Promise Constraint Satisfaction:
Algebraic Structure and a Symmetric Boolean Dichotomy*

Joshua Brakensiek[†]   Venkatesan Guruswami[‡]

Abstract

A classic result due to Schaefer (1978) classifies all constraint satisfaction problems (CSPs) over the Boolean domain as being either in P or NP-hard. This paper considers a promise-problem variant of CSPs called PCSPs. A PCSP over a finite set of pairs of constraints Γ consists of a pair (Ψ_P, Ψ_Q) of CSPs with the same set of variables such that for every $(P, Q) \in \Gamma$, $P(x_{i_1}, \ldots, x_{i_k})$ is a clause of Ψ_P if and only if $Q(x_{i_1}, \ldots, x_{i_k})$ is a clause of Ψ_Q. The promise problem $\mathrm{PCSP}(\Gamma)$ is to distinguish, given (Ψ_P, Ψ_Q), between the cases Ψ_P is satisfiable and Ψ_Q is unsatisfiable. Many problems such as approximate graph and hypergraph coloring as well as the $(2 + \epsilon)$-SAT problem due to Austrin, Guruswami, and Håstad [1] can be placed in this framework.

This paper is motivated by the pursuit of understanding the computational complexity of Boolean promise CSPs, determining for which Γ the associated PCSP is polynomial-time tractable or NP-hard. As our main result, we show that $\mathrm{PCSP}(\Gamma)$ exhibits a dichotomy (it is either polynomial-time tractable or NP-hard) when the relations in Γ are symmetric and allow for negations of variables. In particular, we show that every such polynomial-time tractable Γ can be solved via either Gaussian elimination over $\mathbb{F}_2$ or a linear programming relaxation. We achieve our dichotomy theorem by extending the (weak) polymorphism framework of Austrin, Guruswami, and Håstad which itself is a generalization of the algebraic approach used by polymorphisms to study CSPs. In both the algorithm and hardness portions of our proof, we incorporate new ideas and techniques not utilized in the CSP case.

Contents

1 Introduction

A constraint satisfaction problem (CSP) over domain D is specified by a finite collection Λ of relations over D, and is denoted as $\mathrm{CSP}(\Lambda)$. An instance of $\mathrm{CSP}(\Lambda)$ consists of a set of variables V and a collection of constraints $\{(\tau, P)\}$ where $P \in \Lambda$ and τ is a tuple of k variables where k is the arity of P (i.e., $P \subseteq D^k$). The goal is find an assignment $\sigma : V \to D$ that satisfies all constraints, i.e., $(\sigma(\tau_1), \ldots, \sigma(\tau_k)) \in P$ for each constraint (τ, P). In the optimization version, we seek an assignment that maximizes the number of satisfied constraints.

Constraint satisfaction problems form a rich class of problems, and have played a crucial role in the development of computational complexity theory, starting from the NP-completeness of 3SAT to the PCP theorem to the Unique Games Conjecture, all of which study the intractability of a certain CSP. Despite the large variety of problems that can formulated as a CSP, it is remarkable that CSPs are a class whose computational complexity one can dream of understanding completely, for either the decision or optimization version (including approximability in the latter case). For Boolean CSPs (those over domain $D = \{0, 1\}$), Schaefer [51] proved a dichotomy theorem showing that such CSP is either polynomial time solvable or NP-complete. Further, he gave a characterization of the tractable cases – a Boolean $\mathrm{CSP}(\Lambda)$ is in P in precisely six cases, when every constraint in Λ is (i) satisfied by all 0s, (ii) satisfied by all 1s, (iii) a conjunction of 2CNF clauses, (iv) a conjunction of Horn SAT clauses, (v) a conjunction of dual Horn SAT clauses, and finally (vi) every constraint in Λ is a conjunction of affine constraints over $\mathbb{F}_2$. The Feder-Vardi conjecture [30] states that such a complexity dichotomy holds for every $\mathrm{CSP}(\Lambda)$ over arbitrary finite domains. Besides the Boolean domain, before 2017 it had been proved for a few other cases, including CSPs over a domain of size 3 [20] and conservative CSPs (which contain all unary relations) [22, 21]. The conjecture was recently resolved independently by Bulatov [18] and Zhuk [54].

For the (exact) optimization version, a complete dichotomy theorem was established in [52] showing that for every collection of relations Λ, the associated optimization problem is tractable if and only if a certain basic linear programming relaxation solves it, and it is NP-complete otherwise. The result in fact holds for a generalization of Max CSP called *valued* CSP, where each constraint has a finite weight associated with it, and the goal is to find a minimum value solution. When infinite weights are also allowed (so some constraints have to be satisfied), it was shown that, surprisingly, the dichotomy for ordinary CSPs implies a dichotomy for this more general setting as well [41]. For approximate optimization, a line of work exploring the consequences of Khot's Unique Games Conjecture (UGC) [40] culminated in the striking result [50] (see also [17]) that for every CSP, there is a canonical semidefinite programming relaxation which delivers the optimal worst-case approximation ratio, assuming the UGC.

In this work we are interested in a potential complexity dichotomy for *promise* constraint satisfaction problems (PCSPs). A promise constraint satisfaction problem $\mathrm{PCSP}(\Gamma)$ is specified by a finite collection $\Gamma = \{(P_i, Q_i)\}_i$ of *pairs* of relations with each $P_i \subseteq Q_i$. Let $\Lambda = \{P_i\}_i$ and $\Lambda' = \{Q_i\}_i$, Suppose we are given a satisfiable instance of $\mathrm{CSP}(\Lambda)$ — while finding a satisfying assignment might be NP-hard, can we in polynomial time find a satisfying assignment when the input is treated as an instance of $\mathrm{CSP}(\Lambda')$ (in the obvious way, by replacing each constraint P_i by the corresponding Q_i)? The decision version $\mathrm{PCSP}(\Gamma)$ is the promise problem where given an instance, we need to output Yes on instances that are satisfiable as a $\mathrm{CSP}(\Lambda)$ instance, and output No on instances that are unsatisfiable even as a $\mathrm{CSP}(\Lambda')$ instance.[1] The following challenge drives this work:

Question 1.1. *For which Γ is $\mathrm{PCSP}(\Gamma)$ polynomial-time tractable? For which Γ is $\mathrm{PCSP}(\Gamma)$ NP-hard? Must every Γ fall into one of these two categories?*

---

[1]PCSPs are not to be confused with a promise problem of a similar name studied by Ham and Jackson [34, 33] which deal with a different kind of promise that certain partial assignments to a CSP are 'extendable' to a full solution.

The condition $P_i \subseteq Q_i$ guarantees that the satisfiability of a CSP with clauses in the P_i's implies the satisfiability of the CSP with the P_i's replaced by the corresponding Q_i's. More generally one may consider two not necessarily distinct domains D_1 and D_2, with $P_i \subseteq D_1^k$ and $Q_i \subseteq D_2^k$ and a map $\sigma : D_1 \to D_2$ such that $\sigma(P_i) \subseteq Q_i$ (where $\sigma(P_i) = \{(\sigma(a_1), \ldots, \sigma(a_k)) \mid (a_1, \ldots, a_k) \in P_i\}$. Note that σ is not necessarily injective. We omit the details of this more general presentation in this work (except briefly in the promise graph homomorphism example below).

To demonstrate the depth and far-reaching nature of the above question, we provide some interesting examples which fall under this Promise CSP framework. (Throughout the paper, we will use $[n] = \{1, \ldots, n\}$, $|x|$ to denote the Hamming weight (the number of 1s) of a Boolean vector $x \in \{0, 1\}$, and $e_i \in \{0, 1\}^n$ to denote the unique vector such that $(e_i)_j = 1$ if and only if $j = i$.)

CSPs. Consider a PCSP $\Gamma = \{(P_i, Q_i) : i \in [r]\}$ such that $P_i = Q_i$ for all i. Then $\mathrm{PCSP}(\Gamma)$ is equivalent to the CSP decision problem $\mathrm{CSP}(\Gamma)$. Thus the above question in full generality subsumes the CSP dichotomy theorem as a special case.

Approximate graph coloring. Let $3 \leq c \leq t$ be positive integers, and consider the relations $P = \{(a, b) \in [c]^2 : a \neq b\}$ and $Q = \{(a, b) \in [t]^2 : a \neq b\}$ with $D = \{1, \ldots, t\}$.[2] Then (P, Q)-PCSP is an instance of the approximate graph coloring problem in which one needs to distinguish if the chromatic number of a graph is at most c or at least $t + 1$. The complexity of $\mathrm{PCSP}(P, Q)$ is a notorious open problem; this problem is strongly believed to be NP-hard for all $3 \leq c \leq t$, but the best NP-hardness in various regimes [39, 32, 36, 9] fall woefully short of establishing hardness for all c and t, especially when c is small.

Hypergraph coloring. Generalizations of the coloring problem to the setting of hypergraphs also fall under this framework. The hardness of telling if a 3-uniform hypergraph is 2-colorable or not even t-colorable [29] (for any fixed t) is captured by $\mathrm{PCSP}(P, Q)$ where $P = \{1, 2\}^3 \setminus \{(1, 1, 1), (2, 2, 2)\}$ and $Q = [t]^3 \setminus \{(j, j, j) \mid j \in [t]\}$.

$(2 + \epsilon)$-SAT. This problem studied by [1] corresponds to $\Gamma = \{(P_1, Q_1), (P_2, Q_2)\}$ where $(P_1, Q_1) = (\{x \in \{0, 1\}^{2k+1}, |x| \geq k\}, \{0, 1\}^{2k+1} \setminus \{(0, \ldots, 0)\})$ (where $|x|$ is the Hamming weight of x) and $(P_2, Q_2) = (\{(0, 1), (1, 0)\}, \{(0, 1), (1, 0)\})$. The purpose of (P_2, Q_2) is so that we can refer to some variables as negations of others. This specific $\mathrm{PCSP}(\Gamma)$ was shown to be NP-hard. On the other hand, if we replace P_1 with $\{x \in \{0, 1\}^{2k+1}, |x| \geq k + 1\}$, then $\mathrm{PCSP}(\Gamma)$ has a polynomial-time algorithm.

Hypergraph discrepancy. Let $\Gamma = \{(P, Q)\}$ where $P = \{x \in \{0, 1\}^{2k+1}, |x| \in \{k, k + 1\}\}$ and $Q = \{0, 1\}^{2k+1} \setminus \{0^{2k+1}, 1^{2k+1}\}$. Then $\mathrm{PCSP}(\Gamma)$ was shown to be hard in [1], which means that weak 2-coloring of hypergraphs with minimum discrepancy is hard. On the other hand, if the arity is even and P contains strings of equal number of 0s and 1s, then $\mathrm{PCSP}(\Gamma)$ is tractable.

Promise graph homomorphism. Let G and H be fixed directed graphs for which there is a homomorphism $\phi : G \to H$; that is, for all $(u, v) \in E(G)$, $(\phi(u), \phi(v)) \in E(H)$. Consider the following promise problem: given an input directed graph X, determine whether there exists a homomorphism $X \to G$ or there exists no homomorphism $X \to H$. We coin this question as the "promise digraph homomorphism problem." This question can be encoded as a PCSP as follows. Let $P = E(G)$, the ordered pairs of vertices forming directed edges, and $Q = E(H)$. Any instance $\mathrm{PCSP}(P, Q)$ corresponds to the directed graph whose vertices are the variables and whose directed edges are the clauses.

In the non-promise case (when $G = H$), the problem is known as the "H-coloring" or the "digraph homomorphism" problem (e.g., [35, 30]). Resolving the full CSP dichotomy theorem is equivalent to resolving the special case of digraph homomorphism problems [30]. Special cases which had been resolved before the full dichotomy include the case H is undirected [35] and when H has no sources or sinks [5].

---

[2]Instead of having P ignore a portion of the domain, we could present these more naturally in the homomorphism framework mentioned previously.

Extending the polynomial-time equivalence of CSPs and digraph homomorphism problems, we show in Section 6.5 that every PCSP is polynomial-time equivalent to a promise digraph homomorphism problem. It follows that a dichotomy theorem for promise digraph homomorphism problems is equivalent to a dichotomy theorem for all PCSPs.

Note that even the undirected case of promise digraph homomorphism problem is a substantial generalization of the approximate graph coloring problem, which corresponds to $G = K_c$ and $H = K_t$ being cliques. We conjecture the following very general hardness result for promise graph homomorphism:

Conjecture 1.2. *Let G, H be undirected non-bipartite graphs with a homomorphism from G to H. Then the promise digraph homomorphism problem associated with G and H is NP-hard.*[3]

Since originally posing this conjecture, a number of special cases have been solved, see Section 1.3.

Given that PCSPs generalize CSPs and a dichotomy theorem for CSPs over arbitrary domains is itself a long elusive challenge, in this work we focus on Question 1.1 for relations over the Boolean domain. Even in this restricted setting, Boolean promise CSPs have a richer structure from that of Boolean CSPs (see Section 1.2), rendering proving a generalization of Schaefer's dichotomy quite difficult, if it is even true. In this work, we build the groundwork for the complexity classification of promise PCPs, and prove a dichotomy for the case of *symmetric* Boolean promise CSPs allowing negations (Theorem 1.3 below). Negations can be enforced if $(P, Q) \in \Gamma$ where $P = Q = \{(0, 1), (1, 0)\}$; we say such a Γ *allows negations* or is *folded*. A collection of relation pairs $\Gamma = \{(P_i, Q_i): i \in [r]\}$ is symmetric if each P_i and Q_i is a symmetric relation. A relation P is symmetric if $(a_1, a_2, \ldots, a_l) \in P$ iff $(a_{\pi(1)}, \ldots, a_{\pi(l)}) \in P$ for every permutation $\pi \in S_l$. Note that a symmetric relation $P \subseteq \{0, 1\}^l$ can be specified by a set $S \subseteq \{0, 1, \ldots, l\}$ such that $P = \{x \in \{0, 1\}^l \mid |x| \in S\}$.

Theorem 1.3 (Main). *Let Γ be a symmetric collection of Boolean relation pairs that allows negations. Then $\mathrm{PCSP}(\Gamma)$ is either in P or NP-hard.*

While the symmetry requirement is a significant restriction, it is a natural subclass that still captures several fundamental problems, such as k-SAT, Not-All-Equal-k-SAT, t-out-of-k-SAT, Hypergraph Coloring, Bipartiteness, Discrepancy minimization, etc. In all these cases, whether a constraint is satisfied only depends on the number of variables set to 1 (negations can be enforced via the symmetric relation $\{(0, 1), (1, 0)\}$). Note that Horn SAT is an example of a CSP that is *not* symmetric.

We establish Theorem 1.3 via a characterization of all the tractable cases, and showing that everything else is NP-hard. To describe our results in greater detail, and to highlight the challenges faced in extending Schaefer's theorem to the land of promise CSPs, we now turn to the algebraic approach to study $\mathrm{CSP}(\Lambda)$ via polymorphisms of the underlying relations.

Polymorphisms are operations that preserve membership in a relation. Formally, $f : \{0, 1\}^m \to \{0, 1\}$ is a polymorphism of $P \subseteq \{0, 1\}^k$, denoted $f \in \mathrm{Pol}(P)$, if for all $(a_1^{(i)}, \ldots, a_k^{(i)}) \in P$, $i = 1, 2, \ldots, m$, $\left(f(a_1^{(1)}, a_1^{(2)}, \ldots, a_1^{(m)}), \cdots, f(a_k^{(1)}, a_k^{(2)}, \ldots, a_k^{(m)})\right) \in P$. For a collection Λ of relations, $\mathrm{Pol}(\Lambda) = \cap_{P \in \Lambda} \mathrm{Pol}(P)$. Remarkably, the complexity of $\mathrm{CSP}(\Lambda)$ is completely captured by $\mathrm{Pol}(\Lambda)$. The *Galois correspondence* [37] states that if $\mathrm{Pol}(\Lambda') \subseteq \mathrm{Pol}(\Lambda)$ then any relation in Λ can be build from Λ' using primitive positive reductions (see Section 6.1), from which it follows that $\mathrm{CSP}(\Lambda)$ reduces to $\mathrm{CSP}(\Lambda')$. Note that all dictator functions (called projections in the decision-CSP literature), $f(x_1, \ldots, x_m) = x_j$ for some j, always belong to $\mathrm{Pol}(\Lambda)$.

The algebraic dichotomy theorem, proved by Bulatov [18] and Zhuk [54], states that $\mathrm{CSP}(\Lambda)$ is in P iff $\mathrm{Pol}(\Lambda)$ contains a certain kind of polymorphism known as a weak-near unanimity operator (see [3]). The

---

[3]If either G or H is bipartite, the problem is easy via graph 2-coloring.

NP-hardness half of this theorem was resolved a decade before the full dichotomy (see [3] and [23, 24, 46]): if $\mathrm{Pol}(\Lambda)$ fails to have a weak near-unanimity operator,[4] then there exists an NP-hard reduction.

The algebraic formulation of Schaefer's dichotomy theorem states that a Boolean $\mathrm{CSP}(\Lambda)$ is tractable if $\mathrm{Pol}(\Lambda)$ contains one of the six functions: constant 0, constant 1, Majority on 3 variables, Boolean AND, Boolean OR, or parity of 3 variables;[5] otherwise $\mathrm{CSP}(\Lambda)$ is NP-complete. We refer the reader to the article by Chen [27] for an excellent contemporary treatment of Schaefer's theorem for Boolean domains in the language of polymorphisms. For larger domains, there has been a lot exciting recent progress, including the resolution of the bounded width conjecture by Barto and Kozik [3, 4] proving a precise characterization of when a natural local consistency algorithm works for $\mathrm{CSP}(\Lambda)$ in terms of the structure of $\mathrm{Pol}(\Lambda)$.

Generalizing the situation for CSPs, it is natural to hope the complexity of PCSPs will also have some algebraic structure. Austrin, Guruswami and Håstad [1] pioneered the study of *polymorphisms of PCSPs*.[6] A function $f : \{0,1\}^m \to \{0,1\}$ is polymorphism for a pair of relations (P, Q), denoted $f \in \mathrm{Pol}(P, Q)$, if f maps any m inputs in P to an output in Q. When $P = Q$, this is just the notion of a CSP polymorphism for P.

In Section 6.1, we prove a generalization of the Galois correspondence from CSPs to promise CSPs, establishing that the complexity of a PCSP is captured by its polymorphisms (closely related results, albeit without complexity-theoretic applications in mind, are established by Pippenger [48]). Therefore polymorphisms are the 'right' approach to study the complexity of promise CSPs. When studying promise CSPs under the lens of polymorphisms, however, several challenges surface that didn't exist in the world of CSPs. From an algebraic point of view, polymorphisms are no longer closed under composition (because after one application, we no longer have an assignment in P, but rather for a different relation Q). In universal algebra parlance, polymorphisms of PCSPs do not necessarily form a "clone."[7] The dichotomy theorem for Boolean CSPs can avail of a classification of all Boolean clones which dates back to 1941 [49] (again, see [27] for a crisp presentation). In the world of promise CSPs, polymorphisms belong to a broader class of algebraic structures, and it is a lot more challenging to understand their structure; see Section 1.2.

From a complexity point of view, the distinction between easy and hard is now more nuanced; the existence of any particular polymorphism doesn't itself imply tractability. Indeed, for the $(2+\epsilon)$-SAT problem mentioned earlier, majority of small arity is a polymorphism even though the promise CSP is NP-hard. At an intuitive level, we might expect a PCSP to be easy if there are polymorphisms that "genuinely" depend on a lot of variables, and hard if a few variables exert a lot of influence on the function. The precise way to formalize this notion that captures the boundary between tractable and hard is not yet clear. In [1], hardness was shown when the only polymorphisms were juntas, functions which depends on a bounded number of coordinates; in this work we relax this condition to the existence of a small number of coordinates setting all of which to 0 fixes the function.

In addition to establishing the hardness of many natural PCSPs, we also demonstrate the existence of new polynomial-time tractable PCSPs. As an example, consider a hypergraph $H = (V, E)$ such that all of its edges have bounded valence (but not all the valences need to be the same). Furthermore, for each $e = \{v_1, \ldots, v_k\} \in E$, we specify a hitting number $h_e \in \{1, \ldots, k-1\}$. Then, it is polynomial-time tractable to distinguish between the following two cases (1) there exists a two-coloring of the vertices of H such that for all $e \in E$ the number of vertices of the first color is exactly h_e and (2) every two-coloring of the vertices of H leaves a monochromatic hyperedge. Formally, this is a PCSP with relations of the form $P = \{x \in \{0,1\}^k \mid |x| = a\}$ for any choice of $0 < a < k$ and $Q = \{x \in \{0,1\}^k \mid |x| \in \{1, 2, \ldots, k-1\}\}$. In essence, this PCSP is

---

[4]According to [3], a function $f : D^n \to D$, $n \geq 2$, is a weak near-unanimity operator if for all $a, b \in D$, $f(a, \ldots, a) = a$, a property known as idempotence, and $f(b, a, \ldots, a) = f(a, b, a, \ldots, a) = \cdots = f(a, \ldots, a, b)$. A simple example of a weak near-unanimity operator is the Boolean majority function on an odd number of variables.

