Advances in the Computational Complexity of Holant Problems

By

Tyson Williams

A dissertation submitted in partial fulfillment of

the requirements of the degree of

Doctor of Philosophy

(Computer Science)

at the

UNIVERSITY OF WISCONSIN–MADISON

2015

Date of final oral examination: 5/1/2015

The dissertation is approved by the following members of the Final Oral Committee:
Jin-Yi Cai, Professor, Computer Science
Eric Bach, Professor, Computer Science
Dieter van Melkebeek , Professor, Computer Science
Steffen Lempp, Professor, Mathematics
Joanna Ellis-Monaghan, Professor, Mathematics

To my wife Shannon — who has waited so patiently for the completion of this document.

Acknowledgments

I must begin by thanking my advisor, Jin-Yi Cai. When I came to graduate school to work with you, I had the determination to succeed but lacked the needed skills. You invested a great deal of your time in me and have made me into the researcher that I am today. Although I still make mistakes, you are scrupulous in spotting them and patient as you point them out to me. I am also very grateful for your generous financial support.

To my coauthors (Jin-Yi Cai, Michael Kowalczyk, Heng Guo, and Zhiguo Fu), it is easy for me to say that none of our papers would exist without your involvement. Each of you has had a significant impact on the way I conduct research and on the way I convey results to others. I hope that we have many more opportunities to work together in the future.

To the members of my defense committee (Jin-Yi Cai, Eric Bach, Dieter van Melkebeek, Steffen Lempp, and Joanna Ellis-Monaghan), thank you for the time that you spent serving on my committee. I especially appreciate the detailed comments that I received on a preliminary version of this document. Special thanks goes to Joanna for traveling to my defense.

When I began graduate school, I was fortunate to be surrounded by an engaging group of fellow theory students (Jeff Kinne, Matt Anderson, Siddharth Barman, David Malec, Dalibor Zelený, Seeun (William) Umboh, and Balasubramanian (Balu) Sivan). As I look back, I fondly remember our nerdy conversations punctuated with jokes that only a theorist could appreciate. You may not have realized it at the time, but I looked up to each of you as a role model. There was something true for each of you that I wanted to be true for me as well. By becoming more like each of you, I have become a better person—thank you.

As I complete graduate school, I remember the many theory students who started after me (Heng Guo, Aaron Gorenstein, Gautam Prakriya, Brian Nixon, Xi Wu, Nick Pappas, John (Ben) Miller, Alexander (Alexi) Brooks, Kevin Kowalski, Rex Fernando, Ian Skoch, Bryce Sandlund, and Andrew (Drew) Morgan). We had our fair share of similarly nerdy conversations as well. Just as I looked up to the theory students before me, I considered that each of you might be looking up to me. This has challenged me to be more thoughtful and caring, and ultimately a better person. For this, I am grateful to each of you.

To the members of my former bible study (Tyler Campbell, Kyle Furry, Andy Grimmer, Kerrigan Joseph, Andrew O'Connor, Brady Parks, Adam Post, Henry Weiner, Matt Winzenried) and my current bible study (Shannon Williams, Bryan and Lori Amundson, Ben and Christine Dickinson, Justin and Kate Greer, Scott and Sarah Melin, Andy and Meg Laird, Josh and Melissa Lazare, Matt and Margi Rusten, Brian and Erin Towns, Kenny and Shannon Younger), thank you for all of the much needed social time that we spent growing closer to each other and to God. I always felt refreshed and reenergized after spending time with each of you.

To my parents (Dave and Jenny Williams), you have sacrificed so much so that I could gain and have always been supportive of me and my decisions. As I enter a new stage of life, I am comforted by the knowledge that you will continue to be there for me.

Lastly, to my wife (Shannon Williams), thank you for loving me unconditionally. When I wrong you, you are eager to forgive and quick to forget. I cannot imagine my life without you. You are the best thing to happen to me (even more than getting a PhD!).

Abstract

We study the computational complexity of counting problems defined over graphs. Complexity dichotomies are proved for various sets of problems, which classify the complexity of each problem in the set as either computable in polynomial time or #P-hard. These problems are expressible in the frameworks of counting graph homomorphisms, counting constraint satisfaction problems, or Holant problems. However, the proofs are always expressed within the framework of Holant problems, which contains the other two frameworks as special cases. Holographic transformations are naturally expressed using this framework. They represent proofs that two different-looking problems are actually the same. We use them to prove both hardness and tractability. Moreover, the tractable cases are often stated using a holographic transformation.

