{0,1\}$ be f restricted to the domain $\left\{(v, \overline{y}) \mid \overline{y} \in \mathcal{S}\overline{\#}\right\}$. For each $u \in U$, the equality constraints (the equivalence $\sim$) give rise to a function $f_u\colon \mathcal{S}2\overline{\#} \to \{0,1\}$ such that $f_u \xrightarrow{\pi_{uv}} f_v$ for any neighbour v of u. By the regularity of $(U \cup V, E)$ and richness, we have

$$\mathop{\mathbb{E}}_{u \in U} \mathbb{E}[f_u] \geq 1/d.$$

Denote $\delta \coloneqq 1/(2d)$. By an averaging argument, we obtain a set $\mathbf{Good} \subseteq U$ of fractional size at least δ such that $\mathbb{E}[f_u] \geq \delta$ for each $u \in \mathbf{Good}$.

Fix $u \in \mathbf{Good}$. The equality constraints ensure that, sampling $(\overline{x}^{(1)}, \dots, \overline{x}^{(r)}) \in \mu_n^{\Omega}$ uniformly at random, we have

$$\mathbb{E}\left[\prod_{i=1}^{r} f_u(\overline{x}^{(i)})\right] = 0.$$

We apply Theorem 31 (the invariance principle) to $\{f_u\}_{i=1}^r$, and obtain

$$\mathop{\mathbb{E}}_{(\overline{x}^{(1)}, \dots, \overline{x}^{(r)}) \sim \Omega^{2n}} \left[\prod_{i=1}^{r} T f_u(\overline{x}^{(i)})\right] = o(1)$$

where T is the lifting operator for $2\overline{\#}$. By the Markov properties of T, we have

$$\mathop{\mathbb{E}}_{\overline{x} \sim \rho^{2n}} [T f_u(\overline{x})] = \mathop{\mathbb{E}}_{\mathcal{S}2\overline{\#}} [f_u] \geq \delta.$$

Applying the noise inequality Theorem 32 to $T f_u$, we obtain $\tau, k > 0$ such that

$$\max_{j \in [2n]} \mathbf{Inf}_j^{\leq k}[T f_u; \rho] > \tau.$$

Assign to u any label j that maximises the expression above; to the vertices $u' \in U \setminus \mathbf{Good}$ assign arbitrary labels.

We now want to appeal to Lemma 33 in order to label the vertices in V. Fix any sufficiently small $\overline{\#}$-valid β. To each $v \in V$ assign the label d_{f_v} obtained from Lemma 33.

We shall argue that the value of this labelling is greater than ε, which would conclude the proof. Fix $u \in \mathbf{Good}$ and its label j. For a random neighbour v of u, by the richness of the instance, the constraint π_{uv} is distributed uniformly among all 2-to-1 maps, and $f_u \xrightarrow{\pi_{uv}} f_v$. By Lemma 33, we get that $\pi_{uv}(j)$ is the label assigned to v with probability at least $C(\alpha, \beta, \tau, k)$. Thus, by the biregularity and the richness of Φ, the probability that our labelling satisfies a uniformly random edge (u, v) is at least

$$\frac{|\mathbf{Good}|}{|U|} \cdot C(\alpha, \beta, \tau, k) \geq \delta \cdot C(\alpha, \beta, \tau, k)$$

so it suffices to pick $\varepsilon < \delta \cdot C(\alpha, \beta, \tau, k)$.

The dependency of various parameters used in the soundness proof on one another is a delicate matter, so we shall elaborate. First of all, note that n can be assumed to be sufficiently large as long as $\varepsilon > 0$ is allowed to be chosen arbitrarily small. Recall that we treat A, r, R, and d as constants. Therefore, α and δ are also constants. Then τ and k are obtained from Theorem 32. We find a suitable sufficiently small $\beta > 0$, and pick $\varepsilon > 0$ small enough so that $\varepsilon \leq \delta \cdot C(\alpha, \beta, \tau, k)$. Finally, for this ε the Rich 2-to-1 Conjecture gives n. It remains to note that we must ensure that $\beta n \in \mathbb{N}$: we can change β slightly to arrange that without breaking other dependencies. To see how, after fixing the upper-bound B on β, take $\varepsilon > 0$ extremely tiny, which forces n to be extremely large. Now note that there exists $\ell \in \mathbb{N}$ such that $B/2 \leq \ell/n \leq B$. It suffices to take $\beta = \ell/n$. ◁

6 Conclusions

In the present paper, we initiated the study of Promise MMSNP problems. For the most fundamental case, forbidding monochromatic cliques, we obtained a complexity dichotomy under the Rich 2-to-1 Conjecture. This class of problems includes the notorious Approximate Graph Colouring, which are not yet known to admit an unconditional hardness proof. Nevertheless, some special cases of our Promise MMSNPs do follow from existing unconditional hardness results: for example, if the forbidden cliques only use 2 colours, the hardness can be derived from the dichotomy for symmetric Boolean PCSPs by [35].