[5]The first two cases are CSPs satisfied trivially by the all 0s or all 1s assignment; Majority corresponds to 2SAT; AND and OR to Horn SAT and dual Horn SAT; and Parity to linear equations mod 2.

[6]Austrin, Guruswami and Håstad referred to these objects as *weak polymorphisms.*

[7]Recently, the term *clonoid* or *minion* has been used to describe this large class of families.

4

a hypergraph generalization of what makes 2-coloring for graphs efficient. The algorithm for solving this problem is based on linear programming. Unlike other CSPs and PCSPs, the proof of correctness uses the *Alternating-Threshold polymorphism*, a function which takes as input $x_1, \ldots, x_L \in \{0, 1\}$ (L odd) and returns whether $x_1 - x_2 + \cdots - x_{L-1} + x_L$ is positive. In the Boolean setting for CSPs or PCSPs, this is the first non-symmetric polymorphism known to yield a polynomial time algorithm.[8]

We now informally state the main dichotomy (for a formal statement see Theorem 2.16) in two ways. First, we give an explicit characterization in terms of the structure of the PCSP itself. For simplicity, we only state a subset of the main result in this form. In particular, we allow for the "negation of variables," that if x_i can be used in a clause, then $\bar{x}_i$ can also be used and for the "setting of constants," we can force a variable in a clause to be the constant 0 or 1.

Theorem 1.4. *Let $P \subseteq Q \subset \{0, 1\}^k$ be a symmetric pair of relations. Let Γ contain the promise relation (P, Q) as well as allow for negation of variables–$(\{(0, 1), (1, 0)\}, \{(0, 1), (1, 0)\})$–and the setting of constants– $(\{0\}, \{0\})$ and $(\{1\}, \{1\})$–(e.g. $x_i = 0$). Let $S = \{|x| \mid x \in P\}$ and $T = \{|x| \mid x \in Q\}$. Furthermore, assume that $S \cap \{1, \ldots, k - 1\}$ is nonempty. Then, $\mathrm{PCSP}(\Gamma)$ is polynomial-time tractable if*

a) $S \subseteq \{\ell \in [k] \mid \ell \text{ odd}\} \subseteq T$ or $S \subseteq \{\ell \in \{0\} \cup [k] \mid \ell \text{ even}\} \subseteq T$ or

b) $T \supseteq \{0, 1, \ldots, k\} \cap \{2 \min S - k + 1, \ldots, 2 \max S - 1\}$ or

c) $|S| = 1$ and $T \supseteq \{1, \ldots, k - 1\}$.

Otherwise, $\mathrm{PCSP}(\Gamma)$ is NP-hard.

Second, we give a complete and elegant ormulation of the dichotomy in terms of the polymorphisms of the PCSP instead of the PCSP itself.

Theorem 1.5 (Theorem 2.16). *Let Γ be a family of pairs of symmetric relations which allows for negations as well as the setting of constants. Then, $\mathrm{PCSP}(\Gamma)$ is polynomial-time tractable if*

a) The Parity of L variables or the negation of Parity of L variables is a polymorphism of Γ for all odd L or

b) The Majority of L variables or the negation of Majority of L variables is a polymorphism of Γ for all odd L or

c) The Alternating-Threshold of L variables or the negation of Alternating-Threshold of L variables is a polymorphism of Γ for all odd L.

Otherwise, $\mathrm{PCSP}(\Gamma)$ is NP-hard.

Note that we also need to consider the negations of polymorphisms in the full result. To illustrate this, consider $P = \{x \in \{0, 1\}^5 \mid |x| \in \{2\}\}$ and $Q = \{x \in \{0, 1\}^5 \mid |x| \in \{1, 2, 3, 5\}\}$ and allow negations. It is not hard to show that none of Parity, Majority, or Alternating-Threshold on 11 variables is in $\mathrm{Pol}(P, Q)$, yet $\mathrm{PCSP}(\Gamma)$ is polynomial-time tractable. Instead the negation of Parity or 'anti-Parity' on L variables is a polymorphism of $\mathrm{Pol}(P, Q)$ for all odd L. In Section 2.3, we show how to handle this technical issue.

1.1 Proof Overview

The proof of the main theorem consists of three major parts. First, in Section 3 we show that any PCSP which has one of these families of functions as a polymorphism–Parity, Majority, or Alternating-Threshold, or their negations, which we call *non-idempotent polymorphisms*–has a polynomial time algorithm. The algorithms we demonstrate are quite general in that the only assumption we make is the existence of polymorphisms,

---

[8]If we remove the symmetric condition on the relations, it turns our that non-symmetric polymorphisms are the norm, even in the Boolean case, see Section 6.

in particular we do *not* rely on the symmetry assumption. For Parity, we show that the problem can be reduced to an ordinary CSP with Parity as a polymorphism, and thus Schaefer's theorem can be invoked. For Majority and Alternating-Threshold, such a tactic cannot be used. Instead, we show how these problems can be written as linear programming relaxations. Surprisingly, identical algorithms are used in both cases to solve the decision problem. They do diverge, however, if one desires to use the LP relaxation to also find a solution when the PCSP is satisfiable. To deal with non-idempotent polymorphisms, we show in Section 2.3 how these PCSPs with non-idemotent polymorphisms can be reduced in polynomial time to PCSPs with their negations (that is, the "normal" or idempotent polymorphisms), which we already know are polynomial-time solvable.

Second, in Section 4, for every symmetric PCSP with negations that does not have the entirety of any of the mentioned families of polymorphisms, we show that its polymorphisms are 'lopsided.' More precisely, we show that there exists a constant C, only dependent on the type of the PCSP, such that for all polymorphisms of the PCSP, there are C coordinates such that setting all C of those coordinates to the same value fixes the value of the polymorphism. We say that such polymorphisms are "C-fixing." The general philosophy of the argument is as follows. First, since our Γ fails to have Alternating-Threshold on L variables for some odd L as a polymorphism, there is some $(P, Q) \in \Gamma$ responsible for this exclusion. Using a nuanced combinatorial argument, we attempt to classify the polymorphisms of (P, Q) given that P and Q are symmetric. To simplify the proof, we first show that we may transform (P, Q) into a canonical (P', Q') without losing any polymorphisms (see Lemma 4.5). From this, we show that all polymorphisms f of Γ have the property that either $f(e_i)$ differs from $f(0, \ldots, 0)$ for a bounded number of e_i or a substantial portion of f is structured like the Parity polymorphism. Since we assume that Parity of L' variables is not a polymorphism of Γ for some odd L', we can show that the latter situation is impossible. Using another $(P'', Q'') \in \Gamma$ which fails to have Majority as a polymorphism, and after simplifying (P'', Q'') to a canonical form, we can use arguments inspired from [1] to obtain additional structural information which yields that all polymorphisms are C-fixing. We crucially exploit that (P, Q) and (P'', Q'') are symmetric to get these structural properties, but do not assume anything about the other clauses of Γ.

Finally, since we have pinned down the nature of the polymorphisms in these believed-to-be-hard PCSPs, in Section 5, we prove the NP-hardness of these PCSPs. We prove this by reducing from Label Cover, a well-known problem to reduce from for hardness of approximation proofs. This part of the proof is based on an argument of [1], but we greatly simplify how projection constraints are handled. With this hardness result established, the main theorem is proved.

1.2   Must there be a Dichotomy?

Extending this dichotomy from the symmetric case to the full Boolean case presents significant challenges, some of which perhaps suggest that a dichotomy does not exist. Compared to when this manuscript was first written, the possibility of a dichotomy seems more likely, but it has become more clear what the hurdles will be establishing such a result.

One challenge is the large variety of families of polymorphisms to consider. In Section 6, we provide necessary and sufficient conditions, adapting a result of Pippenger [48], for a family of functions $\mathcal{F}$ to satisfy $\mathcal{F} = \mathrm{Pol}(\Gamma)$ for some PCSP Γ (not necessarily Boolean). These conditions, known as *projection-closure* and *finitization*, are extremely flexible, allowing for an extremely rich variety of polymorphisms. Note that these results liberate us from ever thinking about Γ, and instead we can think entirely in terms of establishing the easiness/hardness of projection-closed, finitized families of functions. There is, however, a caveat: there is a huge amount of freedom in finitizable, projection-closed families of functions!

However, a fair counterargument is that the polymorphisms themselves are too fine-grained of a perspective even for the CSP dichotomy. Rather, the *identities* which the polymorphisms satisfy (e.g.,

whether the polymorphism is symmetric in its coordinates, etc.) are more important (e.g., [6]). In particular, the recent work of Barto, Bulin, Krokhin, and Opršal [2] has shown it suffices to classify PCSPs by their "height-1 identities."

That said, even in the Boolean setting, both the algorithmic and hardness sides of a dichotomy resolution seem more daunting. For an algorithmic example, let p be any prime number and let $S \subset \{0, \ldots, p-1\}$ be a non-empty strict subset. Then, for each $L \in \mathbb{N}$ define $f^{(L)} : \{0,1\}^L \to \{0,1\}$ such that $f^{(L)}(x) = 0$ if $|x|$ mod $p \in S$, and $f^{(L)}(x) = 1$ if $|x|$ mod $p \notin S$. Therefore, if a PCSP Γ has $f^{(L)} \in \text{Pol}(\Gamma)$ for infinitely many L, then $\text{PCSP}(\Gamma)$ can be efficiently solved using Gaussian elimination over $\mathbb{F}_p$! Thus, algorithms for Boolean PCSP seem to require algorithms for CSPs over arbitrarily large domains. The subsequent work of [2] shows that in fact CSP algorithms over infinite domains can be necessary! This alone, combined with the fact that dichotomy theorems for infinite-domain CSPs do not exist in general (e.g., [7]), should give some pause.

The hardness side of the dichotomy is also vastly more complex for PCSPs than for CSPs. The hardness side of the CSP dichotomy theorem was resolved relatively early (e.g., [23]) with a gadget reduction from 3-SAT. On the other hand, even for Boolean PCSPs, the use of a PCP-like theorem for hardness seems necessary. For instance, the hardness of $(2 + \epsilon)$-SAT in [1] relied on the PCP theorem.[9] Some other hardness results for PCSPs are based on very strong, conjectural versions of the PCP theorem. For instance, strong hardness results for approximate graph coloring rely on variants of the Unique Game Conjecture [28], although new paths to circumventing this assumption have recently appeared [2, 42, 53]. A recent work showed that Boolean ordered PCSPs, whose polymorphisms must be monotone functions, exhibit a dichotomy [13], but the hardness side relied on the "Rich 2-to-1 conjecture" due to [16], which is a significant strenghtening of the 2-to-1 conjecture of Khot [40].

This leads the authors to conjecture that if a dichotomy of PCSPs does hold, then the algorithmic side will be resolved (long) before the hardness side. In fact, follow-up work [14] has given a polynomial time (decision) algorithm for any PCSP with infinitely many symmetric polymorphisms; and it is not currently known if there is a tractable Boolean PCSP which is *not* solved by this algorithm.[10]

1.3 Subsequent Work

Since the original version of this paper [10], there have been numerous follow-up works. Most notably, [26] developed a universal algebraic theory of PCSPs. In particular, generalizing the Galois correspondence of this paper, they show that the *identities* which the polymorphisms satisfy suffice to determine the complexity. As a result, they were able to show that 3 versus 5 approximate graph coloring is NP-hard. A subsequent revision [2] contains additional results, including a result that the polynomial-time complexity of 1-in-3-SAT versus NAE-3-SAT cannot be attributed to any finite-domain CSP.

In this work, we assume that negations of variables are allowed in our dichotomy, a recent work [31] removes this assumption by relaxing the notion of C-fixing.

On the complexity of $\text{PCSP}(G, H)$, where $G \to H$ are non-bipartite graphs, [42] showed that this problem is NP-hard whenever $H = K_3$. Interestingly, there hardness proof uses topologically inspired techniques which have been further developed in the recent work of [53]. See also [43].

On the algorithmic side, follow-up works of the authors [11, 12] have fleshed out many of the algorithmic techniques applied in this paper. In particular, [12] show that any PCSP with infinitely many symmetric polymorphisms has an efficient decision algorithm, although constructing an efficient search algorithm in general is still an open question. The exact class of PCSPs for which this algorithm solves was classified in

---

[9]Recent unpublished work of Barto-Kozik shows a "Baby PCP theorem" which might suffice in some cases.

[10]Jakub Opršal showed this algorithm does not work for general non-Boolean (P)CSPs (see remark in [14]); however he conjectures that a natural extension of the algorithm may work for all CSPs [47].

an expanded version [14].

1.4 Organization

In Section 2, we formally define the notion of a PCSP as well as other tools and terminology which we will need in investigating PCSPs. In Section 3, we prove the algorithmic portion of the main theorem. In Section 4, we give a structural characterization of PCSPs that is used to show hardness. In Section 5, we use the results of Section 4 to complete the NP-hardness results of the main theorem. In Section 6, we prove some more general facts about PCSPs, including that polymorphisms precisely capture the computational complexity of PCSPs as well as classify the possible families of polymorphisms of PCSPs.

2 Promise Constraint Satisfaction Problems

We develop a theory of the complexity of promise constraint satisfaction problems (PCSPs) analogous to that of 'ordinary' CSPs such as found in [27]. We need to formally define what we mean by a PCSP.

Definition 2.1. Let D be a finite domain. A *relation* of arity k is a subset $P \subseteq D^k$. A promise relation is a pair of relations (P, Q) of arity k such that $P \subseteq Q$.

We say that a relation is *Boolean* if $D = \{0, 1\}$ (or more generally $|D| = 2$). For a given relation P, we will refer to it both as a subset of D^k as well as its indicator function $P : D^k \to \{0, 1\}$ ($P(x) = 1$ iff $x \in P$). It should be clear from context which notation for P we are using. If (P, Q) is a promise relation then $P(x) = 1 \implies Q(x) = 1$. When $P = Q$, the promise relation (P, Q) is analogous to the relation P in a CSP. In fact, when it is clear that we are referring to promise relations, we let P denoted the promise relation (P, P).

Definition 2.2. Let $(P, Q) \subseteq D^k \times D^k$ be a promise relation. A (P, Q)-*PCSP* is a pair of formulae (Ψ_P, Ψ_Q), each with m clauses on the variables $x_1, \ldots, x_n$ along with a variable-choice function $\ell : [m] \times [k] \to [n]$, such that $\Psi_P(x_1, \ldots, x_n) = \bigwedge_{i=1}^m P(x_{\ell(i,1)}, x_{\ell(i,2)}, \ldots, x_{\ell(i,k)})$ and $\Psi_Q(x_1, \ldots, x_n) = \bigwedge_{i=1}^m Q(x_{\ell(i,1)}, x_{\ell(i,2)}, \ldots, x_{\ell(i,k)})$. Further, let $\Gamma = \{(P_i, Q_i) \subseteq D^{k_i} \times D^{k_i} : i \in [r]\}$ be a set of promise relations over D of possibly distinct arities. For $i \in [r]$ let (Ψ_{P_i}, Ψ_{Q_i}) be a (P_i, Q_i)-PCSP so that each PCSP is on the same variable set $x_1, \ldots, x_n$. A Γ-*PCSP* is then a pair of formula (Ψ_P, Ψ_Q) such that $\Psi_P(x_1, \ldots, x_n) = \bigwedge_{i=1}^r \Psi_{P_i}(x_1, \ldots, x_n)$ and $\Psi_Q(x_1, \ldots, x_n) = \bigwedge_{i=1}^r \Psi_{Q_i}(x_1, \ldots, x_n)$.

We say that (Ψ_P, Ψ_Q) is *satisfiable* if there exists $(x_1, \ldots, x_n) \in D^n$ such that $\Psi_P(x_1, \ldots, x_n) = 1$. That is, Ψ_P is satisfiable in the usual sense. We say that (Ψ_P, Ψ_Q) is *unsatisfiable* if Ψ_Q is unsatisfiable, for all $(x_1, \ldots, x_n) \in D^n$, $\Psi_Q(x_1, \ldots, x_n) = 0$. Since the clauses involve promise relations, any satisfying assignment to Ψ_P is necessarily a satisfying assignment to Ψ_Q, so no Γ-PCSP can be simultaneously satisfiable and unsatisfiable. Despite that, it is possible for the PCSP to be neither satisfiable nor unsatisfiable. As an extreme case, consider $P = \{\}$ and $Q = D^k$ then every (P, Q)-PCSP (Ψ_P, Ψ_Q) has the property Ψ_P is unsatisfiable but Ψ_Q is satisfiable, so the PCSP is neither satisfiable or unsatisfiable. As such, the main computational problem we seek to study is a promise decision problem.

Definition 2.3. Let $\Gamma = \{(P_i, Q_i) \subseteq D^{k_i} \times D^{k_i}\}$ be a set of promise relations. $\mathrm{PCSP}(\Gamma)$ is the following promise decision problem. Given a Γ-PCSP $\Psi = (\Psi_P, \Psi_Q)$, output YES if Ψ is satisfiable and output NO if Ψ is unsatisfiable.

Note that $\mathrm{PCSP}(\Gamma)$ is in promiseNP since we can easily check in polynomial time if an assignment satisfies Ψ_P. We implicitly allow repetition of the variables in a specific clause. We show in Section 6.6 that removing this assumption does not meaningfully change the complexity of the problem.

8

Remark. An equivalent notation which is used in subsequent works (e.g., [2]) is to consider sets $\mathbb{A} = \{P_i : i \in [r]\}$ and $\mathbb{B} = \{Q_i : i \in [r]\}$ and denote $\mathrm{PCSP}(\Gamma)$ by $\mathrm{PCSP}(\mathbb{A}, \mathbb{B})$. We keep with the former notation in this paper.

2.1 Polymorphisms

As it can be quite cumbersome to find a direct NP-hardness reduction for $\mathrm{PCSP}(\Gamma)$, we study the combinatorial properties of a set of functions known as *polymorphisms*, which have served well as a proxy for the computational complexity of PCSPs [1, 9].

Definition 2.4. Let $(P, Q) \in D^k \times D^k$ be a promise relation. A polymorphism of (P, Q) is a function $f : D^L \to D$ such that for all $(x_1^{(1)}, \ldots, x_k^{(1)}), \ldots, (x_1^{(L)}, \ldots, x_k^{(L)}) \in P$ then $(f(x_1^{(1)}, \ldots, x_1^{(L)}), \ldots, f(x_k^{(1)}, \ldots, x_k^{(L)})) \in Q$. Denote this set of polymorphisms as $\mathrm{Pol}(P, Q)$. If $\Gamma = \{(P_i, Q_i): i \in [r]\}$ is a set of promise relations, then $f : D^L \to D$ is a polymorphism of Γ iff f is a polymorphisms of (P_i, Q_i) for all i.

We let $\mathrm{Pol}(\Gamma)$ denote the set of polymorphisms of Γ. Note that the projection maps $\pi_i(x) = x_i$ are polymorphisms of every promise relation. Further note that $\mathrm{Pol}(\Gamma) = \bigcap_{(P_i, Q_i) \in \Gamma} \mathrm{Pol}(P_i, Q_i)$.

When $P_i = Q_i$, these polymorphisms are the polymorphisms studied in the CSP literature (e.g. [27]). Sadly, when $P_i \neq Q_i$, the polymorphisms are no longer easily composable, so we no longer have necessarily that our polymorphisms form a clone. However, we still have one key property of a clone, that the polymorphisms are closed under *projections*[11].

Definition 2.5. Let $f : D^R \to D$ be a polymorphism of a family of promise relations Γ. Let $\pi : [R] \to [L]$ be a map. A *projection* $f^\pi : D^L \to D$ is the map $(f^\pi)(x) = f(y), \forall i, y_i = x_{\pi(i)}$. It is straightforward to verify that $f^\pi \in \mathrm{Pol}(\Gamma)$.

For the remainder of the article, we assume that $D = \{0, 1\}$.

Definition 2.6. Let $f : \{0, 1\}^L \to \{0, 1\}$ be a polymorphism of a family of Boolean promise relations Γ. We say that f is *folded* if $f(x) = \neg f(\bar{x})$ for all $x \in \{0, 1\}^L$. We say that a family of promise relations Γ is folded if all of its polymorphisms are folded.

It is straightforward to show that if Γ contains the NOT relation ($P = Q = \{(0, 1), (1, 0)\}$) then all polymorphisms are folded. Furthermore, note that projections of folded functions are also folded.

We will also view a polymorphism as *generating* a set of promise relations Γ' from a set of relations Γ.

Definition 2.7. Let $f : \{0, 1\}^L \to \{0, 1\}$ be a polymorphism, and let $P \subseteq \{0, 1\}^k$ be a relation. Define $f(P)$ to be

$$f(P) := \{x \in \{0, 1\}^k : \text{exist } x^{(1)}, \ldots, x^{(L)} \in P \text{ such that } x_i = f(x_i^{(1)}, \ldots, x_i^{(L)}) \text{ for all } i \in [k] = \{1, \ldots, k\}\}$$

This could also be written as $f(P) := f(P^L)$. We often state that $x = f(x^{(1)}, \ldots, x^{(L)})$, where $x^{(i)} \in P$, as a shorthand for $x_i = f(x_i^{(1)}, \ldots, x_i^{(L)})$ for all $i \in [k]$. Note that $f \in \mathrm{Pol}(P, Q)$ if and only if $f(P) \subseteq Q$.