The uniting theme in the proofs of every dichotomy is the technical advances achieved in order to prove the hardness. Specifically, polynomial interpolation appears prominently and is indispensable. We repeatedly strengthen and extend this technique and are rewarded with dichotomies for larger and larger classes of problems. We now have a thorough understanding of its power as well as its ultimate limitations. However, fundamental questions remain since polynomial interpolation is intimately connected with integer solutions of algebraic curves and determinations of Galois groups, subjects that remain active areas of research in pure mathematics.

Our motivation for this work is to understand the limits of efficient computation. Without settling the P versus #P question, the best hope is to achieve such complexity classifications.

Contents

Chapter 1

Introduction

The class #P is the counting version of NP. A problem in NP corresponds to a problem in #P by changing the question from "does a solution exist?" to "how many solutions exist?". It was defined by Valiant [125] in order to show that counting perfect matchings is #P-hard over bipartite graphs. Since then, many more problems have been shown to be #P-hard.

Although counting perfect matchings is #P-hard over bipartite graphs, the problem is computable in polynomial time over planar graphs. This was proven by Kasteleyn [89, 88]. Years later, Valiant [128] realized that a nontrivial fraction of quantum computation can be efficiently simulated on a classical computer by reduction to Kasteleyn's algorithm. The graph gadgets in this reduction are called *matchgates*. Immediately afterwards, he [132, 131] introduced holographic transformations to further extend the reach of Kasteleyn's algorithm. This produced polynomial-time algorithms for a number of counting problems over planar graphs for which only exponential-time algorithms were previously known.

These developments generated much excitement [77]. The new polynomial-time algorithms appear so exotic and unexpected, and they solve problems that appear so close to being #P-hard. Could it be that these new algorithmic techniques can efficiently solve everything in #P? Quoting Valiant [131]:

> "The objects enumerated are sets of polynomial systems such that the solvability of any one member would give a polynomial time algorithm for a specific problem. . . .

> the situation with the P = NP question is not dissimilar to that of other unresolved enumerative conjectures in mathematics. The possibility that accidental or freak objects in the enumeration exist cannot be discounted if the objects in the enumeration have not been studied systematically.

Indeed, if any "freak" object exists in this framework, it would collapse #P to P."

Therefore, over the past 10 to 15 years, these algorithm techniques been intensely studied in order to gain a systematic understanding to the limit of the trio of holographic reductions, matchgates, and Kasteleyn's algorithm [127, 24, 26, 42, 133, 43, 95, 106, 107]. Without settling the P versus #P question, the best hope is to achieve a complexity classification. This program finds its sharpest expression in a complexity dichotomy theorem, which classifies *every* problem expressible in a framework as either solvable in P or #P-hard, with nothing in between.

The study of these algorithm techniques has taken place in the counting framework of Holant problems [47]. The framework naturally encodes and expresses the problem of counting perfect matchings as well as Valiant's matchgates and holographic reductions. It is a refinement of counting Constraint Satisfaction Problems (#CSP)—a complete complexity classification for Holant problems implies one for #CSP. Another important special case is counting weighted graph homomorphisms.

We prove dichotomy theorems for various sets of counting problems. The problems in one set can be expressed as counting weighted graph homomorphisms (Chapter 5). The problems in another set can be expressed as #CSP (Chapter 8). The problems in the remaining sets are expressed as Holant problems (Chapters 6, 7, 10, and 11). However, the proofs are always expressed within the framework of Holant problems.

A common theme among the proofs of these dichotomy theorems is the use of polynomial interpolation to prove the hardness. The advances we achieve in strengthening and extending this technique ultimately lead to dichotomy theorems for larger and larger classes of Holant problems. In the context of Holant problems, we now have a thorough understanding of the possible reductions using polynomial interpolation. We also develop many new tools to show that a given interpolation will succeed. However, many questions remain. For some interpolations, success is intimately connected with the number and location of integer solutions of an algebraic curve; still others require the determination of Galois groups. Thus, a complete understanding of polynomial interpolation in the context of Holant problems depends on results in these active areas of research in pure mathematics.

In Chapter 2, we define the framework of Holant problems as well as the two special cases of #CSP and counting graph homomorphisms. In Chapter 3, we explain the most common reductions we use between Holant problems. Some of these reductions are used to prove both hardness and tractability. In Chapter 4, we introduce the known tractable cases and ask many unanswered questions regarding them.