The hardness of the Approximate Graph Colouring was also obtained by [40] from the weaker d-to-1 Conjecture of Khot [44]. Hence the most natural question is whether "richness" is necessary for the Promise MMSNPs studied in this paper. In other words, *do Promise MMSNPs over forbidden monochromatic cliques admit the dichotomy under the 2-to-1 Conjecture?*

These Promise MMSNPs are tightly connected with certain *balanced* hypergraph colouring problems. More concretely, consider $\mathrm{PCSP}(R, \mathsf{NAE}_r)$ where R constitutes the promise colouring. If the promise colouring is *perfectly balanced*, that is, each colour appears the same number of times in every hyperedge, then a proper (NAE_r) colouring can be computed efficiently (see [38]). For multiple special cases, this was proved to be the only tractable case, e.g., [38, 39]. Our result can also be viewed as another special case where a perfect balance of the colours characterises tractability.

There are also some interesting connections with the *topological approach to PCSPs*. Some of the previous results on rainbow-free colouring of hypergraphs are based on topological combinatorics [1]; while our approach is Fourier-analytic, the notion of *reconfigurability* is (perhaps unsurprisingly, given the name) in some sense topological. Indeed, many of the recent applications of topology in (P)CSPs [3, 36, 49, 54] have used *hom-complexes*, a vast generalisation of the reconfiguration graph. (It is worth noting that our Theorem 8 implies all the hardness results that we know of that follow via the topological approach to PCSP, although under the additional hypothesis of the Rich 2-to-1 conjecture.) Such connections between topological and Fourier-analytic approaches to PCSP have been noticed before, e.g. a connection between the work of [49] and Fourier analysis via winding numbers due to Mottet (personal communication). Finding a principled reason why these connections appear may be worth investigating.

Besides the forbidden vertex-coloured patterns, there has been some research on the edge-coloured ones. It would be interesting to obtain similar results for promise problems of forbidden monochromatic cliques of the latter type as well. The forbidden edge-coloured patterns give rise to a larger than MMSNP fragment of Existential Second-Order Logic called *GMSNP*, see [10]. Unlike MMSNP, the computational dichotomy for GMSNP remains unresolved; however, it lies within the scope of the Bodirsky-Pinsker conjecture as well as MMSNP [14]. Recently, the containment problem for two GMSNP formulas was shown to be decidable, and also dependent on the existence of a recolouring [7]. Deciding whether a given graph has an edge 2-colouring that avoids monochromatic triangles is a classic NP-complete problem [37]. It remains NP-hard for an arbitrary (fixed) clique size [27], and for an arbitrary (fixed) number of colours [6]. Nevertheless, its promise extensions present new obstacles, compared to the vertex-coloured cliques. In particular,

- Characterising the existence of a recolouring must depend on the Ramsey numbers, which are very hard to compute.
- The associated finite-domain PCSPs are no longer symmetric: the full symmetric group acting with permutations on the edge set of a clique does not always map the edges of a subclique to the edges of a subclique.