What is the motivation for studying these polymorphisms? Roughly, if Γ has an interesting family of polymorphisms, then we expect for that family to 'beget' a polynomial-time algorithm for $\mathrm{PCSP}(\Gamma)$. The following are examples of families of polymorphisms which will yield algorithms. For all of these functions, we have that our domain is $x \in \{0, 1\}^L$.

- The zero and one functions: $\mathrm{Zero}_L(x) = 0$, $\mathrm{One}_L(x) = 1$.

---

[11]Also called *minors* in the literature.

- The AND and OR functions: $\text{AND}_L(x) = \bigwedge_{i=1}^{L} x_i$, $\text{OR}_L(x) = \bigvee_{i=1}^{L} x_i$.

- The Parity function: $\text{Par}_L(x) = \bigoplus_{i=1}^{L} x_i$. ($L$ odd)

- The Majority function: $\text{Maj}_L(x) = 1$ if $\sum_{i=1}^{L} x_i > L/2$ and 0 otherwise (L odd).

- The Alternating-Threshold function: $\text{AT}_L(x) = 1$ if $\sum_{i=1}^{L}(-1)^{i-1}x_i > 0$ and 0 otherwise (L odd).

Note that except for the Alternating-Threshold operator, all of these polymorphisms appear in the modern treatment of Schaefer's Theorem. Although the Alternating-Threshold operator is a polymorphism of some traditional Boolean CSPs, such as 2-coloring, in those cases it is possible to show that Majority is also present as a polymorphism. We will see later that this is not the case for PCSPs. Note that the arity-3 Alternating-Threshold operator would be considered a *Mal'tsev operator* in traditional CSPs (e.g. [19]).

In addition to these polymorphisms, we also use the prefix 'anti-' to refer to the negations of these functions. The 'anti-' polymorphism will be denoted with a horizontal bar. For example, anti-parity is $\overline{\text{Par}}_L(x) = \neg\,\text{Par}_L(x)$. Note that the One function is the 'anti-Zero' function and vice-versa. These polymorphisms appear due to technicalities of the nature of promise-CSPs. In Section 2.3, we show that these anti-s can be transformed into normal polymorphisms.

In Section 3, we show that if $\text{Pol}(\Gamma)$ contains any one of these infinite families of polymorphisms, then $\text{PCSP}(\Gamma)$ is tractable.

2.2 Decoding

As mentioned in the introduction, one formulation of the Algebraic CSP Dichotomy Theorem is that for any finite set of finite (traditional) relations Γ, the decision problem on the satisfiability of CSPs with clauses from Γ is in P if and only if Γ a weak near-unanimity operator. In the case of Promise CSPs, the picture is known to be not as clean. For example, [1, 9] both study NP-hard PCSPs in which some of the polymorphisms depend non-trivially on multiple coordinates. In both of those works, the polymorphisms depend on a *bounded* number of coordinates, either literally or after correcting some noise. By utilizing these polymorphisms as gadgets in a suitable probabilistically checkable proof, such as Label Cover, hardness was obtained.

In [9], we approached understanding these polymorphisms of NP-hard PCSPs using a *robust decoding* framework which identified influential coordinates in these polymorphism in a manner amenable to Label Cover. In this paper, to identify influential coordinates we will use the concept of a *C-fixing junta*.

Definition 2.8. Let $f : \{0, 1\}^L \to \{0, 1\}$ be a folded polymorphism. We say that a folded polymorphism is a *C-fixing junta* (or just *C-fixing*) if there exists $S \subseteq \{1, \ldots, L\}$ with $|S| \leq C$ such that if $x \in \{0, 1\}^L$ satisfies $x_i = 0$ for all $i \in S$, then $f(x) = f(0, \ldots 0)$.

In Section 5, we show that for any folded family of promise relations Γ all of whose polymorphisms are C-fixing, then $\text{PCSP}(\Gamma)$ is NP-hard. In Section 4, we show for a large class of Γ that their polymorphisms are C-fixing via combinatorial arguments.

2.3 Idempotence

Define a function $f : \{0, 1\}^L \to \{0, 1\}$ to be *idempotent* if $f(0, \ldots, 0) = 0$ and $f(1, \ldots, 1) = 1$. We say that a family of promise relations Γ is idempotent if all polymorphisms are idempotent.

Proposition 2.9. *For any relation $P \subseteq \{0, 1\}^k$ and any idempotent function f, we have that $P \subseteq f(P)$.*

Proof. For every $x \in P$ note that $x_i = f(x_i, \ldots, x_i)$; thus $x \in f(P)$. □

10

We say that f *generates* the promise relation $(P, f(P))$ from P. If $\Gamma = \{P_i : i \in [r]\}$ is a set of relations, then $f(\Gamma) = \{(P_i, f(P_i)) : i \in [r]\}$. Essentially by definition, $f(\Gamma)$ has f as a polymorphism.

We can force the polymorphisms of a family of promise relations Γ to be idempotent by adding in the unary promise relations SET-ZERO $= (\{(0)\}, \{(0)\})$ and SET-ONE $= (\{(1)\}, \{(1)\})$.

Proposition 2.10. *For any family of promise relations Γ, the set of idempotent promise relations of Γ is exactly* $\mathrm{Pol}(\Gamma \cup \{\mathrm{SET\text{-}ZERO}, \mathrm{SET\text{-}ONE}\})$.

Proof. For any idempotent $f \in \mathrm{Pol}(\Gamma)$, we have $f(0, \ldots, 0) = 0$ and $f(1, \ldots, 1) = 1$, so $f \in \mathrm{Pol}(\mathrm{SET\text{-}ZERO})$ and $f \in \mathrm{Pol}(\mathrm{SET\text{-}ONE})$. Likewise, every polymorphism of $\Gamma \cup \{\mathrm{SET\text{-}ZERO}, \mathrm{SET\text{-}ONE}\}$ is idempotent. $\square$

For a relation Q, define $\neg Q = \{\bar{x} : x \in Q\}$. If (P, Q) is a promise relation, it is not longer clear that $(P, \neg Q)$ is a promise relation, because we might not have that $P \subseteq \neg Q$. If we assume *non-degeneracy* and that (P, Q) has a non-idempotent polymorphism, then this is the case.

Definition 2.11. A function $f : \{0,1\}^L \to \{0,1\}$ is *non-degenerate* if $f(0, \ldots, 0) \neq f(1, \ldots, 1)$. A family of promise relations Γ is non-degenerate if all of its polymorphisms are *non-degenerate*.

One can verify that Γ is non-degenerate if and only if $\mathrm{Zero}_1, \mathrm{One}_1 \notin \mathrm{Pol}(\Gamma)$.

Proposition 2.12. *Let (P, Q) be a promise relation with a non-degenerate, non-idempotent polymorphism $f : \{0,1\}^L \to \{0,1\}$. Then, $(P, \neg Q)$ is a promise relation, and $\neg f$ is a idempotent polymorphism of this promise relation.*

Proof. Since f is non-idempotent and non-degenerate, we have that $f(0, \ldots, 0) = 1$ and $f(1, \ldots, 1) = 0$. Thus, for any $x \in \{0,1\}^L$, we have that $f(x, \ldots, x) = \bar{x}$. Since $f \in \mathrm{Pol}(P, Q)$, we thus have that $\neg P \subseteq Q$. Thus, $P \subseteq \neg Q$, so $(P, \neg Q)$ is a promise relation. It is easy to then see that for any $x^1, \ldots, x^L \in P$, since $f(x^1, \ldots, x^L) \in Q$, $\neg f(x^1, \ldots, x^L) \in \neg Q$. Thus, $\neg f$, which is idempotent, is a polymorphism of $(P, \neg Q)$. $\square$

Thus, if a non-degenerate family of promise relations Γ has at least one non-idempotent polymorphism, we may define $\neg \Gamma = ((P_i, \neg Q_i) : (P_i, Q_i) \in \Gamma)$ as another family of promise relations. Note that since Γ always has idempotent polymorphisms (the dictators/projections), $\neg \Gamma$ thus has non-idempotent polymorphisms, so $\neg(\neg \Gamma)$ exists and is equal to Γ. Thus, the idempotent polymorphisms of Γ are exactly the non-idempotent polymorphisms of $\neg \Gamma$ and vice-versa. We can formally show that the idempotent polymorphisms of Γ and $\neg \Gamma$ capture the computational complexity of $\mathrm{PCSP}(\Gamma)$.

Lemma 2.13. *Let Γ be a non-degenerate family of promise relations with at least one non-idempotent polymorphism. Let $\Gamma' = \Gamma \cup \{\mathrm{SET\text{-}ZERO}, \mathrm{SET\text{-}ONE}\}$ and $\Gamma'' = (\neg \Gamma) \cup \{\mathrm{SET\text{-}ZERO}, \mathrm{SET\text{-}ONE}\}$. Then*

1. $\mathrm{Pol}(\Gamma) = \mathrm{Pol}(\Gamma') \cup (\neg \mathrm{Pol}(\Gamma''))$, *where* $\neg \mathrm{Pol}(\Delta) = \{\neg f : f \in \mathrm{Pol}(\Delta)\}$.

2. *If* $\mathrm{PCSP}(\Gamma')$ *or* $\mathrm{PCSP}(\Gamma'')$ *is polynomial-time tractable, then so is* $\mathrm{PCSP}(\Gamma)$.

Proof. 1. In Proposition 2.10, we have that the polymorphisms of $\mathrm{Pol}(\Gamma')$ are exactly the idempotent polymorphisms of $\mathrm{Pol}(\Gamma)$. From Proposition 2.12 and the subsequent discussion, we have that the non-idempotent polymorphisms of Γ are exactly the negations of the idempotent polymorphisms of $\neg \Gamma$ which are the polymorphisms of Γ''. Thus, $\mathrm{Pol}(\Gamma) = \mathrm{Pol}(\Gamma') \cup (\neg \mathrm{Pol}(\Gamma''))$ since every polymorphism of Γ is either idempotent or non-idempotent and every polymorphism of Γ' and Γ'' is idempotent.

2. Since $\Gamma \subseteq \Gamma'$, we have that if $\mathrm{PCSP}(\Gamma')$ is polynomial-time tractable, then $\mathrm{PCSP}(\Gamma)$ is polynomial-tractable by applying the exact same algorithm. Now, assume that $\mathrm{PCSP}(\Gamma'')$ is polynomial-time tractable.

11

Since $\neg\Gamma \subseteq \Gamma''$, we have that $\mathrm{PCSP}(\neg\Gamma)$ is polynomial-time tractable. Consider an instance $\Psi = (\Psi_P, \Psi_Q)$ of $\mathrm{PCSP}(\Gamma)$. Let $\Psi^{\neg} = (\Psi_P, \Psi_{\neg Q})$ be an instance of $\mathrm{PCSP}(\neg\Gamma)$ in which every Q_i clause of Ψ_Q is replaced with a $\neg Q_i$ clause. Clearly Ψ_P is satisfiable if and only if Ψ_P is satisfiable and Ψ_Q is satisfiable if and only if $\Psi_{\neg Q}$ is satisfiable (a satisfying assignment to one is the negation of a satisfying assignment to the other). Thus, Ψ is satisfiable if and only if $\Psi^{\neg}$ is satisfiable. Thus, if we run the algorithm for $\mathrm{PCSP}(\neg\Gamma)$ which decides $\Psi^{\neg}$, we have also solved the problem in polynomial time for Ψ. $\qquad\square$

In the proceeding sections, we utilize this lemma repeatedly so that we do not need to separately consider the non-idempotent polymorphisms.

2.4 Symmetric PCSPs

The primary focus of this paper is the study of Γ in which every relation is *symmetric*.

Definition 2.14. A relation $P \subseteq \{0,1\}^k$ is symmetric if for all $x \in P$ and all permutations $\sigma : \{1,\ldots,k\} \to \{1,\ldots,k\}$, we have that $(x_{\sigma(1)},\ldots,x_{\sigma(k)}) \in P$. We say that a family of promise relations $\Gamma = \{(P_i, Q_i) : i \in [r]\}$ is symmetric if P_i and Q_i are symmetric for all i.

For a symmetric family of promise relations $\Gamma = \{(P_i, Q_i) : i \in [r]\}$, we have that each P_i and Q_i is uniquely determined by its arity and the Hamming weights of the elements. We let $\mathrm{Ham}_k(S) = \{x \in \{0,1\}^k : |x| \in S\}$ denote these sets. For example, $\mathrm{NOT} = \{(0,1),(1,0)\} = \mathrm{Ham}_2(\{1\})$. Furthermore, the idempotence relations SET-ZERO and SET-ONE are also symmetric, so adding these relations to a symmetric family of promise relations preserves that the family is symmetric. The following property of symmetric relations helps us when working with polymorphisms.

Proposition 2.15. *Let P be a symmetric relation. Let $f : \{0,1\}^L \to \{0,1\}$ be any function. Then, $f(P)$ is symmetric.*

Proof. For any $y \in f(P)$ and permutation $\sigma : \{1,\ldots,k\} \to \{1,\ldots,k\}$, consider the $x^1,\ldots,x^L \in P$ such that $f(x^1,\ldots,x^L) = y$. If we apply σ to the coordinates of $x^1,\ldots,x^L$, they will stay in P (since P is symmetric). Furthermore, f applies to these permuted variables with be σ applied to the coordinates of y. $\qquad\square$

In the remainder of the paper, we prove the following result. Note that Theorem 1.5 follows as a corollary.

Theorem 2.16 (Main Result)**.** *Let Γ be a folded, symmetric, finite family of promise relations. If at least one of Par_L, Maj_L, AT_L, $\overline{\mathrm{Par}}_L$, $\overline{\mathrm{Maj}}_L$, or $\overline{\mathrm{AT}}_L$ is a polymorphism of Γ for all odd L, then $\mathrm{PCSP}(\Gamma)$ is polynomial-time tractable. Otherwise, $\mathrm{PCSP}(\Gamma)$ is NP-hard.*

3 Efficient Algorithms

In this section, we show that if a finite collection of promise relations Γ has a polymorphism of a certain kind, then there exists a polynomial-time algorithm for solving $\mathrm{PCSP}(\Gamma)$. Note that we need not assume that the relations of Γ are symmetric. We let k refer to the maximum arity of any relation of Γ.

3.1 Zero, One, AND, OR, Parity

In each of these cases, we will reduce the PCSP Γ to a traditional CSP Γ' with the same polymorphism, which we can then solve in polynomial time by virtue of Schaefer's theorem. See [27].

Lemma 3.1. *Let* $\Gamma = \{(P_i, Q_i) : i \in \{1, \ldots, \ell\}\}$ *be a finite family of promise relations, each of arity at most* k. *Suppose that* Γ *has* f *as a polymorphism, in which* $f \in \{\text{Zero}_1, \text{One}_1, \text{AND}_{2^k}, \text{OR}_{2^k}, \text{Par}_{2^k+1}\}$. *Then,* $\text{PCSP}(\Gamma)$ *is polynomial-time tractable.*

Proof. If for some $(P_i, Q_i) \in \Gamma$, P_i is the empty relation, we can check if our Γ-PCSP has a P_i clause and reject, otherwise, we run the polynomial time algorithm for the promise relation family $\Gamma \setminus \{(P_i, Q_i) : i \in [r]\}$. Thus, we may without loss of generality assume that no P_i of Γ is the empty relation.

For each possible f, we reduce the family of promise relations Γ to $\Gamma' = \{R_i = f(P_i) \cup P_i : i \in \{1, \ldots, \ell\}\}$. We must have that $P_i \subseteq R_i \subseteq Q_i$, so the reduction is immediate (replace each (P_i, Q_i) clause with a corresponding R_i clause). We now show that $\text{CSP}(\Gamma')$ is tractable in each case.

Case 1, $f = \text{Zero}_1$. For all $i \in \{1, \ldots, \ell\}$, note that $f(P_i) = \{(0, \ldots, 0)\}$. Thus, for all $R_i \in \Gamma'$, R_i is closed under Zero_1. Thus Γ' has Zero_1 as a polymorphism and so $\text{CSP}(\Gamma')$ is polynomial-time tractable. Setting every variable to 0 satisfies the Γ'-CSP.

Case 2, $f = \text{One}_1$. This is identical to Case 1, except $f(P_i) = \{(1, \ldots, 1)\}$.

Case 3, $f = \text{AND}_{2^k}$. Since $2^k \geq |P_i|$, the bitwise-AND of every subset of P_i must be in R_i. Thus, we have that R_i must be closed under the AND_2 operator. Thus, Γ' has AND_2 as a polymorphism and is polynomial-time tractable.

Case 4, $f = \text{OR}_{2^k}$. Essentially the same as Case 3.

Case 5, $f = \text{Par}_{2^k+1}$. Since $2^k+1 > |P_i|$, the bitwise-XOR of every odd-sized subset of P_i is in R_i. Thus, R_i is closed under the Par_3 operator (the symmetric difference of 3 odd-sized subsets is an odd-sized subset). Thus, Γ' has Par_3 as a polymorphism and so it is polynomial-time tractable via a Gaussian-elimination algorithm. □

3.2 Majority and Alternating-Threshold

The algorithms in the previous section used the fact that $\text{PCSP}(\Gamma)$ has a tractable CSP Γ' that is 'sandwiched' by Γ. If Γ has the Maj_L or AT_L polymorphism for all odd L, it is no longer always the case that the polymorphisms of Γ contain the polymorphisms of a tractable CSP.[12] Instead, we demonstrate tractability by writing any Γ-PCSP $\Psi = (\Psi_P, \Psi_Q)$ as a linear programming relaxation. This approach generalizes that of [1]. The following is the pseudocode for establishing the existence of a solution.

- Construct the LP relaxation:

  - For each variable x_j of Ψ_P, stipulate that $0 \leq v_j \leq 1$.
  - For each clause $P_i(x_{j_1}, \ldots, x_{j_{k_i}})$ in Ψ_P, stipulate that $(v_{j_1}, \ldots, v_{j_{k_i}})$ is in the convex hull of the elements of P_i.

- For each variable x_j of Ψ_P.

  - Fix $v_j = 0$ (fix no other variables) and re-solve the LP.
  - If no solutions, fix $v_j = 1$ and re-solve the LP.
  - If still no solutions, output 'unsatisfiable.'

- Output 'satisfiable.'

---

[12] This follows from the fact that the composition of Maj_L (or AT_L) with itself is not a higher-arity Maj (or AT) function. In other words, the closure of the family of majority (or alternating-threshold) functions under identification do not form a clone.

Remark. Since Γ is fixed and finite, the size of the LP relaxation is within a constant factor of the size of the instance PCSP. Also, this LP does not require an objective function, since we are only concerned whether the LP has any solution.

Remark. It is worth noting that a different algorithm also exists for the Alternating-Threshold polymorphism. For each $P_i(x_{j_1}, \ldots, x_{j_{k_i}})$ in Ψ_P, write the minimal system of linear equations over $\mathbb{Z}$ such that every element of P_i is a solution (this is known as the *affine hull* of P_i). Then, solve this system of linear equations using Gaussian elimination over $\mathbb{Z}$ (e.g., [38]). Clearly if the system is infeasible, then Ψ_P is unsatisfiable. For any solutions $(v_1, \ldots, v_n)$ to this system, then $(w_1, \ldots, w_n)$ where

$$
w_i = \begin{cases} 1 & v_i \geq 1 \\ 0 & v_i \leq 0 \end{cases}
$$

is a solution to Ψ_Q.

Proof. Note that the algorithm did not distinguish whether the family $\{\mathrm{Maj}_L\}$ or $\{\mathrm{AT}_L\}$ were the polymorphisms. The reason the algorithm works, however, differs for these two cases.

First, assume that Ψ_P is satisfiable. Then, there must exist an integer solution to the linear program. Thus, for each variable x_j, there must the LP must be feasible for at least one of $v_j = 0$ or $v_j = 1$. Therefore, the algorithm always correctly reports satisfiable in this case.

Now, consider the case that Ψ_Q is unsatisfiable. Assume for the sake of contradiction, that our algorithm incorrectly reports satisfiable on input Ψ. Thus, from our checks, we have that there exists a matrix $M \in [0,1]^{n \times n}$ of solutions (on the columns) such that $M_{i,i} \in \{0,1\}$ for all $i \in [n]$. Note that we may assume that the entries of M are rational. Furthermore, any convex combination of these n solutions will yield a new solution to the original LP. In other words, for any column vector $v \in [0,1]^n$, the sum of whose weights is 1, we have that Mv is also a solution to the LP. Now, we split into cases.

Case 1, Maj_L is a polymorphism of Γ for all odd L.

We claim that there is $v \in [0,1]^n$ with sum of coordinates 1 such that $(Mv)_i \neq 1/2$ for all $i \in [n]$. Consider w with the right properties such that Mw has a minimal number of coordinates equal to $1/2$. If the number of such coordinates is 0, we are done. Otherwise, consider a coordinate j such that $(Mw)_j = 1/2$. Let $\epsilon = \min\{|(Mw)_i - 1/2|, (Mw)_i \neq 1/2\}/n$. Consider $w' = (1 - \epsilon/2)w + (\epsilon/2)e_j$, where $e_j \in \{0,1\}^n$ has value 1 in the jth coordinate and 0 everywhere else. Note, then that $|(Mw)_i - (Mw')_i| \leq \epsilon/2$ for all i, so Mw' will not have any new coordinates equal to $1/2$. Furthermore, since $(Me_j)_j$ is an integer (by construction of M), we have that the $(Mw')_j = 1/2 \pm \epsilon/4 \neq 1/2$ also. Thus, Mw' has fewer coordinates equal to $1/2$, violating the minimality of w. Thus, we can find a v such that $(Mv)_i \neq 1/2$ for all $i \in [n]$.