Chapter 2

Definitions and Classes of Counting Problems

Let $\mathbb{N} = \{1, 2, 3, \ldots,\}$ be the set of natural numbers, $\mathbb{Z}$ the set of integers, $\mathbb{Z}^+$ the set of positive integers, and $\mathbb{C}$ the set of complex numbers. For $n \in \mathbb{Z}^+$, let $[n] = \{1, \ldots, n\}$ be the set of integers from 1 to n. We use $\mathbf{GL}_\kappa(\mathbb{C})$ to denote the set of invertible κ-by-κ matrices over $\mathbb{C}$. We use $\mathbf{O}_\kappa(\mathbb{C})$ to denote the set of matrices in $\mathbf{GL}_\kappa(\mathbb{C})$ that are orthogonal (i.e. $TT^\intercal = I_\kappa$).

We typically use polynomial-time Turing reductions to reduce between counting problems. For counting problems P_1 and P_2, we use $P_1 \le_T P_2$ to denote this type of reduction from P_1 to P_2. If $P_1 \le_T P_2$ and $P_2 \le_T P_1$, we write $P_1 \equiv_T P_2$. Some of our reductions are of a more restricted form (such as mapping reductions instead of Turing reductions), but we only point these out in special cases. In such cases, we use $\le$ or $\equiv$ (without the subscript T) to indicate that something is special about the reduction.

The inputs to our counting problems are graphs, which may have self-loops and parallel edges. A graph without self-loops or parallel edges is called a *simple* graph. A *plane* graph is a planar embedding of a planar graph.

All tensor products, which are denoted by $\otimes$, refer to the Kronecker product (of matrices).

2.1 Holant Problems

The framework of Holant problems is defined for functions mapping tuples of length $n \geq 1$ over a domain of finite size κ to a commutative semiring R. We consider the computational complexity of complex-weighted Holant problems (i.e. $R = \mathbb{C}$). For consideration of models of computation, functions take complex algebraic numbers.

Let $\mathcal{F}$ be a set of functions, which are called *signatures* or local constraint functions. A *signature grid* $\Omega = (G, \pi)$ over $\mathcal{F}$ consists of a graph $G = (V, E)$ with a linear order of the incident edges at each vertex and a function π that assigns to each vertex $v \in V$ some $f_v \in \mathcal{F}$. A Holant problem is parametrized by a set of signatures.

Definition 2.1.1. For a set $\mathcal{F}$ of signatures over a domain D of size κ, we define $\text{Holant}_\kappa(\mathcal{F})$ as:

Input: A *signature grid* $\Omega = (G, \pi)$ over $\mathcal{F}$;

Output:

$$\text{Holant}_\kappa(\Omega; \mathcal{F}) = \sum_{\sigma: E \to D} \prod_{v \in V} f_v(\sigma\mid_{E(v)}),$$

where

- $E(v)$ denotes the incident edges of v and
- $\sigma\mid_{E(v)}$ denotes the restriction of σ to $E(v)$ in the linear order at v.

We use $\text{Holant}(\mathcal{F})$ to denote $\text{Holant}_2(\mathcal{F})$. We use G in place of Ω when π is clear from context. We also omit $\mathcal{F}$ in the expression $\text{Holant}_\kappa(\Omega; \mathcal{F})$ when $\mathcal{F}$ is clear from context. When $\mathcal{F}$ is a finite set of signatures, we sometimes omit the curly braces and just list the signatures it contains. For example, if $\mathcal{F} = \{f, g\}$, then instead of writing $\text{Holant}_\kappa(\{f, g\})$, we may also write $\text{Holant}_\kappa(f, g)$. This is especially true when $\mathcal{F}$ is a singleton set.

A signature f of arity n over a domain D of size κ can be denoted by $(f_0, f_1, \ldots, f_{\kappa^n - 1})$, where f_i is the output of f on the ith lexicographical input based on an ordering of the elements in D. This is a listing of its outputs in lexicographical order as in a truth table. It is a vector in $\mathbb{C}^{\kappa^n}$, or a tensor (with a basis) in $(\mathbb{C}^\kappa)^{\otimes n}$. A symmetric signature f of arity n over the Boolean domain can be expressed as $[f_0, f_1, \ldots, f_n]$, where f_w is the value of f on inputs of Hamming weight w. An example is the EQUALITY signature $(=_n) = [1, 0, \ldots, 0, 1]$ of arity n.