References

1. Per Austrin, Amey Bhangale, and Aditya Potukuchi. Improved Inapproximability of Rainbow Coloring. In Shuchi Chawla, editor, *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1479–1495. SIAM, January 2020. doi:10.1137/1.9781611975994.90.
2. Per Austrin, Johan Håstad, and Björn Martinsson. On the Usefulness of Promises. In Kasper Green Larsen and Barna Saha, editors, *Proceedings of the 2026 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 6332–6378. SIAM, January 2026. doi:10.1137/1.9781611978971.228.
3. Sergey Avvakumov, Marek Filakovský, Jakub Oprsal, Gianluca Tasinato, and Uli Wagner. Hardness of 4-Colouring G-Colourable Graphs. In Michal Koucký and Nikhil Bansal, editors, *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, pages 72–83. ACM, June 2025. doi:10.1145/3717823.3718154.
4. Demian Banakh, Alexey Barsukov, and Tamio-Vesa Nakajima. Towards infinite PCSP: a dichotomy for monochromatic cliques, 2026. arXiv:2605.09815.
5. Demian Banakh and Katzper Michno. Boolean PCSPs through the lens of Fourier Analysis, 2026. arXiv:2604.22742.
6. Alexey Barsukov, Antoine Mottet, and Davide Perinti. Edge-coloring problems with forbidden patterns and planted colors, 2026. doi:10.48550/arXiv.2507.19000.
7. Alexey Barsukov, Michael Pinsker, and Jakub Rydval. Containment for Guarded Monotone Strict NP. In *52nd International Colloquium on Automata, Languages, and Programming (ICALP 2025)*, volume 334 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 140:1–140:20, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2025.140.
8. Libor Barto, Diego Battistelli, and Kevin M. Berg. Symmetric Promise Constraint Satisfaction Problems: Beyond the Boolean Case. In *38th International Symposium on Theoretical Aspects of Computer Science (STACS 2021)*, volume 187 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:16, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.STACS.2021.10.
9. Libor Barto, Jakub Bulín, Andrei Krokhin, and Jakub Opršal. Algebraic Approach to Promise Constraint Satisfaction. *Journal of the ACM*, 68(4), July 2021. doi:10.1145/3457606.
10. Meghyn Bienvenu, Balder Ten Cate, Carsten Lutz, and Frank Wolter. Ontology-Based Data Access: A Study through Disjunctive Datalog, CSP, and MMSNP. *ACM Trans. Database Syst.*, 39(4):33:1–33:44, December 2014. doi:10.1145/2661643.
11. Manuel Bodirsky. *Complexity of infinite-domain constraint satisfaction*, volume 52 of *Lecture Notes in Logic*. Cambridge University Press, Cambridge; Association for Symbolic Logic, Ithaca, NY, 2021. doi:10.1017/9781107337534.
12. Manuel Bodirsky and Martin Grohe. Non-dichotomies in Constraint Satisfaction Complexity. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, *Automata, Languages and Programming*, pages 184–196, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg. doi:10.1007/978-3-540-70583-3_16.

13 Manuel Bodirsky and Jan Kára. The complexity of temporal constraint satisfaction problems. *J. ACM*, 57(2):9:1–9:41, January 2010. doi:10.1145/1667053.1667058.

14 Manuel Bodirsky, Simon Knäuer, and Florian Starke. ASNP: a tame fragment of existential second-order logic. In *Beyond the horizon of computability*, volume 12098 of *Lecture Notes in Comput. Sci.*, pages 149–162. Springer, Cham, 2020. doi:10.1007/978-3-030-51466-2_13.

15 Manuel Bodirsky, Florent Madelaine, and Antoine Mottet. A universal-algebraic proof of the complexity dichotomy for Monotone Monadic SNP. In Anuj Dawar and Erich Grädel, editors, *Proceedings of the 33rd Annual ACM/IEEE Symposium on Logic in Computer Science*, LICS '18, pages 105–114. ACM, July 2018. doi:10.1145/3209108.3209156.

16 Manuel Bodirsky, Florent Madelaine, and Antoine Mottet. A proof of the algebraic tractability conjecture for monotone monadic SNP. *SIAM J. Comput.*, 50(4):1359–1409, 2021. doi:10.1137/19M128466X.

17 Manuel Bodirsky, Barnaby Martin, Michael Pinsker, and András Pongrácz. Constraint satisfaction problems for reducts of homogeneous graphs. *SIAM J. Comput.*, 48(4):1224–1264, 2019. doi:10.1137/16M1082974.

18 Manuel Bodirsky, Michael Pinsker, and András Pongrácz. Projective clone homomorphisms. *J. Symb. Log.*, 86(1):148–161, 2021. doi:10.1017/jsl.2019.23.

19 Joshua Brakensiek and Venkatesan Guruswami. Promise constraint satisfaction: Algebraic structure and a symmetric boolean dichotomy. *SIAM Journal on Computing*, 50(6):1663–1700, 2021. doi:10.1137/19M128212X.