Thus, now we know that such a v exists, we may consider $\epsilon = \min\{|(Mv)_j - 1/2|\} > 0$. We may perturb v slightly to v' with all of its coordinates rational so that $(Mv')_j \neq 1/2$ for all j. Since the coefficients of M are rational, we have that $w = Mv$ has rational entries all not equal to $1/2$. We claim that $x_i^* = \lfloor w_i \rceil$ (w_i rounded to the nearest integer) is a satisfying assignment to Ψ_Q. Now, consider any clause $P_i(x_{j_1}^*, \ldots, x_{j_{k_i}}^*)$ of Ψ_P, and enumerate the integral points $x^1, \ldots, x^{|P|} \in P$. Since w is a rational solution to the LP, we have that there exists $\alpha_1, \ldots, \alpha_{|P|} \in \mathbb{Q} \cap [0,1]$ which sum to 1 such that $(w_{j_1}, \ldots, w_{j_{k_i}}) = \alpha_1 x^1 + \cdots + \alpha_{|P|} x^{|P|}$. Pick an integer $N \in \mathbb{N}$ which is a common denominator of $\alpha_1, \ldots, \alpha_{|P|}$. Consider $L = 2N + 1$. Since Maj_L is a polymorphism of (P, Q), we have that that the majority of $2\alpha_1 N$ copies of x^1, up to $2\alpha_{|P|}N$ copies of x^N and an extra copy of x^1 (which has no effect) is in Q. It is easy to verify that this majority is the rounding of the entries of $(w_{j_1}, \ldots, w_{j_{k_i}})$ to the nearest integer. Thus, a satisfying assignment to Ψ_Q exists, a contradiction.

Case 2, AT_L is a polymorphism of Γ for all odd L.

Let $\hat{w}$ be any rational solution to the LP. Using an argument similar to that in Case 1, we may find $v, w \in [0,1]^n \cap \mathbb{Q}^n$ with $\sum_i v_i = 1$ such that $w = Mv$ and $w_i \neq \hat{w}_i$ for all $i \in [n]$ such that $\hat{w}_i \notin \{0,1\}$

(otherwise, it may be the case that $w_i = \hat{w}_i$ for all possible w_i). We now claim that the following is a satisfying assignment to Ψ_Q.

$$\forall i \in [n], \; x_i^* = \begin{cases} 0 & w_i < \hat{w}_i \text{ or } w_i = \hat{w}_i = 0 \\ 1 & w_i > \hat{w}_i \text{ or } w_i = \hat{w}_i = 1 \end{cases}.$$

Consider any clause $P_i(x_{j_1}^*, \ldots, x_{j_{k_i}}^*)$ of Ψ_P and enumerate the integral points $x^1, \ldots, x^{|P|} \in P$. Let $\alpha_1, \ldots, \alpha_{|P|}, \hat{\alpha}_1, \ldots, \hat{\alpha}_{|P|} \in [0,1] \cap \mathbb{Q}$ be the weights such that $(w_{j_1}, \ldots, w_{j_{k_i}}) = \alpha_1 x^1 + \cdots + \alpha_{|P|} x^{|P|}$ and $(\hat{w}_{j_1}, \ldots, \hat{w}_{j_{k_i}}) = \hat{\alpha}_1 x^1 + \cdots + \hat{\alpha}_{|P|} x^{|P|}$. Let N be a common denominator of the α_i's and $\hat{\alpha}_i$'s. For $L = 4N+1$, plug into the odd-indexed entries of AT_L, $2N\alpha_i$ copies of x^i for all $i \in \{1, \ldots, |P|\}$ and one extra copy of x^1 (which will not affect the output of the polymorphism). Into the even-indexed entries plug in $2N\hat{\alpha}_i$ copies of x^i for all $i \in \{1, \ldots, |P|\}$. For each coordinate $\ell \in \{1, \ldots, k\}$, if $w_{j_\ell} = \hat{w}_{j_\ell} \in \{0, 1\}$, then when computing the ℓth coordinate, AT_L will have every input equal to $x_{j_\ell}^*$ and thus will output that same value, as desired. If $w_{j_\ell} < \hat{w}_{j_\ell}$, then there will be strictly more 1s in the even coordinates than in the odd coordinates, so AT_L will output 0 which agrees with our solution $x_{j_\ell}^*$. Finally, if $w_{j_\ell} > \hat{w}_{j_\ell}$, then there will be strictly more 1s in the odd coordinates than in the even coordinates, so AT_L will output 1 which agrees with our solution $x_{j_\ell}^*$. Therefore Ψ_Q is indeed satisfiable, contradiction.

End Cases. $\qquad\square$

Remark. This only checks whether $\Psi = (\Psi_P, \Psi_Q)$ is satisfiable, but does not find a solution when satisfiable. The proof of correctness may be modified to achieve polynomial-time algorithms for finding a satisfying assignment. See also the subsequent work by the same authors [11, 12] which have a substantial discussion on the differences between decision and search as well as streamlined and generalized the algorithms in this section to a wider range of polymorphisms.

3.3 Non-idempotent polymorphisms

Consider a family of promise relations Γ. If Zero_1 or One_1 is a polymorphism of Γ, as previously mentioned, it is polynomial-time tractable. Thus, now consider Γ non-degenerate. What if $\mathrm{Pol}(\Gamma)$ has none of the idempotent families of polymorphisms mentioned in this section, but it has one of the non-idempotent families (such as $\overline{\mathrm{Maj}}_L$ for all odd L)? Then, by Proposition 2.12, the non-idempotent version of this family yields the corresponding idempotent family of polymorphisms of $\Gamma'' = (\neg\Gamma) \cup \{\mathrm{SET\text{-}ZERO}, \mathrm{SET\text{-}ONE}\}$. From the previous sections, we then have that Γ'' is polynomial-time tractable. Therefore, by Lemma 2.13, that Γ itself is polynomial-time tractable. Hence, we have proved the following.

Theorem 3.2. *Let Γ be a finite family of promise relations. If at least one of Zero_L, One_L, AND_L, OR_L, $\overline{\mathrm{AND}}_L$, or $\overline{\mathrm{OR}}_L$ is a polymorphism of Γ for all L, or Par_L, Maj_L, AT_L, $\overline{\mathrm{Par}}_L$, $\overline{\mathrm{Maj}}_L$, or $\overline{\mathrm{AT}}_L$ is a polymorphism of Γ for all odd L, then $\mathrm{PCSP}(\Gamma)$ is polynomial-time tractable.*

Note that we did not assume that Γ is symmetric for our algorithms. That assumption will be incorporated into the NP-hardness arguments.

4 Classification of Polymorphisms of Folded, Symmetric Promise Relations

Consider any family Γ of finitely many symmetric promise relations which contains the NOT relation. We showed in Section 3 if the polymorphisms of Γ contain any of Par_L, Maj_L, AT_L, $\overline{\mathrm{Par}}_L$, $\overline{\mathrm{Maj}}_L$, or $\overline{\mathrm{AT}}_L$ for all odd L, then Γ is polynomial-time tractable. We show in this section that if Γ does not have any of these as polymorphisms for all odd L, then Γ's polymorphisms are rather skewed. Explicitly, we show that every polymorphism $f \in \mathrm{Pol}(\Gamma)$ is a "C-fixing junta." This is a key element of the NP-hardness proof in Section 5.

4.1 PCSP relaxation

In order to simplify our proof as well as to illuminate the crucial role of the promise relations, we introduce the notion of *relaxing* a promise relation.

Definition 4.1. Let $\Gamma = \{(P_i, Q_i): i \in [r]\}$ be a family of promise relations. We say that another family of promise relations Γ' is a *relaxation* of Γ if $\mathrm{Pol}(\Gamma) \subseteq \mathrm{Pol}(\Gamma')$.

Intuitively, a larger set of polymorphisms should make the PCSP easier. In Section 6.1, we confirm this by showing that if $\mathrm{Pol}(\Gamma) \subseteq \mathrm{Pol}(\Gamma')$, then there is a polynomial time reduction from $\mathrm{PCSP}(\Gamma')$ to $\mathrm{PCSP}(\Gamma)$. This fact follows from the *Galois correspondence* of polymorphisms and PCSPs. Therefore, since our aim is to demonstrate the NP-hardness of $\mathrm{PCSP}(\Gamma)$, it suffices to show that $\mathrm{PCSP}(\Gamma')$ is NP-hard for some suitable choice of Γ' that is a relaxation of Γ.

The main insight leading to our choice of Γ' is our over-arching philosophy that *polymorphisms beget algorithms*. Thus, if we ensure Γ' fails to have the polymorphisms which we showed led to polynomial-time algorithms, $\mathrm{Par}_L, \mathrm{AT}_L, \mathrm{Maj}_L$, then $\mathrm{PCSP}(\Gamma')$ should be NP-hard. In the coming subsections, we show exactly which promise relations need to be added to Γ' in order to exclude Parity, Alternating-Threshold, and Majority, while still including all of the idempotent polymorphisms of Γ.

To warm up, here is a claim about such relaxations in the symmetric case. Intuitively, this relation says we can reduce the arity of any symmetric relation in a way which respects the symmetric structure.

Claim 4.2. *Let (P, Q) be a symmetric promise relation of arity k. Let $P = \mathrm{Ham}_k(S), Q = \mathrm{Ham}_k(T)$, where $S \subseteq T \subseteq \{0, \ldots, k\}$. Then, each idempotent polymorphisms (P, Q) is a polymorphism of $(\mathrm{Ham}_{k-1}(S \setminus \{k\}), \mathrm{Ham}_{k-1}(T \setminus \{k\}))$.*

For this proof, and throughout the paper we shall use the following notation. For $S \subseteq \{1, \ldots, L\}$, we let $e_S \subseteq \{0, 1\}^L$ be such that $(e_S)_i = 1$ if and only if $i \in S$. If $S = \{i\}$ is a single element, we let $e_i = e_S$.

Proof. Let $f \in \mathrm{Pol}(P, Q)$ be any idempotent polymorphism of arity L. Consider $x^1, \ldots, x^{k-1} \in \{0, 1\}^L$ such that for all $i \in \{1, \ldots, L\}, |(x_i^1, \ldots, x_i^{k-1})| \in S \setminus \{k\}$. This implies that $|(x_i^1, \ldots, x_i^{k-1}, 0)| \in S$ for all i, so since $f \in \mathrm{Pol}(P, Q)$,
$$|(f(x^1), \ldots, f(x^{k-1}), f(0 \ldots 0))| \in T$$
Thus, since f is idempotent, $|(f(x^1), \ldots, f(x^{k-1}))| \in T \setminus \{k\}$. Thus, $f \in \mathrm{Pol}(\mathrm{Ham}_{k-1}(S \setminus \{k\}), \mathrm{Ham}_{k-1}(T \setminus \{k\})$. $\qquad\square$

Let P be any relation of arity k, and let $S \subseteq \{1, \ldots, k\}$ be any subset. Then, define
$$\mathrm{flip}_S(P) = \{y \in \{0, 1\}^k \ : \ y \oplus e_S \in P\}.$$

Note that $\neg P = \mathrm{flip}_{[k]}(P)$.

Claim 4.3. *Let (P, Q) be a promise relation of arity k, and let $S \subseteq \{1, \ldots, k\}$. Then, (P, Q) and $(\mathrm{flip}_S(P), \mathrm{flip}_S(Q))$ have identical folded polymorphisms.*

Proof. Consider any $f \in \mathrm{Pol}(P, Q)$ of arity L which is folded. Pick $x^1, \ldots, x^k \in \{0, 1\}^L$ such that $(x_j^1, \ldots, x_j^k) \in \mathrm{flip}_S(P)$ for all $j \in \{1, \ldots, L\}$. Then, consider $y^1, \ldots, y^k$ such that $y^i = \neg x^i$ if $i \in S$ and $y^i = x^i$ otherwise. Then, for all $j \in \{1, \ldots, L\}, (y_j^1, \ldots, y_j^k) \in P$. Thus, $(f(y^1), \ldots, f(y^k)) \in Q$. Due to folding, we have that $(f(x^1), \ldots, f(x^k)) \in \mathrm{flip}_S(Q)$. Thus, the folded polymorphisms of (P, Q) are polymorphisms of $(\mathrm{flip}_S(P), \mathrm{flip}_S(Q))$. By a symmetric argument, we may deduce that the folded polymorphisms of (P, Q) and $(\mathrm{flip}_S(P), \mathrm{flip}_S(Q))$ are identical. $\qquad\square$

Remark. Note that unless $S = \{\}$ or $S = \{1, \ldots, k\}$, then $\text{flip}_S(P)$ and $\text{flip}_S(Q)$ may not be symmetric if P and Q were originally symmetric. That said, in most applications, S will be one of these symmetry-preserving choices.

We can combine these two claims to get a natural corollary. This result tells us that we can shift down the Hamming weights of a symmetric, folded promise relation.

Claim 4.4. *Let (P, Q) be a symmetric promise relation of arity k. Let $P = \text{Ham}_k(S), Q = \text{Ham}_k(T)$, where $S \subseteq T \subseteq \{0, \ldots, k\}$. Then, the idempotent, folded polymorphisms of (P, Q) are polymorphisms of $(\text{Ham}_{k-1}(\{\ell : \ell \geq 0, \ell + 1 \in S\}), \text{Ham}_{k-1}(\{\ell : \ell \geq 0, \ell + 1 \in T\}))$.*

Proof. Apply Claim 4.3 on $\{1, \ldots, k\}$ to reduce the idempotent, polymorphisms of (P, Q) to $(\text{Ham}_k(\{k - \ell : \ell \in S\}), \text{Ham}_k(\{k - \ell : \ell \in T\}))$. Then, we apply Claim 4.2 to reduce further to $(\text{Ham}_{k-1}(\{k - \ell : \ell \in S\} \cap \{0, \ldots, k-1\}), \text{Ham}_{k-1}(\{k - \ell : \ell \in T\} \cap \{0, \ldots, k-1\}))$. Finally, we use Claim 4.3 again to reduce the idempotent, folded polymorphisms of (P, Q) to $(\text{Ham}_{k-1}(\{\ell : \ell \geq 0, \ell + 1 \in S\}), \text{Ham}_{k-1}(\{\ell : \ell \geq 0, \ell + 1 \in T\}))$, as desired. $\square$

In the following sections, we will repeatedly use the claims to the reduce the (P, Q) of our Γ to some simpler promise relations for which we can analyze the folded, idempotent polymorphisms.

4.1.1 Alternating-Threshold-excluding relaxation

Lemma 4.5. *Let Γ be a symmetric, folded, idempotent family of promise relations such that $\text{AT}_L \notin \text{Pol}(\Gamma)$ for some odd positive integer L, then $\Gamma' = \{(P, Q)\}$ is a relaxation of Γ, in which either*

$$P = \text{Ham}_k(\{1\}), \qquad Q = \text{Ham}_k(\{0, 1, \ldots, k - 2, k\}), k \geq 3, \text{ or}$$
$$P = \text{Ham}_k(\{0, b\}), \qquad Q = \text{Ham}_k(\{0, \ldots, k - 1\}), k \geq 2, b \in \{1, \ldots, k - 1\}.$$

Proof. As $\text{AT}_L \notin \text{Pol}(\Gamma)$, there is $(P, Q) \in \Gamma$ such that $\text{AT}_L \notin \text{Pol}(P, Q)$. Define $\text{AT}(P) = \bigcup_{L \in \mathbb{N}, \text{odd}} \text{AT}_L(P)$. Since $\text{AT}_L \notin \text{Pol}(P, Q)$ for some odd L, we have that $\text{AT}(P) \not\subseteq Q$. We claim the following.

Claim 4.6. *Consider $k \geq 1$, then*

1. $\text{AT}(\text{Ham}_k(\{0\})) = \text{Ham}_k(\{0\})$

2. $\text{AT}(\text{Ham}_k(\{k\})) = \text{Ham}_k(\{k\})$

3. $\text{AT}(\text{Ham}_k(\{0, k\})) = \text{Ham}_k(\{0, k\})$

4. $\text{AT}(\text{Ham}_k(\{\ell\})) = \text{Ham}_k(\{1, \ldots, k - 1\}), k \geq 2, \ell \in \{1, \ldots, k - 1\}$

5. $\text{AT}(\text{Ham}_k(\{\ell_1, \ell_2\})) = \{0, 1\}^k, k \geq 2, \{\ell_1, \ell_2\} \neq \{0, k\}, \ell_1 \neq \ell_2$

Proof. Facts 1-3 are easy to verify since AT_L is idempotent for all odd L.

For Fact 4, consider $\ell, \ell' \in \{1, \ldots, k - 1\}$. We claim that $\text{Ham}_k(\{\ell'\}) \subseteq \text{AT}(\text{Ham}_k(\{\ell\}))$. Pick $L = 2\ell'(k-\ell')+1$. It suffices to pick $x^1, \ldots, x^L \in \text{Ham}_k(\{\ell\})$ such that $\text{AT}_L(x^1, \ldots, x^L) = (1, \ldots, 1, 0, \ldots, 0)$, where the output has Hamming weight ℓ'. Let $x^1 = (1, \ldots, 1, 0, \ldots, 0)$, of Hamming weight ℓ, and let $x^2 = (0, \ldots, 0, 1, \ldots, 1)$, of Hamming weight ℓ. Let $x^1, x^3, \ldots, x^{L-2}$ be all possible permutations of x^1 in which the first ℓ' coordinates are cyclically shifted and the last $k - \ell'$ coordinates are cyclically shifted. There

may be repetition, but each repetition should appear an equal number of times. Likewise, let $x^2, x^4, \ldots, x^{L-1}$ be the same kind of permutations but of x^2. Let $x^L = 1$. It is easy to verify that

$$j \in \{1, \ldots, \ell'\}, \quad \sum_{i=1,\text{odd}}^{L-2} x_j^i = (k - \ell') \min(\ell, \ell')$$

$$j \in \{\ell' + 1, \ldots, k\}, \quad \sum_{i=1,\text{odd}}^{L-2} x_j^i = \ell' \max(0, \ell - \ell')$$

$$j \in \{1, \ldots, \ell'\}, \quad \sum_{i=1,\text{even}}^{L-1} x_j^i = (k - \ell') \max(0, \ell + \ell' - k)$$

$$j \in \{\ell' + 1, \ldots, k\}, \quad \sum_{i=1,\text{even}}^{L-1} x_j^i = \ell' \min(\ell, k - \ell')$$

Thus,

$$j \in \{1, \ldots, \ell'\}, \sum_{i=1}^{L} (-1)^{i-1} x_j^i = (k - \ell')(\min(\ell, \ell') - \max(0, \ell + \ell' - k)) + x_j^L \geq (k - \ell') + x^L > 0$$

$$j \in \{\ell' + 1 \ldots, k\}, \sum_{i=1}^{L} (-1)^{i-1} x_j^i = \ell'(\max(0, \ell - \ell') - \max(\ell, k - \ell')) + x_j^L < -\ell' + x_j^L \leq 0.$$

Therefore, $\mathrm{AT}_L(x^1, \ldots, x^L) \in \mathrm{Ham}_k(\{\ell'\})$, as desired. By Proposition 2.15, $\mathrm{Ham}_k(\{\ell'\}) \subseteq \mathrm{AT}(\mathrm{Ham}_k(\{\ell\}))$, for all $\ell, \ell' \in \{1, \ldots, k-1\}$, as desired.

Now, to finish Fact 4, we seek to show that $\mathrm{Ham}_k(\{0\}) = \{(0, \ldots, 0)\} \not\subseteq \mathrm{Ham}_k(\{\ell\})$. Assume for the sake of contradiction, there exists L odd and $x^1, \ldots, x^L \in \mathrm{Ham}_k(\{\ell\})$ such that $\mathrm{AT}_L(x^1, \ldots, x^L) = (0, \ldots, 0)$. Then, we have that for all $i \in \{1, \ldots, k\}$, $\sum_{j=1}^{L} (-1)^{j-1} x_i^j \leq 0$. Summing over all i, we have that $0 \geq \sum_{j=1}^{L} (-1)^{j-1} \sum_{i=1}^{k} x_i^j = \sum_{j=1}^{L} (-1)^{j-1} \ell = \ell$, a contradiction. Likewise, if $\mathrm{AT}_L(x^1, \ldots, x^L) = (1, \ldots 1)$. We would have that $k \leq \sum_{j=1}^{L} (-1)^{j-1} \sum_{i=1}^{L} x_i^j = \sum_{j=1}^{L} (-1)^{j-1} \ell = \ell$, which is also a contradiction. Thus, we have shown fact 4.