We give examples of Holant problems using a symmetric signature f of arity n. Since f is the only available signature and its arity is n, every vertex of the graph G must be of degree n. Here are four examples over the Boolean domain.

$$\operatorname{Holant}_2(G; f) \text{ counts } \begin{cases} \text{matchings} & \text{in } G \text{ when } f = \textsc{At-Most-One}_n; \\ \text{perfect matchings} & \text{in } G \text{ when } f = \textsc{Exact-One}_n; \\ \text{cycle covers} & \text{in } G \text{ when } f = \textsc{Exact-Two}_n; \\ \text{edge covers} & \text{in } G \text{ when } f = \textsc{Or}_n. \end{cases}$$

An example for any domain of size κ is that $\operatorname{Holant}_\kappa(G; \textsc{All-Distinct}_n)$ counts edge colorings in G using a most κ colors. Each of these five examples are expressed as Holant problem in a straightforward manner. A less obvious example is that $\operatorname{Holant}_2(G; \frac{1}{2}[3, 0, 1, 0, 3])$ counts Eulerian orientations of a 4-regular graph G.

As stated above, one can view a signature as a tensor along with a basis. When doing so, a signature grid is equivalent to a tensor network, and the Holant of the signature grid is equal to the scalar that remains after contracting all edges in the corresponding tensor network. Tensors provide a way to express basis-free representations and are useful when the concepts being studied are invariant under a change of basis. However, one must chose a basis when considering the computational complexity of contracting tensor networks because the complexity depends on the choice of basis. An important special case of Definition 2.1.1 is evaluating the partition function of the edge coloring model, which is a graph polynomial. For more about the partition function of the edge coloring model and the contraction of tensor networks, see [114, Chapter 3].

A *planar signature grid* is a signature grid such that its underlying graph is planar and for some planar embedding, for every vertex v, the linear order of the incident edges at v agrees with the counterclockwise order of the incident edge at v in the embedding. We use $\text{Pl-Holant}_\kappa(\mathcal{F})$ to denote the restriction of $\operatorname{Holant}_\kappa(\mathcal{F})$ to planar signature grids. For signature sets $\mathcal{F}$ and $\mathcal{G}$, a *bipartite signature grid* over $(\mathcal{F} \mid \mathcal{G})$ is a signature grid $\Omega = (H, \pi)$ over $\mathcal{F} \cup \mathcal{G}$, where $H = (V, E)$ is a bipartite graph with bipartition $V = (V_1, V_2)$ such that $\pi(V_1) \subseteq \mathcal{F}$ and $\pi(V_2) \subseteq \mathcal{G}$. Signatures in $\mathcal{F}$

are considered as row vectors (or covariant tensors); signatures in $\mathcal{G}$ are considered as column vectors (or contravariant tensors) [57]. We use $\text{Holant}_\kappa(\mathcal{F} \mid \mathcal{G})$ to denote the restriction of $\text{Holant}_\kappa(\mathcal{F} \cup \mathcal{G})$ to bipartite signature grids over $(\mathcal{F} \mid \mathcal{G})$. A *planar bipartite signature grid* is one that is both planar and bipartite. We use $\text{Pl-Holant}_\kappa(\mathcal{F} \mid \mathcal{G})$ to denote the restriction to these signature grids.

A signature f of arity n is *degenerate* if there exist unary signatures $u_j \in \mathbb{C}^2$ $(1 \leq j \leq n)$ such that $f = u_1 \otimes \cdots \otimes u_n$. A symmetric degenerate signature has the form $u^{\otimes n}$. Replacing such signatures by n copies of the corresponding unary signature does not change the Holant value. Replacing a signature $f \in \mathcal{F}$ by a constant multiple cf, where $c \neq 0$, does not change the complexity of $\text{Holant}_\kappa(\mathcal{F})$. It introduces a nonzero factor to $\text{Holant}_\kappa(\Omega; \mathcal{F})$.

We allow $\mathcal{F}$ to be an infinite set. For $\text{Holant}_\kappa(\mathcal{F})$ to be tractable, the problem must be computable in polynomial time even when the description of the signatures in the input Ω are included in the input size. In contrast, we say $\text{Holant}_\kappa(\mathcal{F})$ is #P-hard if there exists a finite subset of $\mathcal{F}$ for which the problem is #P-hard. We say a signature set $\mathcal{F}$ is tractable (resp. #P-hard) if the corresponding counting problem $\text{Holant}_\kappa(\mathcal{F})$ is tractable (resp. #P-hard). Similarly for a signature f, we say f is tractable (resp. #P-hard) if $\{f\}$ is. We also speak of a signature or signature set as being tractable or #P-hard for the Holant problem defined over planar, bipartite, or planar and bipartite graphs. The class of graphs should be clear from context.