20 Joshua Brakensiek and Venkatesan Guruswami. The quest for strong inapproximability results with perfect completeness. *ACM Trans. Algorithms*, 17(3):27:1–27:35, 2021. doi:10.1145/3459668.

21 Joshua Brakensiek, Venkatesan Guruswami, and Sai Sandeep. Conditional Dichotomy of Boolean Ordered Promise CSPs. *TheoretiCS*, 2, January 2023. doi:10.46298/theoretics.23.2.

22 Alex Brandts and Stanislav Živný. Beyond PCSP(1-in-3, NAE). *Inform. and Comput.*, 289, November 2022. doi:10.1016/j.ic.2022.104954.

23 Mark Braverman, Subhash Khot, Noam Lifshitz, and Dor Minzer. An invariance principle for the multi-slice, with applications. *Adv. Math.*, 480:Paper No. 110460, 80, 2025. doi:10.1016/j.aim.2025.110460.

24 Mark Braverman, Subhash Khot, and Dor Minzer. On Rich 2-to-1 Games. In *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, volume 185 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:20, Dagstuhl, Germany, 2021. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ITCS.2021.27.

25 Andrei Bulatov, Peter Jeavons, and Andrei Krokhin. Classifying the complexity of constraints using finite algebras. *SIAM J. Comput.*, 34(3):720–742, 2005. doi:10.1137/S0097539700376676.

26 Andrei A. Bulatov. A Dichotomy Theorem for Nonuniform CSPs. In Chris Umans, editor, *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 319–330. IEEE, October 2017. doi:10.1109/FOCS.2017.37.

27 Stefan A. Burr. On the computational complexity of Ramsey-type problems. In *Mathematics of Ramsey theory*, volume 5 of *Algorithms Combin.*, pages 46–52. Springer, 1990. doi:10.1007/978-3-642-72905-8_5.

28 Maria Chudnovsky, Linda Cook, James Davies, and Sang-il Oum. Reuniting χ-boundedness with polynomial χ-boundedness. *J. Combin. Theory Ser. B*, 176:30–73, 2026. doi:10.1016/j.jctb.2025.08.002.

29 Lorenzo Ciardo, Marcin Kozik, Andrei Krokhin, Tamio-Vesa Nakajima, and Stanislav Živný. 1-in-3 vs. not-all-equal: Dichotomy of a broken promise. *ACM Trans. Comput. Logic*, 26(2), April 2025. doi:10.1145/3719007.

30 Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional hardness for approximate coloring. *SIAM Journal on Computing*, 39(3):843–873, 2009. doi:10.1137/07068062X.

31 Irit Dinur, Oded Regev, and Clifford Smyth. The hardness of 3-uniform hypergraph coloring. *Combinatorica*, 25(5):519–535, 2005. `doi:10.1007/s00493-005-0032-4`.

32 P. Erdős. Graph theory and probability. *Canadian J. Math.*, 11:34–38, 1959. `doi:10.4153/CJM-1959-003-9`.

33 Ronald Fagin. Generalized first-order spectra and polynomial-time recognizable sets. In *Complexity of computation (Proc. SIAM-AMS Sympos., New York, 1973)*, volume Vol. VII of *SIAM-AMS Proc.*, pages 43–73. Amer. Math. Soc., Providence, RI, 1974.

34 Tomás Feder and Moshe Y. Vardi. The computational structure of monotone monadic SNP and constraint satisfaction: A study through datalog and group theory. *SIAM J. Comput.*, 28(1):57–104, 1998. `doi:10.1137/S0097539794266766`.

35 Miron Ficak, Marcin Kozik, Miroslav Olšák, and Szymon Stankiewicz. Dichotomy for Symmetric Boolean PCSPs. In *46th International Colloquium on Automata, Languages, and Programming (ICALP 2019)*, volume 132 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 57:1–57:12, Dagstuhl, Germany, 2019. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/LIPIcs.ICALP.2019.57`.

36 Marek Filakovský, Tamio-Vesa Nakajima, Jakub Opršal, Gianluca Tasinato, and Uli Wagner. Hardness of linearly ordered 4-colouring of 3-colourable 3-uniform hypergraphs. *ACM Trans. Comput. Theory*, 18(2), May 2026. `doi:10.1145/3779121`.

37 Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. A Series of Books in the Mathematical Sciences. W. H. Freeman, 1979. A guide to the theory of NP-completeness.