For Fact 5, since we know that $\{\ell_1, \ell_2\} \neq \{0, k\}$, we know that at least one of ℓ_1 and ℓ_2 is strictly between 1 and $k - 1$. Thus, by fact 4, $\mathrm{Ham}_k(\{1, \ldots, k-1\}) \subseteq \mathrm{AT}(\mathrm{Ham}_k(\{\ell_1, \ell_2\}))$. Therefore, it suffices to prove that $(0, \ldots, 0), (1, \ldots, 1) \in \mathrm{AT}(\mathrm{Ham}_k(\{\ell_1, \ell_2\}))$. Assume that $\ell_1 < \ell_2$ and consider $L = 4k + 1$. To show $(0, \ldots, 0) \in \mathrm{AT}(\mathrm{Ham}_k(\{\ell_1, \ell_2)\})$, pick $x^1, \ldots, x^{4k+1}$ such that x^j has Hamming weight ℓ_1 when j is odd and Hamming weight ℓ_2 when j is even. Let $x^1 = (1, \ldots, 1, 0, \ldots, 0)$, and let $x^3, x^5, \ldots$ be successive cyclic shifts. Likewise, let $x^2 = (1, \ldots, 1, 0, \ldots, 0)$ (with the appropriate number of 1s), and let $x^4, \ldots$ be successive cyclic shifts. Then, it is easy to see that for all $i \in \{1, \ldots, k\}$, we have that $\sum_{j=1}^{4k+1} (-1)^{i-1} x_i^j$ is $2\ell_1 - 2\ell_2 < 0$ or $2\ell_1 - 2\ell_2 + 1 < 0$ (because we have $2k + 1$ odd-indexed terms but $2k$ even-indexed terms). Thus, $\mathrm{AT}_{4k+1}(x^1, \ldots, x^{4k+1}) = (0, \ldots, 0)$. If we do the same construction but swap ℓ_1 and ℓ_2, we would then have that $2\ell_1 - 2\ell_2, 2\ell_1 - 2\ell_2 + 1 > 0$, so $\mathrm{AT}_{4k+1}(x^1, \ldots, x^{4k+1}) = (1, \ldots, 1)$. Thus, Fact 5 is shown. $\qquad \square$

Now, let $a \in \{0, \ldots, k\}$ be such that $\mathrm{Ham}_k(\{a\}) \subseteq \mathrm{AT}(P)$ but $\mathrm{Ham}_k(\{a\}) \not\subseteq Q$. Such an a must exist by Proposition 2.15. Note that since $P \subseteq Q$, we must have that $\mathrm{Ham}_k(\{a\}) \not\subseteq P$. We divide the remaining analysis into two cases.

Case 1, $a \in \{1, \ldots, k-1\}$.

Then, by Fact 3 of the claim, there must exists $\ell \in \{1, \ldots, k-1\}$ such that $\mathrm{Ham}_k(\{\ell\}) \subseteq P$ (else, $P \subseteq \mathrm{Ham}_k(\{0, k\})$ and then $\mathrm{Ham}_k(\{a\}) \not\subseteq \mathrm{AT}(P)$.) Let $P' = \mathrm{Ham}_k(\{\ell\})$ and let $Q' = \mathrm{Ham}_k(\{0, \ldots, k\} - \{a\})$.

Since $P' \subseteq P \subseteq Q \subseteq Q'$, every polymorphism of (P, Q) is a polymorphism of (P', Q'). Furthermore, by Fact 4, $\mathrm{Ham}_k(\{a\}) \subseteq \mathrm{AT}(P')$ but $\mathrm{Ham}_k(\{a\}) \not\subseteq Q'$. Thus, (P', Q') does not admit $\mathrm{AT}_{L'}$ as a polymorphism for some L'.

We may assume without loss of generality that $\ell < a$. Else we may apply Claim 4.3 with $S = [k]$. Let $k' = a + 1$. From Claim 4.2, applied $k - k'$ times, we have that all of the idempotent polymorphisms of (P', Q') are idempotent polymorphisms of

$$P^{(2)} = \mathrm{Ham}_{k'}(\{\ell\}), Q^{(2)} = \mathrm{Ham}_{k'}(\{0, \ldots, k'\} - \{a\}).$$

Likewise, applying Claim 4.4 $\ell - 1$ times, all of the folded polymorphisms of $P^{(2)}, Q^{(2)}$ are polymorphisms of

$$P^{(3)} = \mathrm{Ham}_{k''}(\{\ell'\}), Q^{(3)} = \mathrm{Ham}_{k''}(\{0, \ldots, k''\} - \{a'\})$$

where $k'' = a - \ell + 2 \geq 3$. Note that $\ell' = 1$ and $a' = k'' - 1$. Thus, the idempotent, folded polymorphisms of Γ are polymorphisms of

$$P^{(4)} = \mathrm{Ham}_{k''}(\{1\}), Q^{(4)} = \mathrm{Ham}_{k''}(\{0, \ldots, k'' - 2, k''\}).$$

Case 2, $a \in \{0, k\}$.

Without loss of generality, we may assume that $a = 0$. Otherwise, we may replace (P, Q) with $(\mathrm{flip}_{[k]}(P), \mathrm{flip}_{[k]}(Q))$, which preserves the folded, idempotent polymorphisms of Γ. Since $\mathrm{Ham}_k(\{0\}) \subseteq \mathrm{AT}(P)$ but $\mathrm{Ham}_k(\{0\}) \not\subseteq P \subseteq Q$, we must be in Fact 5 of Claim 4.6. That is, there must be $\ell_1, \ell_2 \in \{0, \ldots, k\}$ distinct and $\{\ell_1, \ell_2\} \neq \{0, k\}$ such that $\mathrm{Ham}_k(\{\ell_1, \ell_2\}) \subseteq P$. Like in Case 1, relax (P, Q) to $P' = \mathrm{Ham}_k(\{\ell_1, \ell_2\})$ and $Q' = \mathrm{Ham}_k(\{1, \ldots, k\})$. Let $k' = \max(\ell_1, \ell_2)$, and apply Claim 4.2 $k - k'$ times to yield

$$P^{(2)} = \mathrm{Ham}_{k'}(\{\min(\ell_1, \ell_2), k'\}), Q^{(2)} = \mathrm{Ham}_{k'}(\{1, \ldots, k'\}).$$

Then, applying Claim 4.3 with $S = \{1, \ldots, k'\}$, we get that

$$P^{(3)} = \mathrm{Ham}_{k'}(\{0, b\}), Q^{(3)} = \mathrm{Ham}_{k'}(\{0, \ldots, k' - 1\}), b = k' - \min(\ell_1, \ell_2) \in \{1, \ldots, k' - 1\}, k' \geq 2$$

has as polymorphisms the folded, idempotent polymorphisms of Γ, as desired.

End Cases. □

4.1.2 Majority-excluding relaxation

Lemma 4.7. *Let Γ be a symmetric, folded, idempotent family of promise relations such that $\mathrm{Maj}_L \notin \mathrm{Pol}(\Gamma)$ for some odd positive integer L, then $\Gamma' = \{(P, Q)\}$ is a relaxation of Γ, in which either*

$$P = \mathrm{Ham}_k(\{(k+1)/2\}), \qquad Q = \mathrm{Ham}_k(\{0, 1, \ldots, k-1\}), \ (k \geq 3 \ odd), \ or$$
$$P = \mathrm{Ham}_k(\{1, k\}), \qquad Q = \mathrm{Ham}_k(\{0, 1, \ldots, k\} - \{b\}), k \geq 3, b \in \{2, \ldots, k-1\}.$$

Proof. The proof proceeds in a similar manner to Lemma 4.5. Define $\mathrm{Maj}(P) = \bigcup_{L \in \mathbb{N}, \mathrm{odd}} \mathrm{Maj}_L(P)$. We begin with the analogue of Claim 4.6 for the Majority operation.

Claim 4.8. *Consider $k \geq 1$. If $P \subseteq \mathrm{Ham}_k(\{0, k\})$, then $\mathrm{Maj}(P) = P$. Otherwise, if $P = \mathrm{Ham}_k(S)$ is symmetric but $S \setminus \{0, k\}$ is nonempty, then*

$$\mathrm{Maj}(P) = \mathrm{Ham}_k(\{0, \ldots, k\} \cap \{2(\min S) - k + 1, \ldots, 2(\max S) - 1\}).$$

Proof. As before, the case $S \subseteq \{0, k\}$ is easy.

We first, show that $\mathrm{Maj}(P) \subseteq \mathrm{Ham}_k(\{0, \ldots, k\} \cap \{2 \min S - k + 1, \ldots, 2 \max S - 1\})$. For any $b \in \{0, \ldots, k\}$, such that $\mathrm{Ham}_k(\{b\}) \subseteq \mathrm{Maj}(P)$, there is L odd and $x^1, \ldots, x^L \in P$ such that $\mathrm{Maj}_L(x^1, \ldots, x^L)$ has Hamming weight b. Assume without loss of generality the coordinates equal to 1 are the first b ones. Thus, we have that for all $i \in \{1, \ldots, b\}$, $\sum_{j=1}^{L} x_i^j \geq (L+1)/2$.. Thus,

$$\sum_{i=1}^{k} \sum_{j=1}^{L} x_i^j \geq b(L+1)/2.$$

Thus, by the pigeonhole principle, there is some x^j such that its Hamming weight is at least $b(L+1)/(2L) \leq \max S$. Thus, $b \leq 2L \max S/(L+1) < 2 \max S$. Therefore, $b \leq 2 \max S - 1$, as desired. Using the fact that $\sum_{j=1}^{L} x_i^j \leq (L-1)/2$ for all $i \in \{b+1, \ldots, k\}$, we have that some x^j has Hamming weight at most $b + (k-b)(L-1)/(2L) \geq \min S$ (we add b since all of the first b coordinates may be 1s). Thus, $b \geq 2L \max S/(L-1) - k(L-1)/(L+1) > 2 \max S - k$. Therefore, $b \geq 2 \max S + 1 - k$, as desired. Thus, $\mathrm{Maj}(P) \subseteq \mathrm{Ham}_k(\{0, \ldots, k\} \cap \{2 \min S - k + 1, \ldots, 2 \max S - 1\})$.

Now, we show the reverse direction, that every $b \in \{0, \ldots, k\} \cap \{2 \min S - k + 1, \ldots, 2 \max S - 1\}$ can be obtained as a Hamming weight. Assume that there is $\ell \in S \cap \{1, \ldots, k-1\}$ (so $k \geq 2$). For ease of notation, let $s = \min S, t = \max S$, so $s \leq \ell \leq t$. To start, we show if $b \in \{0, \ldots, k\} \cap \{\ell, \ldots, 2 \max S - 1\}$, then $\mathrm{Ham}_k(\{b\}) \subseteq \mathrm{Maj}(P)$.

First, if $b \geq \max S$, consider $L = 2b+1$. We now seek to pick $x^1, \ldots, x^L$ of Hamming weight $\max S$ such that $\mathrm{Maj}_L(x^1, \ldots, x^L)$ has Hamming weight b. Let $x^1 = (1, \ldots, 1, 0, \ldots, 0)$ of the suitable Hamming weight, and for all $j \geq 2$, let x^j be the cyclic shift of x^{j-1} in the first $b \geq \max S$ coordinates. For all $i \in \{b+1, \ldots, k\}$, $\sum_j x_i^j = 0$, so $\mathrm{Maj}_L(x_i^1, \ldots, x_i^L) = 0$. For all $i \in \{1, \ldots, b\}$, $\sum_j x_i^j = 2 \max S + x_j^L \geq b + 1 + x_j^L > L/2$. Thus, $\mathrm{Maj}_L(x_j^1, \ldots, x_j^L) = 1$, so $\mathrm{Maj}_L(x^1, \ldots, x^L)$ has Hamming weight b.

Otherwise, if $b \in \{\ell, \ell+1, \ldots, \max S - 1\}$, consider now $L = 2b - 1 \geq 1$. Let $x^1 = \cdots = x^{b-1} = (1, \ldots, 1, 0, \ldots, 0)$, with Hamming weight $\max S$. Let $x^b = (1, \ldots, 1, 0, \ldots, 0)$ of Hamming weight ℓ, and let $x^{b+1}, \ldots, x^{2b-1}$ be x^b except that the first b coordinates are cyclically shifted. If $i \in \{1, \ldots, b\}$ then $\sum_i x_i^j = b - 1 + \ell \geq b > L/2$ since $\ell \geq 1$. If $i \in \{b+1, \ldots, k\}$, then $\sum_i x_i^j \leq b-1 < L/2$. Thus, $\mathrm{Maj}_L(x^1, \ldots, x^L)$ has Hamming weight b, as desired. Thus, $\mathrm{Ham}_k(\{0, \ldots, k\} \cap \{\ell, \ldots, 2 \max S - 1\}) \subseteq \mathrm{Maj}(P)$.

By an analogous argument, we may show that $\mathrm{Ham}_k(\{0, \ldots, k\} \cap \{2 \min S - k + 1, \ldots, \ell\}) \subseteq \mathrm{Maj}(P)$. A simple route to this is reversing the notions of 0 and 1 in our previous construction. □

Consider $b \in \{0, \ldots, k\}$ such that $\mathrm{Ham}_k(\{b\}) \subseteq \mathrm{Maj}(P)$ but $\mathrm{Ham}_k(\{b\}) \not\subseteq Q, P$. From Claim 4.8, $P \setminus \mathrm{Ham}_k(\{0, k\})$ must be nonempty. Thus, there is $\ell \in \{1, \ldots, k-1\}$ such that $\mathrm{Ham}_k(\{\ell\}) \subseteq P$. We may assume without loss of generality that $\ell < b$ as $\ell \neq b$ and we can apply Claim 4.3 to (P, Q) to get $(\mathrm{flip}_{[k]}(P), \mathrm{flip}_{[k]}(Q))$, which does not change the folded, idempotent polymorphisms.

Let $S \subseteq \{0, \ldots, k\}$ be such that $P = \mathrm{Ham}_k(S)$. Since $\mathrm{Ham}_k(\{b\}) \subseteq \mathrm{Maj}(P)$ and $\ell < b$, we have by the claim that $\mathrm{Ham}_k(\{b\}) \subseteq \mathrm{Maj}(\mathrm{Ham}_k(\{\ell, \max S\}))$. Thus, we can relax to $P' = \mathrm{Ham}_k(\{\ell, \max S\})$ and $Q' = \mathrm{Ham}_k(\{0, \ldots, k\} - \{b\})$ while still preserving the idempotent, folded polymorphisms of Γ. We again diverge into two cases.

Case 1, $b > \max S$.

We may relax to

$$P^{(2)} = \mathrm{Ham}_k(\{\max S\}), Q^{(2)} = \mathrm{Ham}_k(\{0, \ldots, k\} - \{b\}).$$

Let $k' = b$, and apply Claim 4.2 $k - k'$ times to relax the folded, idempotent polymorphisms to

$$P^{(3)} = \mathrm{Ham}_{k'}(\{\max S\}), Q^{(3)} = \mathrm{Ham}_{k'}(\{0, \ldots, k' - 1\}).$$

Recall that $\max S < k' = b \le 2\max S - 1$ by Claim 4.8. Thus, $k'' = 2(k' - \max S) + 1 \le k'$. Applying Claim 4.4 $k' - k''$ times, we then get that the folded, idempotent polymorphisms of Γ are also polymorphisms of

$$P^{(4)} = \mathrm{Ham}_{k''}(\{(k''+1)/2\}), Q^{(4)} = \mathrm{Ham}_{k''}(\{0, \ldots, k'' - 1\}), k'' \ge 3.$$

This establishes the first case of the lemma.

Case 2, $\ell < b < \max S$.

Letting $k' = \max S$ and applying Claim 4.2 $k - k'$ times, we get that the idempotent, folded polymorphisms of Γ are also polymorphisms of

$$P^{(2)} = \mathrm{Ham}_{k'}(\{\ell, k'\}), Q^{(2)} = \mathrm{Ham}_{k'}(\{0, \ldots, k'\} - \{b\}).$$

Now, consider $k'' = k' - \ell + 1$, and apply Claim 4.4 $k' - k''$ times to get that the idempotent, folded, polymorphisms of Γ are also polymorphisms of

$$P^{(3)} = \mathrm{Ham}_{k''}(\{1, k''\}), Q^{(3)} = \mathrm{Ham}_{k''}(\{0, \ldots, k''\} - \{b'\}), b' \in \{2, \ldots, k'' - 1\}.$$

Note that $k'' \ge 3$; therefore the second case of the lemma is also fully established.

End Cases. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

Remark. We do not have a section specific to Parity-avoiding promise relations. It is convenient enough for us to reason about excluding Parity along with another familiy of polymorphisms (e.g., Lemma 4.10).

4.2 Idempotent case

We now seek to establish that if a symmetric, idempotent, folded family of promise relations Γ avoids $\mathrm{Par}_{L_1}, \mathrm{AT}_{L_2}, \mathrm{Maj}_{L_3}$ as polymorphisms for some odd L_1, L_2, L_3, then the polymorphisms are C-fixing for some suitable constant $C(\Gamma)$. Note that this C may depend on L_1, L_2, L_3, but if we pick L_1, L_2, L_3 to be minimal, then these also depend only on Γ. Our first step is to establish the following lemma in additive combinatorics.

Lemma 4.9. *Let $S_0, S_1 \subseteq \mathbb{Z}_{\ge 0}$ such that $0 \in S_0$ and $1 \in S_1$. Assume that there exists a positive integer n such that for all $a \in S_0$ and $b_1, \ldots, b_n \in S_1$ (not necessarily distinct)*

$$b_1 + \cdots + b_n \in S_0 \tag{1}$$

$$a + b_1 + \cdots + b_{n-1} \in S_1. \tag{2}$$

If n is odd, then there is $A(n) \in \mathbb{Z}_{\ge 0}$ such that $A(n) \in S_0 \cap S_1$. Otherwise, if n is even, there is $d(n) \in \mathbb{Z}_{\ge 0}$ such that S_0 contains all even integers at least $d(n)$ and S_1 contains all odd integers at least $d(n)$.

Proof. If $n = 1$, then by (2), we have that $0 \in S_1$. Thus, we can set $A(1) = 0$. Now assume $n \ge 2$. We can easily deduce the following facts

$$\forall x \in S_0, x + n - 1 \in S_1 \qquad\qquad (a = x \text{ and } b_1, \ldots, b_{n-1} = 1 \text{ in } (2)) \tag{3}$$

$$\forall y \in S_1, y + n - 1 \in S_0 \qquad\qquad (b_1 = y \text{ and } b_2, \ldots, b_{n-1} = 1 \text{ in } (1)) \tag{4}$$

$$\forall y \in S_1, y + n - 2 \in S_1 \qquad\qquad (a = 0, b_1 = y, \text{ and } b_2, \ldots, b_{n-1} = 1 \text{ in } (2)) \tag{5}$$

In particular, we may deduce that

$$\forall x \in S_0, x + 2n - 2 \in S_0 \qquad\qquad (3 \text{ and } 4) \tag{6}$$

$$\forall x \in S_0, x + 3n - 4 \in S_0 \qquad\qquad (3, 5, \text{ and } 4) \tag{7}$$

$$\forall y \in S_1, y + n - 2 \in S_1 \qquad\qquad (5) $$

$$\forall y \in S_1, y + 2n - 2 \in S_1 \qquad\qquad (4 \text{ and } 3) \tag{8}$$

Note that if $n \geq 3$ is odd, then $\gcd(2n - 2, 3n - 4) = 1$. Therefore, by (6) and (7) and that $0 \in S_0$, we may deduce by the standard analysis to the Frobenius coin problem (e.g., [15]) (colloquially known as the Chicken McNugget Theorem) that S_0 contains all sufficiently large positive integers. Likewise, since $\gcd(n - 2, 2n - 2) = 1$ and (5) and (8) hold, we have that S_1 contains all sufficiently large positive integers. Hence, there must exist $A(n) \in \mathbb{N}$ such that $A(n) \in S_0 \cap S_1$.

If $n \geq 2$ is even, then $\gcd(2n - 2, 3n - 4) = \gcd(n - 2, 2n - 2) = 2$. Since $0 \in S_0$, and $1 \in S_1$, we may then deduce by the same theorem that S_0 will contain all sufficiently large even numbers and that S_1 will contain all sufficiently large odd numbers. Thus, we may select $d(n)$ accordingly. □

Remark. Consider the modification that there is some positive integer N such that $\max S_0, \max S_1 \leq N$ with the stipulation that (1) and (2) only apply when the sums are at most N. The theorem still holds as long as, when n is odd, the original $A(n)$ is at most N, or when n is even, the original $d(n)$ is at most N. These follow from the fact that $A(n)$ and $d(n)$ are independent of S_1, S_2.

With this established, we can now deduce significant structural properties of the polymorphisms of Alternating-Threshold and Parity-avoiding families of promise relations. These arguments have connections to those in [1], but differ significantly in details.