2.2 Constraint Satisfaction Problems

Like a Holant problem, a counting Constraint Satisfaction Problem (#CSP) [55] is parametrized by a set of local constraint functions $\mathcal{F}$ that we also call signatures. It is denoted by $\#\text{CSP}_\kappa(\mathcal{F})$ when the signatures in $\mathcal{F}$ are defined over a domain D of size κ. An instance of $\#\text{CSP}_\kappa(\mathcal{F})$ is a set C of clauses. Each clause is a constraint $f_c \in \mathcal{F}$ of some arity m together with its m inputs variables $x_{i_1}, \ldots, x_{i_m}$. The output is

$$\sum_{x_1, \ldots, x_n \in D} \prod_{(f_c, x_{i_1}, \ldots, x_{i_m}) \in C} f_c(x_{i_1}, \ldots, x_{i_m}). \tag{2.2.1}$$

The canonical example of a #CSP is #SAT, or Boolean satisfiability, the problem of counting satisfying assignments to a given Boolean (i.e. $\kappa = 2$) formula. As a #CSP, it is denoted by $\#\text{CSP}_2(\mathcal{F})$ with $\mathcal{F} = \{\textsc{Or}_n \mid n \geq 1\} \cup \{\neq_2\}$, where $\textsc{Or}_n$ is the $\textsc{Or}$ function of arity n and $(\neq_2) = [0, 1, 0]$ is the binary DISEQUALITY function. Here are several more well-known examples of CSPs over the Boolean domain and their corresponding constraint set $\mathcal{F}$:

$$
\begin{array}{rll}
\textsc{Sat} & \text{has} & \mathcal{F} = \{\textsc{Or}_n \mid n \geq 1\} \cup \{\neq_2\} \\
3\textsc{Sat} & \text{has} & \mathcal{F} = \{\textsc{Or}_3, \neq_2\} \\
1\text{-}\textsc{in}\text{-}3\textsc{Sat} & \text{has} & \mathcal{F} = \{\textsc{Exact-One}_3, \neq_2\} \\
\text{NAE-}3\textsc{Sat} & \text{has} & \mathcal{F} = \{\textsc{Not-All-Equal}_3, \neq_2\} \\
\\
\textsc{Mon-Sat} & \text{has} & \mathcal{F} = \{\textsc{Or}_n \mid n \geq 1\} \\
\textsc{Mon}\text{-}3\textsc{Sat} & \text{has} & \mathcal{F} = \{\textsc{Or}_3\} \\
\textsc{Mon}\text{-}1\text{-}\textsc{in}\text{-}3\textsc{Sat} & \text{has} & \mathcal{F} = \{\textsc{Exact-One}_3\} \\
\textsc{Mon}\text{-NAE-}3\textsc{Sat} & \text{has} & \mathcal{F} = \{\textsc{Not-All-Equal}_3\}
\end{array}
$$

Notice that prefixes like "3" and "MON" (for monotone) are used to say which constraints are *not* present in $\mathcal{F}$.

By $\#\text{CSP}_\kappa^d(\mathcal{F})$, we denote the special case of $\#\text{CSP}_\kappa(\mathcal{F})$ in which every variable must appear some multiple of d times. Note that $\#\text{CSP}_\kappa(\mathcal{F})$ is the same as $\#\text{CSP}_\kappa^d(\mathcal{F})$ with $d = 1$. We can express $\#\text{CSP}_\kappa^d(\mathcal{F})$ as a Holant problem. An instance of $\#\text{CSP}_\kappa^d(\mathcal{F})$ has the following bipartite view. Create a node for each variable and each clause. Connect a variable node to a clause node if the variable appears in the clause. This bipartite graph is also known as the *constraint graph*. To each variable vertex, we assign the EQUALITY signature of the appropriate arity. To each clause vertex, we assign the constraint used in that clause. Under this view, we see that

$$\#\text{CSP}_\kappa^d(\mathcal{F}) \equiv_T \text{Holant}_\kappa(\mathcal{EQ}_d \mid \mathcal{F}), \tag{2.2.2}$$

where $\mathcal{EQ}_d = \{=_{dk} \mid k \geq 1\}$ is the set of EQUALITY signatures of whose arities are a multiple of d. We denote by Pl-$\#\text{CSP}_\kappa^d(\mathcal{F})$ the restriction of $\#\text{CSP}_\kappa^d(\mathcal{F})$ to inputs with a planar constraint graph. The construction above also shows that

$$\text{Pl-\#CSP}^d_\kappa(\mathcal{F}) \equiv_T \text{Pl-Holant}_\kappa(\mathcal{EQ}_d \mid \mathcal{F}). \tag{2.2.3}$$

If $d \in \{1, 2\}$, then more is true.