38 Venkatesan Guruswami and Euiwoong Lee. Strong inapproximability results on balanced rainbow-colorable hypergraphs. *Combinatorica*, 38(3):547–599, 2018. `doi:10.1007/s00493-016-3383-0`.

39 Venkatesan Guruswami and Rishi Saket. Hardness of Rainbow Coloring Hypergraphs. In Satya V. Lokam and R. Ramanujam, editors, *37th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2017)*, volume 93 of *LIPIcs*, pages 33:1–33:15, Dagstuhl, Germany, 2017. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/LIPIcs.FSTTCS.2017.33`.

40 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 (ICALP 2020)*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 62:1–62:12, Dagstuhl, Germany, 2020. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. `doi:10.4230/LIPIcs.ICALP.2020.62`.

41 Wilfrid Hodges. *A shorter model theory*. Cambridge University Press, 1997.

42 Ravindran Kannan and Achim Bachem. Polynomial algorithms for computing the smith and hermite normal forms of an integer matrix. *SIAM Journal on Computing*, 8(4):499–507, 1979. `doi:10.1137/0208040`.

43 Michał Karpiński and Krzysztof Piecuch. On vertex coloring without monochromatic triangles. In *Computer science—theory and applications*, volume 10846 of *Lecture Notes in Comput. Sci.*, pages 220–231. Springer, Cham, 2018. `doi:10.1007/978-3-319-90530-3_19`.

44 Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thiry-Fourth Annual ACM Symposium on Theory of Computing*, STOC '02, pages 767–775. ACM, 2002. `doi:10.1145/509907.510017`.

45 Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal Inapproximability Results for MAX-CUT and Other 2-Variable CSPs? *SIAM J. Comput.*, 37(1):319–357, January 2007. `doi:10.1137/S0097539705447372`.

46 Subhash Khot and Oded Regev. Vertex cover might be hard to approximate to within $2-\varepsilon$. *J. Comput. Syst. Sci.*, 74(3):335–349, 2008. `doi:10.1016/j.jcss.2007.06.019`.

47 Michael Kompatscher and Trung Van Pham. A complexity dichotomy for poset constraint satisfaction. *J. Appl. Logics*, 5(8):1663–1695, 2018. `arXiv:1603.00082`.

48 Andrei Krokhin and Danny Vagnozzi. Approximating 1-in-3 SAT by linearly ordered hypergraph 3-colouring is NP-hard, 2025. doi:10.48550/arXiv.2508.14606.

49 Andrei A. Krokhin, Jakub Oprsal, Marcin Wrochna, and Stanislav Živný. Topology and Adjunction in Promise Constraint Satisfaction. *SIAM J. Comput.*, 52(1):38–79, 2023. doi:10.1137/20M1378223.

50 Gábor Kun. Constraints, MMSNP and expander relational structures. *Comb.*, 33(3):335–347, 2013. doi:10.1007/s00493-013-2405-4.

51 Richard E. Ladner. On the structure of polynomial time reducibility. *J. ACM*, 22(1):155–171, 1975. doi:10.1145/321864.321877.

52 Florent Madelaine. On the Containment of Forbidden Patterns Problems. In David Cohen, editor, *Principles and Practice of Constraint Programming – CP 2010*, volume 6308 of *Lecture Notes in Computer Science*, pages 345–359. Springer, 2010. doi:10.1007/978-3-642-15396-9_29.

53 Florent Madelaine and Iain A. Stewart. Constraint satisfaction, logic and forbidden patterns. *SIAM J. Comput.*, 37(1):132–163, January 2007. doi:10.1137/050634840.

54 Sebastian Meyer and Jakub Opršal. A topological proof of the Hell-Nešetřil dichotomy. In Yossi Azar and Debmalya Panigrahi, editors, *Proceedings of the 2025 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4507–4519. SIAM, January 2025. doi:10.1137/1.9781611978322.154.

55 Elchanan Mossel. Gaussian bounds for noise correlation of functions. *Geometric and Functional Analysis*, 19(6):1713–1756, March 2010. doi:10.1007/s00039-010-0047-x.

56 Antoine Mottet and Michael Pinsker. Smooth approximations: an algebraic approach to CSPs over finitely bounded homogeneous structures. *J. ACM*, 71(5):36:1–36:47, 2024. doi:10.1145/3689207.