Lemma 4.10. *Let Γ be a symmetric, folded, idempotent family of promise relations such that $\mathrm{Par}_{L_1}, \mathrm{AT}_{L_2} \notin \mathrm{Pol}(\Gamma)$ for some odd positive integers L_1, L_2. Then, there exists $c(\Gamma) \in \mathbb{N}$ such that for all $L \in \mathbb{N}$ and for all $f : \{0, 1\}^L \to \{0, 1\} \in \mathrm{Pol}(\Gamma)$,*

$$|\{i \in \{1, \ldots, L\} : f(e_i) = 1\}| \leq c(\Gamma).$$

Proof. Fix $f \in \mathrm{Pol}(\Gamma)$ of arity L, and let $A = \{i \in \{1, \ldots, L\} : f(e_i) = 1\}$. Assume for the sake of contradiction that $|A|$ grows arbitrarily large. Define $S_0, S_1 \subseteq \{0, 1, \ldots, |A|\}$ as follows.

$$S_i = \{j : \text{for all } T \subseteq A \text{ of size } j, f(e_T) = i\}, \ i \in \{0, 1\}.$$

We seek to show that there exists $n(\Gamma)$ for which (1) and (2) hold, so that that we may invoke Lemma 4.9 on S_0 and S_1. Either, we have some f such that $S_0 \cap S_1$ is nonempty, an immediate contradiction. Or, f will have structure similar to that of parity, which we can then use to contradict that $\mathrm{Par}_{L_1} \notin \mathrm{Pol}(\Gamma)$.

To achieve the first goal, which is to show that $n(\Gamma)$ exists which satisfies (1) and (2), we utilize Lemma 4.5 to deduce a symmetric (P, Q), independent of L, such that $f \in \mathrm{Pol}(P, Q)$. Let k be the arity of P and Q. The proof now proceeds into two cases.

Case 1, $k \geq 3$, $P = \mathrm{Ham}_k(\{1\}), Q = \mathrm{Ham}_k(\{0, \ldots, k - 2, k\})$.

Let $n = k - 1$. For any $b_1, \ldots, b_{k-1} \in S_1$ such that $b_1 + \cdots + b_{k-1} \leq |A|$, consider any $T \subseteq A$ of size $b_1 + \ldots + b_{k-1}$. Partition $T = T_1 \cup T_2 \cup \cdots \cup T_{k-1}$ such that $|T_i| = b_i$ for all i. Let $T_k := [L] \setminus T$. Consider the k-tuple of L-tuples

$$(e_{T_1}, e_{T_2}, \ldots, e_{T_{k-1}}, e_{T_k}).$$

For every $i \in \{1, \ldots, L\}$, there is exactly one $j \in [k]$ with $i \in T_j$. Thus, since $f \in \mathrm{Pol}(P, Q)$, we have that

$$(f(e_{T_1}), f(e_{T_2}), \ldots, f(e_{T_{k-1}}), f(e_{T_k}))$$

has Hamming weight not equal to $k - 1$. Since $f(e_{T_i}) = 1$ for all $i \in [k - 1]$, we must then have that $f(e_{T_k}) = 1$. Since f is folded and $T = [L] \setminus T_k$, we can thus deduce that $f(e_T) = 0$, as desired. Since the choice of $T \subseteq A$ is arbitrary except for size, we have that $b_1 + \cdots + b_{k-1} \in S_0$, so (1) holds.

Now, consider any $a \in S_0$ and $b_1, \ldots, b_{k-2} \in S_1$ such that $a + b_1 + \cdots + b_{k-2} \leq |A|$. Again, consider any $T \subseteq A$ of size $a + b_1 + \cdots + b_{k-2}$. Partition $T = T_0 \cup T_1 \cup \cdots \cup T_{k-2}$ such that $|T_0| = a$ and $|T_i| = b_i$ for all other i. Let $T_{k-1} := [L] \setminus T$ Note again that the k-tuple of L-tuples

$$(e_{T_0}, \ldots, e_{T_{k-2}}, e_{T_{k-1}})$$

22

has for every $i \in \{1, \ldots, L\}$ has exactly one $j \in \{0, 1, \ldots, k-1\}$ with $j \in T_i$. Thus, we may again deduce that since $f \in \mathrm{Pol}(P, Q)$,

$$(f(e_{T_0}), \ldots, f(e_{T_{k-2}}), f(e_{T_{k-1}}))$$

has Hamming weight not equal to $k-1$. Since exactly $k-2$ of the first $k-1$ entries are equal to 1, we must have $f(e_{T_{k-1}}) = 0$. Thus, $f(e_T) = 1$. Since the choice of $T \subseteq A$ is arbitrary except for size, $a + b_1 + \cdots + b_{k-2} \in S_1$, so (2) also holds, as desired.

Case 2, $k \geq 2$, $P = \mathrm{Ham}_k(\{0, b\})$, $Q = \mathrm{Ham}_k(\{0, \ldots, k-1\})$, $b \in \{1, \ldots, k-1\}$.

Let $n = k - b + 1$. For any $b_1, \ldots, b_n \in S_1$ such that $b_1 + \cdots + b_n \leq |A|$. Consider any $T \subseteq A$ of size $b_1 + \ldots + b_n$. Partition $T = T_1 \cup T_2 \cup \cdots \cup T_n$ such that $|T_i| = b_i$ for all i. Consider the k-tuple

$$(e_{T_1}, e_{T_2}, \ldots, e_{T_n}, e_T, \ldots, e_T),$$

where e_T appears $b - 1 \geq 0$ times. We can verify that for each $i \in T$, there are exactly b tuples in this list with 1 in the ith coordinate. For any $i \notin T$, there are 0 tuples with 1 in the ith coordinate. Thus,

$$(f(e_{T_1}), \ldots, f(e_{T_n}), f(e_T), \ldots, f(e_T)) \in Q.$$

Since $f(e_{T_1}) = \cdots = f(e_{T_n}) = 1$, to avoid a contradiction, we must have that $f(e_T) = 0$, so $b_1 + \cdots + b_n \in S_0$.

For any $a \in S_0$ and $b_1, \ldots, b_{n-1} \in S_1$ such that $a + b_1 + \cdots + b_{n-1} \leq |A|$, consider $T \subseteq A$ of size $a + b_1 + \cdots + b_{n-1}$. Partition $T = T_0 \cup T_1 \cup \cdots \cup T_{n-1}$ such that $|T_0| = a$ and $|T_i| = b_i$ for all other i. It is easy to check that the following is a valid k-tuple

$$(e_{T_1}, \ldots, e_{T_{n-1}}, \neg e_{T_0}, \ldots, \neg e_{T_0}, \neg e_T),$$

where the are $k - n = b - 1$ copies of $\neg e_{T_0}$. Thus, since f applies to the first $k - 1$ tuples is equal to 1, $f(\neg e_T) = 0$, which implies by folding that $f(e_T) = 1$. Therefore, $a + b_1 + \cdots + b_{n-1} \in S_1$, as desired.

End Cases

Thus, we have established that the conditions of Lemma 4.9 hold for some $n(\Gamma)$. As stated at the beginning of the proof, we may apply the lemma to see that if $|A|$ grows arbitrarily large, then either $S_0 \cap S_1$ is nonempty for some f, which is an immediate contradiction, or S_0 contains all even integers between $d(n)$ and $|A|$ and S_1 contains all odd integers between $d(n)$ and $|A|$. To obtain a contradiction in this second case, we show that Par_{L_1} is a minor of f. Observe that $f(e_A) = 1$, implying that $|A| \in S_1$, so $|A|$ is odd. Assume that $|A| \geq L_1(d(n) + 1)$, then we can partition A into $A_1 \cup \cdots \cup A_{L_1}$ such that for all $i \in [L_1]$, $|A_i|$ is odd and at least $d(n)$. Now consider, the map $\pi : [L] \to [L_1]$ defined as follows:

$$\pi(i) = \begin{cases} j & i \in A_j \\ 1 & i \notin A. \end{cases}$$

We claim that $f^\pi = \mathrm{Par}_{L_1}$. Since f is folded, f^π is also folded. Thus, it suffices to check that $f^\pi(x) = \mathrm{Par}_{L_1}$ when $x_1 = 0$. In that case,

$$\begin{aligned} f^\pi(x) &= f(e_{\bigcup_{i:x_i=1} A_i}) \\ &= \begin{cases} 0 & \sum_{i:x_i=1} |A_i| \text{ is even} \\ 1 & \sum_{i:x_i=1} |A_i| \text{ is odd} \end{cases} \\ &= \mathrm{Par}_{L_1}(x). \end{aligned} \qquad (|A_i| \text{ odd})$$

The second line follows from the fact that each A_i has size at least $d(n)$. Therefore, $\mathrm{Par}_{L_1} = f^\pi \in \mathrm{Pol}(\Gamma)$, a contradiction.

Thus, $|A|$ is bounded, as desired. $\qquad \square$

From this lemma, we can make an even stronger conclusion.

Corollary 4.11. *Let Γ have the same properties as in Lemma 4.10. Let $f : \{0,1\}^L \to \{0,1\} \in \mathrm{Pol}(\Gamma)$ be any polymorphism and let $S_1,\ldots,S_\ell$ be disjoint subsets of $\{1,\ldots,L\}$ such that $f(e_{S_i}) = 1$ for all $i \in \{1,\ldots,\ell\}$. Then, $\ell \le c(\Gamma)$, where $c(\Gamma)$ is the same as in Lemma 4.10.*

Proof. Choose projection $\pi : \{1,\ldots,L\} \to \{1,\ldots,L\}$ such that for all $i \in \{1,\ldots,\ell\}$ and all $j \in S_i$, $\pi(j) = \min(S_i)$ and otherwise is the identity map. Consider $g = f^\pi$ which must also be a polymorphism of Γ. It is easy then to see that for all $i \in \{1,\ldots,\ell\}$, $g(e_{\min(S_i)}) = f(e_{S_i})$. Thus, $\ell \le c(\Gamma)$ by applying Lemma 4.10 to g. $\qquad\square$

Lemma 4.12. *Let Γ be a symmetric, folded, idempotent family of promise relations such that $\mathrm{Par}_{L_1}, \mathrm{AT}_{L_2}, \mathrm{Maj}_{L_3} \notin \mathrm{Pol}(\Gamma)$ for some odd positive integers L_1, L_2, L_3. Then, there exists $C(\Gamma) \in \mathbb{N}$ such that for all $f \in \mathrm{Pol}(\Gamma)$, f is $C(\Gamma)$-fixing.*

Proof. Fix $f \in \mathrm{Pol}(\Gamma)$ of arity f. Pick a promise relation (P,Q) of arity k as guaranteed by Lemma 4.7 such that $f \in \mathrm{Pol}(P,Q)$ for all $f \in \mathrm{Pol}(\Gamma)$.

Case 1, $k \ge 3$ odd, $P = \mathrm{Ham}_k(\{(k+1)/2\}), Q = \mathrm{Ham}_k(\{0,\ldots,k-1\})$.

This case builds on techniques from Lemmas 4.2 and 5.4 of [1].

Let $B \subseteq \{1,\ldots,L\}$ be the set of coordinates for which f is *somewhere-increasing*. That is, $B = \{i \in \{1,\ldots,L\} : \exists S \subseteq \{1,\ldots,L\}, f(e_{S\setminus i}) = 0, f(e_S) = 1\}$.

We claim that $f(e_S) = 1$, for all $S \supseteq B$. Otherwise if $f(e_S) = 0$, consider a sequence of subsets $S = S_0 \subset S_1 \subset \cdots \subset S_\ell = [L]$ such that $S_{i+1} \setminus S_i$ is always a singleton. Since $f(e_{[L]}) = 1$, there exists at least one i such that $f(S_i) = 0$ and $f(S_{i+1}) = 1$, but the unique $j \in S_{i+1} \setminus S_i$ must be an element of B, contradiction.

Therefore, f is $|B|$-fixing. Thus, if we deduce that $|B|$ is bounded by some C for all f, then we know that all polymorphisms of Γ are C-fixing.

Let $a = (k-1)/2 \ge 1$. If $|B| < a$ then we are done. Otherwise, we claim that for every subset $S \subseteq B$ of size a, we have that $f(e_S) = 1$. Let $S = \{i_1,\ldots,i_a\}$, and let $x^1, y^1, \ldots, x^a, y^a$ be witnesses for $i_1,\ldots,i_a \in B$. That is, $f(x^j) = 0, f(y^j) = 1, x^j_{i_j} = 0, y^j_{i_j} = 1$, and x^j and y^j are identical in all other coordinates. Consider now the k tuples

$$(\neg x^1, y^1, \ldots, \neg x^a, y^a, \neg e_S)$$

It is easy to verify that in each coordinate $i \in \{1,\ldots,L\}$, exactly $a + 1 = (k+1)/2$ of these tuples have their ith coordinate equal to 1. Thus, since $f \in \mathrm{Pol}(P,Q)$, we have that not all of $f(\neg x^1), f(y^1), \ldots, f(\neq x^a), f(y^a), f(\neg e_S)$ are equal to 1. Thus, since the first $2a$ are equal to 1, we have that $f(\neg e_S) = 0$, so $f(e_S) = 1$, as desired.

It is easy now to see that $|B| < (c(\Gamma) + 1)a$, else we may construct disjoint $S_1,\ldots,S_{c(\Gamma)+1} \subseteq B$ of size equal to a, so $f(e_{S_1}), \ldots, f(e_{S_{c(\Gamma)+1}})$, violating Corollary 4.11. Thus, $|B|$ is bounded, so all f are C-fixing for some $C(\Gamma)$ independent of f.

Case 2, $k \ge 3$, $P = \mathrm{Ham}_k(\{1,k\}), Q = \mathrm{Ham}_k(\{0,\ldots,k\} \setminus \{b\}), b \in \{2,\ldots,k-1\}$.

Call $S \subseteq \{1,\ldots,L\}$ minimal if $f(e_S) = 1$ but $f(e_{S'}) = 0$ for all $S' \subset S$. We claim that if S is minimal, then $|S| < b$. Assume for contradiction that S is minimal but $|S| \ge b$. Thus, we may find nonempty disjoint $S_1 \cup \cdots \cup S_b = S$. For each i, note that $f(e_{S\setminus S_i}) = 0$, so $f(e_{([L]\setminus S)\cup S_i}) = 1$ by folding. Furthermore, $f(e_{[L]\setminus S}) = 0$. Thus, consider the k-tuple

$$(e_{([L]\setminus S)\cup S_1}, \ldots, e_{([L]\setminus S)\cup S_b}, e_{[L]\setminus S}, \ldots, e_{[L]\setminus S}).$$

where $e_{[L]\setminus S}$ appears $k - b$ times. It is easy to see that if $i \in S$, then the ith coordinate is equal to 1 in exactly one element of this k-tuple, otherwise the ith coordinate is equal to 1 in every k-tuple. Thus, the ith

coordinates belong to P for all $i \in [L]$. Since, $f \in \mathrm{Pol}(P, Q)$, we then have that

$$(f(e_{([L]\setminus S)\cup S_1}), \dots, f(e_{([L]\setminus S)\cup S_b}), f(e_{[L]\setminus S}), \dots, f(e_{[L]\setminus S})) \in Q.$$

But, the k-tuple has Hamming weight b, a contradiction. Thus, every minimal set has size strictly less than b. Construct a sequence of subsets of $[L]$ as follows. Let T_1 be a minimum sized subset of f such that $f(e_{T_1}) = 1$. Let T_2 be a minimum sized subset, if it exists of f disjoint from T_1 such that $f(e_{T_2}) = 1$, and so forth. Note that each of this subsets is minimal. If we can construct $T_{c(|\Gamma|)+1}$, then $T_1, \dots, T_{c(|Gamma|)+1}$ violate Corollary 4.11. Otherwise, there is some $i < c(|\Gamma|)+1$ such that setting the coordinates of $T_1 \cup \dots \cup T_i$ to 0 fixes f, and so f is $c(|\Gamma|)(b-1)$-fixing. Either way we are done.[13] $\qquad\square$

4.3 Non-idempotent case

Now, assume that our folded, symmetric family Γ of promise relations has non-idempotent polymorphisms. If any polymorphism f has the property that $f(0, \dots, 0) = f(1, \dots, 1)$, then folding is violated. Thus, Γ is non-degenerate, so we may apply Lemma 2.13 to yield that every polymorphism of Γ is a polymorphism of the idempotent family Γ' or it is the negation of a polymorphism of the idempotent family Γ''. Thus, if Γ avoids Parity, Majority, Alternating-Threshold, as well as their antis, then Γ' and Γ'' both avoid Parity, Majority, and Alternating-Threshold. By the previous section, the polymorphisms of Γ' and Γ'' are C-fixing for some sufficiently large C. Since negating a folded polymorphism does not change that it is C-fixing, we have shown the following.

Theorem 4.13. *Let Γ be a finite, folded, symmetric family of promise relations. Assume there exist odd $L_1, \dots, L_6$ such that Par_{L_1}, AT_{L_2}, Maj_{L_3}, $\overline{\mathrm{Par}}_{L_4}$, $\overline{\mathrm{AT}}_{L_5}$, and $\overline{\mathrm{Maj}}_{L_6}$ are not polymorphisms of Γ. Then there exists $C(\Gamma)$ such that all polymorphisms of Γ are C-fixing.*

5 Hardness Arguments

Now that we have a rather strong classification of polymorphisms for folded, symmetric PCSPs, we are in a good position to interface it with a reduction from Label Cover to actually demonstrate NP-hardness.

Definition 5.1. An instance of Label Cover is based on a bipartite graph $G = (U, V, E)$. Each edge $e = (u, v)$ is associated with a projection $\pi_e : [R] \to [L]$ for some positive integers R and L. A labeling is a pair of maps $\sigma_V : V \to [R]$, $\sigma_U : U \to [L]$. A labeling *satisfies* the instance if for all $(u, v) \in E$, $\pi_{(u,v)}(\sigma_V(v)) = \sigma_U(u)$.

The PCP theorem combined with parallel repetition gives the following well-known hardness of Label Cover which is the starting point for most inapproximability results.

Proposition 5.2. *For any $\eta > 0$, given an instance of Label Cover it is NP-hard to distinguish between the two cases:*

- *Completeness: There exists a labeling σ_V, σ_U that satisfies every edge.*

- *Soundness: No labeling σ_V, σ_U can satisfy a fraction η of the edges.*

Theorem 5.3. *Let Γ be a folded, finite family of promise relations. Suppose that there exists a universal constant $C = C(\Gamma) < \infty$ such that every polymorphism of Γ is $C(\Gamma)$-fixing. Then $\mathrm{PCSP}(\Gamma)$ is NP-hard.*

---

[13]This proof idea originated in a subsequent paper [31].

Proof. The proof is via reduction from the hardness of Label Cover as stated in Proposition 5.2, for the parameter $\eta = 1/C^2$. The proof is a simplification of the proof of Theorem 1.1 of [1].

Let $G = (U, V, E)$ be our instance with maps $\pi_e : [R] \to [L]$. As noted in Remark 4.7 of [1], L and R are functions of η and thus are independent of the size of G. We now create a Γ-PCSP $\Psi = (\Psi_P, \Psi_Q)$. For each $u \in U$, identify the vertex with 2^L variables which we denote by $f_u(x)$ where $x \in \{0,1\}^L$ and $f_u : \{0,1\}^L \to \{0,1\}$. For all $(P, Q) \in \Gamma$ and $x^1, \ldots, x^L \in P$ (possibly with repetition) we enforce the constraint
$$P(f_u(x_1^1, \ldots, x_1^L), \ldots, f_u(x_k^1, \ldots, x_k^L))$$
in Ψ_P, with the corresponding constraint in Ψ_Q. From the perspective of Ψ_Q, f_u is a polymorphism of Γ. Likewise, for each $v \in V$, identify 2^R variables which we denote by $f_v(y)$ where $y \in \{0,1\}^R$ and $f_v : \{0,1\}^R \to \{0,1\}$. Again, using the constraints of Γ, we may specify that f_v is a polymorphism from the perspective of Ψ_Q.

Next, we specify the edge constraints, which we do in a manner greatly simplifying that of [1]. For each $e = (u, v) \in E$ and for any $x \in \{0,1\}^L$ and $y \in \{0,1\}^R$ such that $x_{\pi_e(i)} = y_i$ for all $i \in [R]$, we specify that $f_u(x) = f_v(y)$. Note that Γ might not have an equality constraint, but we can implicitly introduce one by using the same variable for $f_u(x)$ and $f_v(y)$ when constructing Ψ. For a specific $(u, v) \in E$, these constraints maintain that $f_v^{\pi_{(u,v)}} = f_u$ (in both Ψ_P and Ψ_Q).

To show that this is a valid reduction, we need to show that both completeness and soundness hold (see Lemmas 4.5 and 4.6 of [1]).

- Completeness: If there exists a labeling σ_U, σ_V satisfying every edge of the Label Cover instance, let $f_u(x) = x_{\sigma_U(u)}$ and $f_v(x) = x_{\sigma_V(v)}$. These satisfy the constraints for Ψ_P since dictators are polymorphisms of (P, P) (as well as (P, Q)) for all $(P, Q) \in \Gamma$. The equal constraints are also satisfied since if $x_{\pi_e(i)} = y_i$ for some $e = (u, v) \in E$, then $f_u(x) = x_{\sigma_U(u)} = y_{\sigma_V(v)} = f_v(y)$, as desired. Thus, Ψ_P is satisfiable when our Label Cover instance is satisfiable.

- Soundness: Assume for the sake of contradiction, that a satisfying assignment to Ψ_Q exists. For each $u \in U$, $v \in V$, f_u and f_v are C-fixing. Thus, we may define $S_u \subseteq [L], S_v \subseteq [R]$ such that $f_u(x) = f_u(1, \ldots, 1)$ and $f_v(y) = f_v(1, \ldots, 1)$ if $x_i = 1$ for $i \in S_u$ and $y_j = 1$ for $j \in S_v$. Since S_u is C-fixing, we can let $|S_u| \leq C$. By the same logic, $|S_v| \leq C$.

  We claim that for every edge $e = (u, v) \in E$, $S_u \cap \pi_e[S_v]$ is nonempty (where $\pi_e[S_v] = \{\pi(s) : s \in S_v\}$). By virtue of the equality constraints, $f_u = f_v^{\pi_e}$; thus we have that $f_u(e_{\pi[S_v]}) = f_v(e_{S_v}) = f_v(1, \ldots, 1) = f_u(1, \ldots, 1)$. Thus, as Γ is folded, f_u must be folded, so $f_u(e_{[L]\setminus\pi_e[S_v]}) = \neg f_u(1, \ldots, 1)$. If S_u and $\pi_e[S_v]$ were disjoint, then $S_u \subseteq [L] \setminus \pi_e[S_v]$, so by the definition of S_u, $f_u(e_{[L]\setminus\pi_e[S_v]}) = f_u(1, \ldots, 1)$, contradiction. Thus, S_u and $\pi_e[S_v]$ intersect non-trivially.

  Due to this fact, we can show a η-approximate labeling exists for our label cover instance as in [1] and typical for Label Cover reductions. For each $u \in U$, select $\sigma_U(u)$ uniformly at random from S_u. Likewise, for each $v \in V$, select $\sigma_V(v)$ uniformly at random from S_v. Since for any given $e = (u, v) \in E$ we have that S_u and $\pi_e[S_v]$ have a common intersection and both sets have size at most C, $\sigma_U(u) = \pi_e(\sigma_V(v))$ with probability at least $\eta \geq 1/C^2$. Thus, the expected number of constraints satisfied by a random labeling is at least η. Hence, there exists a labeling which satisfies at least η-fraction of the constraints, as desired.

Thus, we have completed our reduction, so $\mathrm{PCSP}(\Gamma)$ is NP-hard. $\qquad\square$

Hence, we have completed the proof of Theorem 2.16 by combining Theorems 3.2, 4.13, and 5.3.

6 General Theory of Promise CSPs

This section contains a number of additional observations about Promise CSPs which help to build a broader theory of Promise CSPs.

6.1 Galois Correspondence of Polymorphisms

In this part, we show that for any finite family Γ of promise relations of any finite arity, we show that $\mathrm{Pol}(\Gamma)$ captures the computational complexity of $\mathrm{PCSP}(\Gamma)$ in the following precise sense. To do this, we show a *Galois correspondence* between families of polymorphisms and families of promise relations that are closed under a form of reduction.

Theorem 6.1. *Let Γ and Γ' be families of promise relations such that $\mathrm{Pol}(\Gamma) \subseteq \mathrm{Pol}(\Gamma')$. Then, there is a polynomial-time reduction from $\mathrm{PCSP}(\Gamma')$ to $\mathrm{PCSP}(\Gamma)$.*

This result is the promise-analogue to Theorem 3.16 of [27], originally established by [37], which holds for traditional CSPs. Our proof has similar structure to that of [27]. Pippenger [48] in Section 2 of his paper proves a variation of the Galois correspondence between promise relations and their polymorphisms, although not in this particular complexity-theoretic formulation.

In fact, we show the polynomial-time reduction is of a very local form. Let $\mathrm{EQUAL} = \{(i, i) : \in D\}$ be the relation which specifies that two variables are equal. Since we have been allowing repetition of variables, this relation has been essentially implicit.

Definition 6.2. Let Γ be a finite family of promise relations. We say that a promise relation $(P', Q') \in D^k \times D^k$ is *positive primitive promise definable* (shortened to *ppp-definable*) from Γ if there exists a $\Gamma \cup \{\mathrm{EQUAL}\}$-PCSP $\Psi = (\Psi_P, \Psi_Q)$ on $k + \ell$ variables such that

- For all $(x_1, \ldots, x_k) \in P'$, there exists $(y_1, \ldots, y_\ell)$ such that $(x_1, \ldots, x_k, y_1, \ldots, y_\ell)$ is a satisfying assignment to Ψ_P.

- For all satisfying assignments $(z_1, \ldots, z_{k+\ell})$ to Ψ_Q, $(z_1, \ldots, z_k) \in Q'$.

We say that a finite family of promise relations Γ' is ppp-definable from Γ if every $(P', Q') \in \Gamma'$ is ppp-definable from Γ.

In particular, note that if (P, Q) and (P', Q') have the same arity and $P' \subseteq P \subseteq Q \subseteq Q'$ then (P', Q') is ppp-definable from (P, Q) by letting $(\Psi_P, \Psi_Q) = (P, Q)$. We also note that ppp-definability is reflexive (Γ is ppp-definable from Γ) and transitive: if Γ' is ppp-definable from Γ and Γ'' is ppp-definable from Γ' then Γ'' is ppp-definable from Γ. We have that ppp-definability is a formalization of the notion of a *gadget reduction* in [1] (see Proposition 3.1).

Our notion of ppp-definability is a direct generalization of the notion of *pp-definability* for normal CSP relations defined in [27]. If Γ' is ppp-definable from Γ, there is a corresponding polynomial-time reduction from $\mathrm{PCSP}(\Gamma')$ to $\mathrm{PCSP}(\Gamma)$ by replacing each $(P', Q') \in \Gamma'$ clause with a corresponding (Ψ_P, Ψ_Q) clause (adding in any auxiliary variables), which can can be implemented with clauses from Γ and yields only a constant-factor blowup. It is straightforward to verify that this reduction is valid. As noted in [27], this reduction can be done in logarithmic space.

In establishing the Galois correspondence, one important ppp-definition from Γ is the promise relation of polymorphisms of Γ. This definition is instrumental in showing that every relaxation of PCSP is also a ppp-definition and vice versa.

Proposition 6.3. *Let L be a positive integer. The following promise relation $S_L \subseteq T_L \subseteq D^{D^L}$ is ppp-definable from Γ:*

$$S_L = \{f : D^L \to D \; : \; f \in \mathrm{Pol}(P, P) \text{ for all } (P, Q) \in \Gamma\}$$
$$T_L = \{f : D^L \to D \; : \; f \in \mathrm{Pol}(P, Q) \text{ for all } (P, Q) \in \Gamma\},$$

where we identify a function $f \in D^L \to D$ as a vector of $|D|^L$ variables.

Proof. Using the definition of a polymorphism, one can specify that f is a polymorphism of $\mathrm{Pol}(P, Q)$ of specific arity in terms of a fixed number of Q-clauses. Replacing those Q-clauses with P-clauses exactly characterizes that $f \in \mathrm{Pol}(P, P)$. □

With these facts established, we may now prove the theorem. The proof is quite similar to and was inspired by the second half of Theorem 3.13 of [27].

Proof of Theorem 6.1. It suffices to show that every promise relation $(P', Q') \in \Gamma'$ is ppp-definable from Γ. Let k be the arity of (P', Q') and let $m = |P'|$. Let $x^1, \ldots, x^m$ be some ordering of the elements of P'. Define $y^1, \ldots, y^k \in D^m$ such that $y^i_j = x^j_i$ for all $i \in [k], j \in [m]$. Now from Proposition 6.3, we have that (S_m, T_m) is ppp-definable from Γ. Now, consider the following promise relation (S'_m, T'_m) of arity k.

$$S'_m = \{(f(y^1), \ldots, f(y^k)) \; : \; f \in S_m\}$$
$$T'_m = \{(f(y^1), \ldots, f(y^k)) \; : \; f \in T_m\}.$$

We have that (S'_m, T'_m) is ppp-definable from (S_m, T_m) since every $x \in S'_m$ can be built up into a corresponding element of S_m and every $y \in T_m$ can be stripped down to an element of T'_m. Note that this is the case even if $y^i = y^j$ for some distinct $i, j \in [k]$ by using the EQUAL relation.

We claim that $P' \subseteq S'_m \subseteq T'_m \subseteq Q'$. First, for all $i \in [m]$, consider the unique projection map $\pi_i : D^m \to D$ given by $\pi_i(y) = y_i$. Clearly $\pi_i \in S_m$. Thus, $(\pi_i(y^1), \ldots, \pi_i(y^k)) = (y^1_i, \ldots, y^k_i) = x^i \in S'_m$. Thus, $P' \subseteq S'_m$. Second, we can see that $S'_m \subseteq T'_m$ since $S_m \subseteq T_m$. Third, note that $T'_m \subseteq T_m(P') \subseteq \mathrm{Pol}(\Gamma)(P')$. Since $\mathrm{Pol}(P', Q') \supseteq \mathrm{Pol}(\Gamma)$, we have that $Q' \supseteq \mathrm{Pol}(\Gamma)(P')$. Thus, $T'_m \subseteq Q'$.

Thus, therefore (P', Q') is ppp-definable from (S'_m, T'_m). By transitivity, we have that (P', Q') is ppp-definable from Γ, so Γ' is ppp-definable from Γ. □

6.2 Polymorphism-only description of PCSPs

In this section, we establish a necessary and sufficient set of conditions on a set $\mathcal{F}$ of functions over domain D for which there exists some finite Γ such that $\mathcal{F} = \mathrm{Pol}(\Gamma)$. Pippenger [48] proved such a characterization in the case that Γ may or may not *have infinitely many relations* using similar ideas. Recall the definition of a projection of a polymorphism.

Definition 6.4. Let $f : D^L \to D$ be a function, and let $\pi : [L] \to [R]$ be any map which we call a *projection*. The projection of f with respect to π is the function $f^\pi : D^R \to D$ such that for all $y \in D^R$, $f^\pi(y) = f(x)$, where $x \in D^L$ is the unique L-tuple such that

$$x_i = y_{\pi(i)}, \text{ for all } i \in [L].$$

Note that in a projection it might be the case that $R \geq L$. We say that a family $\mathcal{F}$ of functions over domain D is *projection-closed* if for all $L, R \in \mathbb{N}$, all $f \in \mathcal{F}$ of arity L, and all maps $\pi : [L] \to [R]$, $f^\pi \in \mathcal{F}$.

Another technical property we require of $\mathcal{F}$ is that it is *finitizable*. This means there exists some finite arity $R \in \mathbb{N}$, called the *finitized arity* such that $f : D^L \to D$ is an element of $\mathcal{F}$ if and only if for all $\pi : [L] \to [R]$, the projection f^π is an element of $\mathcal{F}$. Intuitively, this finitization property says that some finite arity of $\mathcal{F}$ captures all of the meaningful information about what is contained in $\mathcal{F}$. This is directly analogous to the property that our set Γ of promise relations is finite.

Surprisingly, these two properties–that $\mathcal{F}$ is projection-closed and finitizable–perfectly capture the families of the form $\mathrm{Pol}(\Gamma)$ for some Γ as long as we stipulate that $\mathcal{F}$ contains the *identity function*: $\mathrm{id}_D : D \to D$ such that $\mathrm{id}_D(x) = x$ for all $x \in D$.[14]

Theorem 6.5. *Let $\mathcal{F}$ be a family of functions over domain D. Then, there exists a finite family Γ of promise relations such that $\mathcal{F} = \mathrm{Pol}(\Gamma)$ if and only if $\mathcal{F}$ is both projection-closed and finitizable and $\mathrm{id}_D \in \mathcal{F}$.*

We start by showing that these two properties are necessary.

Claim 6.6. *Let $\Gamma = \{(P_i, Q_i) : P_i \subseteq Q_i \subseteq D^{k_i}\}$ be a finite family of promise relations with domain D. Then, $\mathrm{Pol}(\Gamma)$ is both projection-closed and finitizable.*

Proof. projection-closed: Let $f : D^L \to D$ be a polymorphism of Γ and let $\pi : [L] \to [R]$ be a map. We claim that $f^\pi : D^R \to D$ is also a polymorphism of Γ. Consider all (P_i, Q_i) and $y^{(1)}, \ldots, y^{(R)} \in P_i$. We need to show that $f^\pi(y^{(1)}, \ldots, y^{(R)}) \in Q_i$. Consider $x^{(1)}, \ldots, x^{(L)} \in P_i$ such that $x^{(j)} = y^{(\pi(j))}$ for all $j \in [L]$. From the definition of f^π it is then easy to see that

$$f^\pi(y^{(1)}, \ldots, y^{(R)}) = f(x^{(1)}, \ldots, x^{(L)}) \in Q_i,$$

as desired.

finitizable: Let $R = \max_{(P_i, Q_i) \in \Gamma} |P_i|$. Crucially, this maximum exists since Γ is finite. Since $\mathrm{Pol}(\Gamma)$ is projection-closed, for all $f \in \mathrm{Pol}(\Gamma)$ of arity L and all $\pi : [L] \to [R]$, we have that $f^\pi \in \mathrm{Pol}(\Gamma)$.

Now, consider any $f \notin \mathrm{Pol}(\Gamma)$ of arity L, we would like to show that there exists $\pi : [L] \to [R]$ such that $f^\pi \notin \mathrm{Pol}(\Gamma)$. Since $f \notin \mathrm{Pol}(\Gamma)$, there exists $(P_i, Q_i) \in \Gamma$ and $x^{(1)}, \ldots, x^{(L)} \in P_i$ such that $f(x^{(1)}, \ldots, x^{(L)}) \notin Q_i$. Since $R \geq |P_i|$, there exists an injective map $\sigma : P_i \to [R]$. Let $\pi : [L] \to [R]$ be $\pi(i) = \sigma(x^{(i)})$. By nature of π, we can select $y^{(1)}, \ldots, y^{(R)} \in P_i$ such that $y^{(r)} = x^{(\pi^{-1}(r))}$ for all $r \in \mathfrak{I}(\pi)$ and $y^{(r)} = 1$ otherwise. (If $r \in R$ is not in the image of π, then we may make an arbitrary choice.) Note that if $\pi(j_1) = \pi(j_2)$ then $x^{(j_1)} = x^{(j_2)}$ so this choice of $y^{(j)}$'s is well-defined. From the definition of a projection,

$$f^\pi(y^{(1)}, \ldots, y^{(R)}) = f(x^{(1)}, \ldots, x^{(L)}) \notin Q_i,$$

as desired. Therefore $\mathrm{Pol}(\Gamma)$ is finitizable. □

Note that since we stipulate that $P \subseteq Q$ for all $(P, Q) \in \Gamma$, we immediately have that id_D is a polymorphism of Γ. Much more difficultly, we show that these two properties are sufficient.

Lemma 6.7. *Let $\mathcal{F}$ be a domain-D family of functions which is both projection-closed and finitizable as well as has id_D as an element. Then, there exists a family Γ of finitely many promise relations such that $\mathrm{Pol}(\Gamma) = \mathcal{F}$.*

Proof. Let $R \in \mathbb{N}$ be the finitized arity of $\mathcal{F}$. Identify the integers of $[|D|^R]$ with elements of D^R. Our choice of Γ will consist of a single promise relation $P \subseteq Q \subseteq D^{|D|^R}$, where each $f \in D^{[|D|^R]}$ will be identified with a function $f : D^R \to D$ in the canonical way. We let $f \in P$ if and only if there exists $j \in [R]$ such that

---

[14]If we broaden our definition of PCSPs (as mentioned in the introduction) so that instead of $P \subseteq Q$, there is some unary map $\sigma : D_1 \to D_2$ such that $\sigma(P) \subseteq Q$, then the condition $\mathrm{id}_D \in \mathcal{F}$ can be replaced with $\sigma \in \mathcal{F}$ for some unary function σ.

$f(x) = x_j$ for all $x \in D^R$. We let $Q = \{f \in \mathcal{F} \mid f \text{ has arity } R\}$. Since $\mathcal{F}$ has the identity function and is projection-closed, we have that $P \subseteq Q$. Thus, Γ is a finite promise relation.

Now that we have constructed Γ, we need to show $\mathrm{Pol}(\Gamma) = \mathcal{F}$. Enumerate the elements of P as $y^{(1)}, \ldots, y^{(R)}$, where $y^{(j)}(x) = x_j$ for all $j \in [R]$ and $x \in D^R$. With this enumeration, we have the property that for all $g \in \mathrm{Pol}(P, Q)$ of arity R, $g(y^{(1)}, \ldots, y^{(R)}) = g$. This is because for all $x \in D^R$,

$$g(y^{(1)}(x), \ldots, y^{(R)}(x)) = g(x_1, \ldots, x_R) = g(x).$$

(Thus the $y^{(i)}$'s are like a long code test.)

First, we show that $\mathrm{Pol}(P, Q) \subseteq \mathcal{F}$. Consider any $f \in \mathrm{Pol}(P, Q)$ of arity L as well as any $\pi : [L] \to [R]$. Pick $x^{(1)}, \ldots, x^{(L)} \in P$ such that $x^{(j)} = y^{(\pi(j))}$ for all $j \in [L]$. Thus,

$$Q \ni f(x^{(1)}, \ldots, x^{(L)}) = f^\pi(y^{(1)}, \ldots, y^{(R)}) = f^\pi.$$

Thus, $f^\pi \in Q \subseteq \mathcal{F}$ for all $\pi : [L] \to [R]$. Thus, $f \in \mathcal{F}$ since $\mathcal{F}$ is finitizable, as desired.

Last, we show that $\mathcal{F} \subseteq \mathrm{Pol}(P, Q)$. For every $f \in \mathcal{F}$, we need to show that for all $x^{(1)}, \ldots, x^{(L)} \in P$, we have that $f(x^{(1)}, \ldots, x^{(L)}) \in Q$. Since $y^{(1)}, \ldots, y^{(R)}$ is an enumeration of the elements of P, there is a unique $\pi : [L] \to [R]$ such that $x^{(j)} = y^{(\pi(j))}$ for all $j \in [L]$. Then, we have that

$$f(x^{(1)}, \ldots, x^{(L)}) = f^\pi(y^{(1)}, \ldots, y^{(R)}) = f^\pi.$$

Since $\mathcal{F}$ is projection closed, $f^\pi \subseteq \mathcal{F}$. Therefore, $f^\pi \in Q$ because f^π has arity R. Thus, $f(x^{(1)}, \ldots, x^{(L)}) \in Q$, as desired.

Hence, $\mathcal{F} = \mathrm{Pol}(P, Q)$. □

Claim 6.6 and Lemma 6.7 together establish Theorem 6.5.

6.3 Analogous characterization for CSPs

We now extend Theorem 6.5 to show that the same characterization holds for CSPs as long as we add the condition that our set of functions form a *clone* (defined below). For our purposes, a CSP is a PCSP Γ in which $P = Q$ for all $(P, Q) \in \Gamma$.

As known in the CSP literature (e.g., [27]), the family of polymorphisms of a CSP Λ have the additional property it is a *clone*.[15] That is, for all $f \in \mathrm{Pol}(\Lambda)$ of arity L_1, and all $g_1, \ldots, g_{L_1} \in \mathrm{Pol}(\Lambda)$ of arity L_2, we have that $h(x^{(1)}, \ldots, x^{(L_1)}) := f(g_1(x^{(1)}), \ldots, g_{L_1}(x^{(L_1)}))$ is a polymorphism of Λ of arity $L_1 L_2$. It turns out this property is necessary and sufficient for characterizing CSPs from their polymorphisms.

Lemma 6.8. *Let $\mathcal{F}$ be a family of functions over the domain D. Then, there exists a CSP Λ such that $\mathcal{F} = \mathrm{Pol}(\Lambda)$ if and only if $\mathcal{F}$ is finitizable, a clone, and contains the identity.*

Proof. As stated previously, $\mathrm{Pol}(\Lambda)$ is finitizable, a clone, and contains the identity. Thus, it suffices to show the converse.

Assume that $\mathcal{F}$ finitizes at arity R. As shown in Lemma 6.7, $\mathcal{F} = \mathrm{Pol}(P, Q)$, where $P \subseteq Q \subseteq D^{D^R}$. In this case, P are the R projection functions from D^R to D and Q is the set of arity-R functions of $\mathcal{F}$. Since, we now have that $\mathcal{F}$ is a clone, we claim that $\mathcal{F} = \mathrm{Pol}(Q, Q)$.

First, we have that $\mathrm{Pol}(Q, Q) \subseteq \mathrm{Pol}(P, Q) = \mathcal{F}$ since $P \subseteq Q$, so membership in $\mathrm{Pol}(Q, Q)$ is a more strict condition. To show the reverse inclusion $\mathcal{F} \subseteq \mathrm{Pol}(Q, Q)$, consider any $f \in \mathcal{F}$ of arity L. We need to show for all $g_1, \ldots, g_L \in Q$, we have that $f(g_1, \ldots, g_L) \in Q$. Since $\mathcal{F}$ is a clone, we immediately have that $f(g_1, \ldots, g_L) \in \mathcal{F}$. Furthermore, $f(g_1, \ldots, g_L)$ has arity R so $f(g_1, \ldots, g_L) \in Q$. □

---

[15] As pointed out to the authors after writing this, the collection of clones which correspond to $\mathrm{CSP}(\Gamma)$ for finite Γ are "finitely related clones" (see, e.g., [45].)