57 Tamio-Vesa Nakajima, Zephyr Verwimp, Marcin Wrochna, and Stanislav Živný. Complexity of Approximate Conflict-Free, Linearly-Ordered, and Nonmonochromatic Hypergraph Colourings. In *52nd International Colloquium on Automata, Languages, and Programming (ICALP 2025)*, volume 334 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 169:1–169:10, Dagstuhl, Germany, 2025. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.ICALP.2025.169.

58 Tamio-Vesa Nakajima and Stanislav Živný. Linearly ordered colourings of hypergraphs. *ACM Trans. Comput. Theory*, 14(3–4), February 2023. doi:10.1145/3570909.

59 Omer Reingold. Undirected connectivity in log-space. *J. ACM*, 55(4), September 2008. doi:10.1145/1391289.1391291.

60 Sai Sandeep. Almost optimal inapproximability of multidimensional packing problems. In *2021 IEEE 62nd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 245–256. IEEE, 2021. doi:10.1109/FOCS52979.2021.00033.

61 Dmitriy Zhuk. A Proof of CSP Dichotomy Conjecture. In Chris Umans, editor, *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 331–342. IEEE, October 2017. doi:10.1109/FOCS.2017.38.

62 Dmitriy Zhuk. A proof of the CSP dichotomy conjecture. *J. ACM*, 67(5):30:1–30:78, 2020. doi:10.1145/3402029.

A Noise inequalities

Correlation plays an important role in our analysis. For a distribution Ω over $A \times B$, it is denoted by $\rho(A, B; \Omega) \in [0, 1]$. We will not define it formally here; the intuition behind $\rho(A, B; \Omega)$ is that it measures the correlation between the first and second coordinates of Ω.

The reason BKLM-connected relations are interesting is due to the following correlation bound that they provide:

▶ **Lemma 35** (Lemma 2.9 in [55]). *Let Ω be a distribution over $A \times B$ whose support forms a connected bipartite graph. Then $\rho(A, B; \Omega) < 1$.*

For a distribution Ω over $A_1 \times \cdots \times A_r$, its correlation vector $\overline{\rho} \in [0,1]^{r-1}$ is defined as

$$\rho_i = \rho(A_1 \times \cdots \times A_i, A_{i+1} \times \cdots \times A_r; \Omega)$$

where Ω is treated as a binary distribution over $(A_1 \times \cdots \times A_i) \times (A_{i+1} \times \cdots \times A_r)$.

Finally, we need Gaussian stability measures. Let $\mathcal{N}$ be the distribution function of a standard Gaussian. We define for $\rho, \mu, \nu \in [0,1]$

$$\Gamma_\rho(\mu, \nu) = \Pr\left[X \leq \mathcal{N}^{-1}(\mu), Y \geq \mathcal{N}^{-1}(1-\nu)\right]$$

where (X, Y) are ρ-correlated[15] Gaussian variables. Note that Γ_ρ is an increasing function in μ and ν. The only fact needed about Γ_ρ is the following.

▶ **Observation 36** (See [30]). *If $\mu > 0$ and $\rho < 1$, then $\Gamma_\rho(\mu, \mu) > 0$.*

The definition of Γ is extended to more than 2 variables in a recursive way. Namely, given $\overline{\rho} \in [0,1]^{r-1}$ and $\overline{\mu} \in [0,1]^r$ for $r \geq 3$ we define

$$\Gamma_{\overline{\rho}}(\mu_1, \ldots, \mu_r) \coloneqq \Gamma_{\rho_1}\left(\mu_1, \Gamma_{\rho_2, \ldots, \rho_{r-1}}(\mu_2, \ldots, \mu_r)\right).$$

▶ **Observation 37.** *If $\mu_1, \ldots, \mu_r > 0$ and $\rho_1, \ldots, \rho_{r-1} < 1$, then $\Gamma_{\overline{\rho}}(\mu_1, \ldots, \mu_r) > 0$.*

Proof. We prove it by induction. For the base case $r = 2$, recall from Observation 36 that

$$\Gamma_\rho(\mu_1, \mu_2) \geq \Gamma_\rho(\mu, \mu) > 0$$

where $\mu \coloneqq \min\{\mu_1, \mu_2\} > 0$. For the inductive step, we have

$$\Gamma_{\overline{\rho}}(\mu_1, \ldots, \mu_r) \coloneqq \Gamma_{\rho_1}\left(\mu_1, \Gamma_{\rho_2, \ldots, \rho_{r-1}}(\mu_2, \ldots, \mu_r)\right) > 0$$

because $\Gamma_{\rho_2, \ldots, \rho_{r-1}}(\mu_2, \ldots, \mu_r) > 0$ by the inductive assumption. ◀

Theorem 32 follows from the powerful noise inequalities discovered by Mossel [55], which we state here for completeness.