6.4 Significance toward establishing complexity of PCSPs

These results liberate us from ever thinking about Γ, and instead we can think entirely in terms of establishing the easiness/hardness of projection-closed, finitized families of functions. As discussed in the introduction, subsequent work (e.g., [2]) has shown that it suffices to only consider the *identities* which the polymorphisms satisfy. Another point of consideration is the case in which there are *infinitely* many relations in our (P)CSPs (although keeping a finite domain). As a computational problem, one can define the (non-uniform) computational complexity of a PCSP Γ, in the style of the compactness theorem, to be the supremum of the computational complexities of all finite subsets $\Gamma' \subseteq \Gamma$. Another common (uniform) definition is that the relations used in any particular CSP are encoded as part of the input (using some canonical encoding). The *local-global conjecture* (e.g., [8]) states that these two notions of complexity should be identical for infinite case. Such a conjecture could also be made for PCSPs, although we doubt the veracity of such a claim for the following reason. Once we allow infinitely many relations into our PCSPs, the possible characterizations of polymorphisms expand to all projection-closed families (that is, the finitization condition can be dropped). As a result, it seems quite tempting that an NP-intermediate PCSP could be constructed by adapting the techniques used to prove Ladner's theorem [44].

6.5 Every PCSP is Equivalent to a Promise Digraph Homomorphism.

We show that for any finite $\mathrm{PCSP}(\Gamma)$ over a domain D, there exist directed graphs H_Γ, H_Γ' with a homomorphism from H_Γ to H_Γ' such that $\mathrm{PCSP}(\Gamma)$ is polynomial time equivalent to the promise digraph homomorphism problem on (H_Γ, H_Γ'). Below is a formal definition of the promise digraph problem.

Definition 6.9. Let H, H' be a pair of directed graphs such there is a homomorphism $\phi : H \to H'$. The promise decision problem $\mathrm{PDGH}(H, H')$ (promise directed graph homomorphism) is given a directed graph G distinguish between the two possible cases.

YES. There is a homomorphism from G to H.

NO. There is no homomorphism from G to H'.

Theorem 6.10. *Let D be a finite domain, and let Γ be any promise relation over D. Then, there exists a pair of directed graphs H_Γ and H_Γ' such that $H_\Gamma \subseteq H_\Gamma'$ and $\mathrm{PCSP}(\Gamma)$ is polynomial-time equivalent to $\mathrm{PDGH}(H_\Gamma, H_\Gamma')$.*

Our proof is an adaptation of corresponding result for CSPs by Feder and Vardi, see Theorem 11 of [30]. We cite the following slightly stronger claim which is implied in the proof of Theorem 11. We let $\mathrm{Hom}(H)$ denote the digraph graph homomorphism problem of determining whether there exists a homomorphism from an input graph G to H. To do this, we need to define a notion of an "oblivious reduction."

Definition 6.11. Let $\Gamma = \{R_i\}$ be a CSP over a domain D. Let Ψ be an instance of $\mathrm{CSP}(\Gamma)$, the *underlying colored directed hypergraph* of Ψ is the hypergraph whose vertices are the variables of Ψ and whose (directed) hyperedges correspond to the tuples on which the relations of Γ are applied, with the hyperedge having color i when R_i is applied.

Definition 6.12. Let Γ be a CSP over a domain D and let H be a directed graph dependent on D.

A reduction $\Psi \mapsto G_\Psi$ from $\mathrm{CSP}(\Gamma)$ to $\mathrm{Hom}(H)$ is *oblivious* if $G_\Psi = \sigma(D, X)$ for some function σ independent of Γ applied to the domain D and underlying colored directed hypergraph X of Ψ.

Conversely, a reduction $G \mapsto \Psi_G$ from $\mathrm{Hom}(H)$ to $\mathrm{CSP}(\Gamma)$ is *oblivious* if $\Psi_G = \eta(D, G)$ for some function η independent of H.

Theorem 6.13 (Feder and Vardi, [30])**.** *Let Γ be a CSP over a domain D such that $\Gamma = (R, S)$, where R has arity 1 and S has arity 2 with the property that the projections of the first and second coordinate of S are both all of D. Then, there is a digraph H_Γ such that $\mathrm{CSP}(\Gamma)$ is equivalent to the graph homomorphism problem on H_Γ. Furthermore, the reductions mapping instances of $\mathrm{CSP}(\Gamma)$ to those of $\mathrm{Hom}(H_\Gamma)$ and the reductions mapping instances of $\mathrm{Hom}(H_\Gamma)$ to $\mathrm{CSP}(\Gamma)$ are both oblivious in the sense of Definition 6.12.*

Below we sketch the details necessary to extend this argument to PCSPs.

Theorem 6.14. *Let Γ be a PCSP, then there exists a pair of digraphs H and H' such there is a homomorphism from H to H' and $\mathrm{PCSP}(\Gamma)$ is polynomial time equivalent to $\mathrm{PDGH}(H, H')$.*

Proof. The proof technique borrows significantly from [30].

Assume that our PCSP is $\Gamma = \{(P_i, Q_i) \colon i \in [r]\}$. Let $P' = \prod_i P_i$ and $Q' = \prod_i Q_i$. Let $\Gamma' = \{(P', Q')\}$. It is easy to check that $\mathrm{PCSP}(\Gamma)$ is polynomial time equivalent to $\mathrm{PCSP}(\Gamma')$.

Assume the single relation of Γ' has arity k. Now, view Γ' as a CSP over the domain $D' = D^k$. Note that $\mathrm{PCSP}(\Gamma')$ over the domain D' is not necessarily equivalent to $\mathrm{PCSP}(\Gamma')$ over the domain D, since we lose the ability to specify the same variable in different coordinates. To mitigate this, we add an additional "shift operator" $S = \{(x, y) : x, y \in D^k; x_i = y_{i+1}, i \in \{1, \ldots, k-1\}\}$. Then, we have that $\mathrm{PCSP}(\Gamma' \cup \{(S, S)\})$ over domain D' is polynomial-time equivalent $\mathrm{PCSP}(\Gamma')$ over domain D. See [30] for more details.

Then, applying Theorem 6.13, to $\{P', S\}$ and $\{Q', S\}$ over domain D', we obtain digraphs $H_{P'}$ and $H_{Q'}$ such that $\mathrm{CSP}(\{P', S\})$ is polynomial-time equivalent to $\mathrm{Hom}(H_{P'})$ and $\mathrm{CSP}(\{Q', S\})$ is polynomial-time equivalent to $\mathrm{Hom}(H_{Q'})$. From the nature of the construction of $H_{P'}$ and $H_{Q'}$ in [30], we know that $H_{P'} \subseteq H_{Q'}$, which trivially implies there is homomorphism between them.

Since the reductions to $\mathrm{Hom}(H_{P'})$ and $\mathrm{Hom}(H_{Q'})$ are oblivious, any instance Ψ of $\mathrm{PCSP}(\Gamma' \cup \{(S, S)\})$ reduces to a digraph G_Ψ such that there is a homomorphism from G_Ψ to $H_{P'}$ if and only if Ψ_P is satisfiable. Likewise, there is a homomorphism from G_Ψ to $H_{Q'}$ if and only if Ψ_Q is satisfiable. Thus, any instance Ψ of $\mathrm{PCSP}(\Gamma' \cup \{(S, S)\})$ reduces to an instance G_Ψ of $\mathrm{PDGH}(H_{P'}, H_{Q'})$.

Similarly, the reductions from $\mathrm{Hom}(H_{P'})$ and $\mathrm{Hom}(H_{Q'})$ are oblivious, so any instance G of $\mathrm{PDGH}(H_{P'}, H_{Q'})$ reduces to an instance Ψ_G of $\mathrm{PCSP}(\Gamma' \cup \{(S, S)\})$. $\qquad\square$

Remark. The theorem also holds when the notion of PCSP is extended to when the P_i's and Q_i's are related by a homomorphism instead of inclusion.

Remark. For CSPs that it is known that any relation Γ there is a digraph H_Γ such that $\mathrm{CSP}(\Gamma)$ is *logspace*-equivalent to $\mathrm{Hom}(H_\Gamma)$ [25]. The authors conjecture that this is also the case for PCSPs.

6.6 Lack of Repetition Does Not Make Things Harder

For a set Γ of promise relations, let $\mathrm{PCSP}_R(\Gamma)$ be the promise decision problem analogous to $\mathrm{PCSP}(\Gamma)$ except that each clause has at most one copy of each variable. We show that the two problems are polynomial-time equivalent using a simple combinatorial argument, simplifying the argument used in [1] for establishing the NP-hardness of "balanced 2-coloring" versus "weak 2-coloring" of $2k + 1$-uniform hypergraphs.

Theorem 6.15. *For all finite $\Gamma = \{(P_i, Q_i) \in D^{k_i} \times D^{k_i}\}$, $\mathrm{PCSP}_R(\Gamma)$ is polynomial-time equivalent to $\mathrm{PCSP}(\Gamma)$*

Proof. $\mathrm{PCSP}_R(\Gamma)$ trivially reduces to $\mathrm{PCSP}(\Gamma)$ since any instance of $\mathrm{PCSP}_R(\Gamma)$ is an instance of $\mathrm{PCSP}(\Gamma)$. Thus, we now consider the harder case. Let $\Psi = (\Psi_P, \Psi_Q)$ be a $\Gamma - \mathrm{PCSP}$ with m clauses on the variable set $x_1, \ldots, x_n$. Let k be the maximum arity of any promise relation of Γ (this is a constant). For our reduction,

replace each variable x_i with $|D|k$ 'copies' $x_i^{(1)}, \ldots, x_i^{(|D|k)}$. Replace each clause $P_i(x_{j_1}, \ldots, x_{j_{k_i}})$ of Ψ_P with a conjunction of at most $(|D|k)^{k_i}$ clauses, $P_i(x_{j_1}^{(a_1)}, \ldots, x_{j_{k_i}}^{(a_{k_i})})$ in which we remove the clauses with a repeated variable. Call this new formula Ψ_P^R. Perform an identical reduction of Ψ_Q to Ψ_Q^R. We can see that $\Psi^R = (\Psi_P^R, \Psi_Q^R)$ is a valid Γ-PCSP without repetition and the size of this PCSP is only a constant factor larger than the size of Ψ.

Now we show that this is a valid reduction. First, if Ψ is satisfiable, there is an assignment to the variables $x_1, \ldots, x_n$ which satisfies Ψ_P. If we let each copy $x_i^{(j)}$ have the same value of x_i, then we yield a satisfying assignment of Ψ_P^R. It suffices then to show that if Ψ is unsatisfiable, then Ψ^R is unsatisfiable. This is equivalent to showing that if Ψ_Q^R is satisfiable, then Ψ_Q is also satisfiable. Assume we have a satisfying assignment of Ψ_Q^R. For each of the variables x_i of Ψ_Q, set x_i to be the most frequently occurring value in the multiset $\{x_i^{(j)} : j \in \{1, \ldots, |D|k\}\}$ (break ties arbitrarily). Crucially note that this most frequently occurring value occurs at least k times. We claim that this choice of the x_i satisfies Ψ_Q. For each clause $Q_i(x_{j_1}, \ldots, x_{j_{k_i}})$, we can find a corresponding clause $Q_i(x_{j_1}^{(a_1)}, \ldots, x_{j_{k_i}}^{(a_{k_i})})$ in Ψ_Q^R such that $x_{j_\ell}^{a_\ell} = x_j$ for all j (this is possible without repetition since there are at least k distinct choices for a_ℓ). Since Ψ_Q^R is satisfied, this particular repetition-free clause is satisfied, so the corresponding clause in Ψ_Q is satisfied. Thus, we have found a satisfying assignment for Ψ_Q. Thus, $\mathrm{PCSP}_R(\Gamma)$ and $\mathrm{PCSP}(\Gamma)$ are polynomial-time equivalent. $\square$

Acknowledgments

References

[1] Per Austrin, Venkatesan Guruswami, and Johan Håstad. $(2+\epsilon)$-Sat Is NP-hard. *SIAM J. Comput.*, 46(5):1554–1573, 2017.

[2] Libor Barto, Jakub Bulín, Andrei Krokhin, and Jakub Opršal. Algebraic approach to promise constraint satisfaction. *arXiv:1811.00970 [cs, math]*, 2019.

[3] Libor Barto and Marcin Kozik. Constraint satisfaction problems of bounded width. In *Proceedings of the 50th IEEE Symposium on Foundations of Computer Science*, pages 595–603, Oct 2009.

[4] Libor Barto and Marcin Kozik. Constraint satisfaction problems solvable by local consistency methods. *Journal of the ACM*, 61(1):3:1–3:19, January 2014.

[5] Libor Barto, Marcin Kozik, and Todd Niven. The CSP Dichotomy Holds for Digraphs with No Sources and No Sinks (A Positive Answer to a Conjecture of Bang-Jensen and Hell). *SIAM Journal on Computing*, 38(5):1782–1802, 2009.

[6] Libor Barto, Andrei Krokhin, and Ross Willard. Polymorphisms, and How to Use Them. In Andrei Krokhin and Stanislav Zivny, editors, *The Constraint Satisfaction Problem: Complexity and Approximability*, volume 7 of *Dagstuhl Follow-Ups*, pages 1–44. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, Dagstuhl, Germany, 2017.

[7] Manuel Bodirsky. Complexity classification in infinite-domain constraint satisfaction. *arXiv preprint arXiv:1201.0856*, 2012.

[8] 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: 35th International Colloquium, ICALP 2008, Reykjavik, Iceland, July 7-11, 2008, Proceedings, Part II*, pages 184–196, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[9] Joshua Brakensiek and Venkatesan Guruswami. New Hardness Results for Graph and Hypergraph Colorings. In Ran Raz, editor, *31st Conference on Computational Complexity (CCC 2016)*, volume 50 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 14:1–14:27, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[10] Joshua Brakensiek and Venkatesan Guruswami. Promise Constraint Satisfaction: Structure Theory and a Symmetric Boolean Dichotomy. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2018, New Orleans, LA, USA, January 7-10, 2018*, pages 1782–1801, 2018.

[11] Joshua Brakensiek and Venkatesan Guruswami. An Algorithmic Blend of LPs and Ring Equations for Promise CSPs. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '19, pages 436–455, Philadelphia, PA, USA, 2019. Society for Industrial and Applied Mathematics.

[12] Joshua Brakensiek and Venkatesan Guruswami. Symmetric Polymorphisms and Efficient Decidability of Promise CSPs. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 297–304, 2020.

[13] Joshua Brakensiek, Venkatesan Guruswami, and Sai Sandeep. Conditional dichotomy of boolean ordered promise csps. *Electron. Colloquium Comput. Complex.*, 28:26, 2021.

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

[15] Alfred Brauer and James E Shockley. On a problem of Frobenius. *J. reine angew. Math*, 211(19):215–220, 1962.

[16] Mark Braverman, Subhash Khot, and Dor Minzer. On rich 2-to-1 games. In James R. Lee, editor, *12th Innovations in Theoretical Computer Science Conference, ITCS 2021, January 6-8, 2021, Virtual Conference*, volume 185 of *LIPIcs*, pages 27:1–27:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2021.

[17] Jonah Brown-Cohen and Prasad Raghavendra. Combinatorial optimization algorithms via polymorphisms. *CoRR*, abs/1501.01598, 2015.

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

[19] Andrei Bulatov and Víctor Dalmau. A Simple Algorithm for Mal'tsev Constraints. *SIAM J. Comput.*, 36(1):16–27, July 2006.

[20] Andrei A. Bulatov. A dichotomy theorem for constraint satisfaction problems on a 3-element set. *Journal of the ACM*, 53(1):66–120, 2006.

[21] Andrei A. Bulatov. Complexity of conservative constraint satisfaction problems. *ACM Trans. Comput. Log.*, 12(4):24, 2011.

[22] Andrei A. Bulatov. Conservative constraint satisfaction re-revisited. *J. Comput. Syst. Sci.*, 82(2):347–356, 2016.

[23] Andrei A. Bulatov, Peter Jeavons, and Andrei A. Krokhin. Classifying the Complexity of Constraints Using Finite Algebras. *SIAM J. Comput.*, 34(3):720–742, 2005.

[24] Andrei A. Bulatov, Andrei A. Krokhin, and Peter Jeavons. Constraint satisfaction problems and finite algebras. In Ugo Montanari, José D. P. Rolim, and Emo Welzl, editors, *Automata, Languages and Programming: 27th International Colloquium, ICALP 2000 Geneva, Switzerland, July 9–15, 2000 Proceedings*, pages 272–282, Berlin, Heidelberg, 2000. Springer Berlin Heidelberg.

[25] Jakub Bulín, Dejan Delic, Marcel Jackson, and Todd Niven. A finer reduction of constraint problems to digraphs. *Logical Methods in Computer Science*, 11(4:18):1–33, 2015.

[26] Jakub Bulín, Andrei Krokhin, and Jakub Opršal. Algebraic Approach to Promise Constraint Satisfaction. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, pages 602–613, New York, NY, USA, 2019. ACM.

[27] Hubie Chen. A rendezvous of logic, complexity, and algebra. *ACM Comput. Surv.*, 42(1):2:1–2:32, December 2009.

[28] Irit Dinur, Elchanan Mossel, and Oded Regev. Conditional Hardness for Approximate Coloring. *SIAM Journal on Computing*, 39(3):843–873, 2009.

[29] Irit Dinur, Oded Regev, and Clifford D. Smyth. The hardness of 3-uniform hypergraph coloring. *Combinatorica*, 25(5):519–535, 2005.

[30] 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.

[31] Miron Ficak, Marcin Kozik, Miroslav Olšák, and Szymon Stankiewicz. Dichotomy for Symmetric Boolean PCSPs. In Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi, editors, *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 fuer Informatik.

[32] Venkatesan Guruswami and Sanjeev Khanna. On the hardness of 4-coloring a 3-colorable graph. *SIAM J. Discrete Math.*, 18(1):30–40, 2004.

[33] Lucy Ham. Gap theorems for robust satisfiability: Boolean CSPs and beyond. *Theoretical Computer Science*, 676:69–91, May 2017.

[34] Lucy Ham and Marcel Jackson. All or nothing: toward a promise problem dichotomy for constraint problems. In *International Conference on Principles and Practice of Constraint Programming*, pages 139–156. Springer, 2017.

[35] Pavol Hell and Jaroslav Nešetřil. On the complexity of *H*-coloring. *Journal of Combinatorial Theory, Series B*, 48(1):92 – 110, 1990.

[36] Sangxia Huang. Improved hardness of approximating chromatic number. In Prasad Raghavendra, Sofya Raskhodnikova, Klaus Jansen, and JoséD.P. Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, volume 8096 of *Lecture Notes in Computer Science*, pages 233–243. Springer Berlin Heidelberg, 2013.

[37] Peter Jeavons. On the algebraic structure of combinatorial problems. *Theoretical Computer Science*, 200(1–2):185 – 204, 1998.

[38] 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.

[39] Sanjeev Khanna, Nathan Linial, and Shmuel Safra. On the hardness of approximating the chromatic number. *Combinatorica*, 20(3):393–415, 2000.

[40] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings on 34th Annual ACM Symposium on Theory of Computing*, pages 767–775, 2002.

[41] Vladimir Kolmogorov, Andrei Krokhin, and Michal Rolínek. The complexity of general-valued CSPs. In *Proceedings of the 56th IEEE Symposium on Foundations of Computer Science*, pages 1246–1258, Oct 2015.

[42] Andrei Krokhin and Jakub Opršal. The complexity of 3-colouring *H*-colourable graphs. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1227–1239. IEEE, 2019.

[43] Andrei Krokhin, Jakub Opršal, Marcin Wrochna, and Stanislav Živnỳ. Topology and adjunction in promise constraint satisfaction. *arXiv preprint arXiv:2003.11351*, 2020.

[44] Richard E. Ladner. On the Structure of Polynomial Time Reducibility. *Journal of the ACM*, 22(1):155–171, 1975.

[45] Petar Marković, Miklós Maróti, and Ralph McKenzie. Finitely related clones and algebras with cube terms. *Order*, 29(2):345–359, 2012.

[46] Miklós Maróti and Ralph McKenzie. Existence theorems for weakly symmetric operations. *Algebra universalis*, 59(3):463–489, 2008.

[47] Jakub Opršal. the promised open problems, 2021. `https://community.dur.ac.uk/jakub.oprsal/open-problems/`.

[48] Nicholas Pippenger. Galois theory for minors of finite functions. *Discrete Mathematics*, 254(1):405–419, June 2002.

[49] Emil L Post. *The two-valued iterative systems of mathematical logic*. Number 5 in Annals of Mathematics Studies. Princeton University Press, 1941.

[50] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the 40th Annual ACM Symposium on Theory of Computing*, pages 245–254, 2008.

[51] Thomas J. Schaefer. The Complexity of Satisfiability Problems. In *Proceedings of the Tenth Annual ACM Symposium on Theory of Computing*, STOC '78, pages 216–226, New York, NY, USA, 1978. ACM.

[52] Johan Thapper and Stanislav Živný. The complexity of finite-valued CSPs. *Journal of the ACM*, 63(4):37:1–37:33, 2016. Preliminary version in STOC'13.

[53] Marcin Wrochna and Stanislav Živný. Improved hardness for *H*-colourings of *G*-colourable graphs. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1426–1435, 2020.

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