▶ **Theorem 38** (Proposition 6.4 in [55]). *Let Ω be a distribution over $\prod_{i=1}^r A_i$ such that*

$$\alpha \coloneqq \min_{\overline{a} \in \mathsf{supp}(\mathbf{P})} \Omega(\overline{a}) \leq \frac{1}{2}$$

and

$$\rho \coloneqq \max_j \rho\left(A_j, \prod_{i \neq j} A_i; \Omega\right) < 1.$$

For every $\varepsilon > 0$ there exists $\tau > 0$ such that if $f_i : A_i^n \to [0,1]$ for $i \in [r]$ satisfy

$$\mathbf{Inf}_j^{\leq \log(1/\tau)/\log(1/\alpha)}[f_i; \Omega_i] \leq \tau$$

for all $i \in [r]$ and $j \in [n]$, then

$$\Gamma_{\overline{\rho}}\left(\mathop{\mathbb{E}}_{\Omega_1^n}[f_1], \ldots, \mathop{\mathbb{E}}_{\Omega_r^n}[f_r]\right) - \varepsilon \leq \mathop{\mathbb{E}}_{(\overline{x}^{(1)}, \ldots, \overline{x}^{(r)}) \sim \Omega^n}\left[\prod_{i=1}^r f_i(\overline{x}^{(i)})\right]$$

where Ω_i denotes the marginal distribution of Ω to i.

---

[15] The exact definition of ρ-correlated variables is not important here; we refer the curious reader to [55].

We are in a position to prove Theorem 32. We state it again, for the reader's convenience.

▶ **Theorem 32** (Corollary of [55, Proposition 6.4]). *Let Ω be a distribution over A^r whose support is a symmetric BKLM-connected relation of arity $r \geq 2$ and size ≥ 2.*

For all $\delta > 0$ there exist $\tau, k > 0$ such that if $g : A^n \to [0,1]$ with $\mathbb{E}_{\overline{x}\sim\Omega_i^n}[g(\overline{x})] \geq \delta$ for each $i \in [r]$, and

$$\mathop{\mathbb{E}}_{(\overline{x}^{(1)},\ldots,\overline{x}^{(r)})\sim\Omega^n}\left[\prod_{i=1}^{r} g(\overline{x}^{(i)})\right] = o(1)$$

then

$$\max_{i\in[r]}\max_{j\in[n]} \mathbf{Inf}_j^{\leq k}[g;\Omega_i] > \tau.$$

Proof of Theorem 32. We apply Theorem 38 with $A_1 = \cdots = A_r = A$ and $f_1 = \cdots = f_r = g$. Clearly α, the minimum probability of any atom in Ω, satisfies $\alpha \leq 1/2$. Since the support of Ω is BKLM-connected and symmetric, we deduce from Lemma 35 that the correlation

$$\rho\left(A_j, \prod_{i\neq j} A_i; \Omega\right) = \rho < 1$$

for any $j \in [r]$. Hence the assumptions of Theorem 38 are satisfied.

We will argue, by contraposition, that it is possible to find $\varepsilon > 0$ small enough so that

$$\Gamma_{\overline{\rho}}\left(\mathop{\mathbb{E}}_{\Omega_1}[g],\ldots,\mathop{\mathbb{E}}_{\Omega_r}[g]\right) - \varepsilon > \mathop{\mathbb{E}}_{(\overline{x}^{(1)},\ldots,\overline{x}^{(r)})\sim\Omega^n}\left[\prod_{i=1}^{r} g(\overline{x}_i)\right] = o(1).$$

Then Theorem 38 gives us $\tau > 0$, we define $k = \log(1/\tau)/\log(1/\alpha)$, and that would conclude the proof.

Since the support of Ω is BKLM-connected, we know from Lemma 35 that $\overline{\rho} \in [0,1)^{r-1}$. It suffices to note that $\Gamma_{\overline{\rho}}(\mathbb{E}_1[g],\ldots,\mathbb{E}_r[g]) > 0$ by Observation 37. ◀