Lemma 2.2.1. *Let $\mathcal{F}$ be a set of signatures over a domain of size κ. If $d \in \{1, 2\}$, then*

$$\#\text{CSP}^d_\kappa(\mathcal{F}) \equiv_T \text{Holant}_\kappa(\mathcal{EQ}_d \cup \mathcal{F}) \qquad \textit{and} \qquad \text{Pl-\#CSP}^d_\kappa(\mathcal{F}) \equiv_T \text{Pl-Holant}_\kappa(\mathcal{EQ}_d \cup \mathcal{F}).$$

Proof. By (2.2.2) and (2.2.3), it suffices to show

$$\text{Holant}_\kappa(\mathcal{EQ}_d \mid \mathcal{F}) \equiv_T \text{Holant}_\kappa(\mathcal{EQ}_d \cup \mathcal{F}) \qquad \text{and} \qquad \text{Pl-Holant}_\kappa(\mathcal{EQ}_d \mid \mathcal{F}) \equiv \text{Pl-Holant}_\kappa(\mathcal{EQ}_d \cup \mathcal{F}).$$

In both cases, the reduction from left to right in the second equivalence is trivial; just ignore the bipartite restriction. For the other direction, we take a signature grid for the problem on the right and create a bipartite signature grid for the problem on the left such that both signature grids have the same Holant value up to an easily computable factor. If the initial graph is planar, then the final graph will also be planar, so this will prove both equivalences.

If two signatures in $\mathcal{F}$ are assigned to adjacent vertices, then we subdivide all edges between them and assign the binary Equality signature $=_2 \in \mathcal{EQ}_d$ to all new vertices. Suppose Equality signatures $=_n, =_m \in \mathcal{EQ}_d$ are assigned to adjacent vertices connected by ℓ edges. If $n = m = \ell$, then we simply remove these two vertices. The Holant of the resulting signature grid differs from the original by a factor of κ. Otherwise, we contract all ℓ edges and assign $=_{n+m-2k} \in \mathcal{EQ}_d$ to the new vertex. □

2.3 Graph Homomorphism Problems

Given two graphs G and H, a graph homomorphism from G to H is a map σ from the vertex $V(G)$ to the vertex set $V(H)$ such that the edge $(u, v) \in E(G)$ is mapped to the edge $(\sigma(u), \sigma(v)) \in E(H)$. Then one can define a counting problem by asking for the number of homomorphisms from G to

(a) #VertexCover (b) #Antichain

Figure 2.1: Target graphs H and the combinatorial counting problems they define as #H-coloring problems.

H. One can even fix H and just consider G as the input. This variant is known as the #H-coloring problem, and H is called the target graph. This name comes from the following fact. If $H = K_n$, the complete graph on n vertices, then the #H-coloring problem is to compute the number of vertex colorings of G using at most n colors (see Figure 2.2).

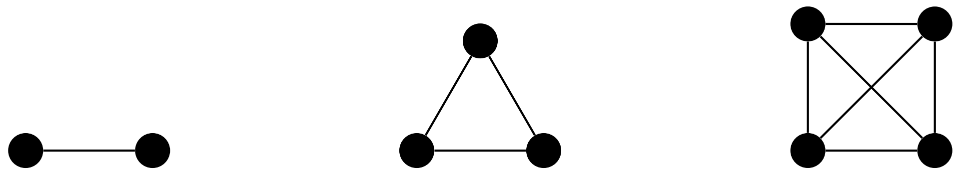
There are two more combinatorial problem that one can express. If H is the two-vertex graph connected by a single edge and one vertex has one self-loop (see Figure 2.1a), then the #H-coloring problem is to compute the number of vertex covers (or equivalently, the number of independent sets). Suppose H is the two-vertex directed graph connected by a single directed edge and both vertices have one directed self-loop (see Figure 2.1b). Then the #H-coloring problem takes directed graphs as input. If the input is a directed acyclic graph, then it defines a partial order, and the #H-coloring problem is to compute the number of antichains (or equivalently, the number of lower sets, or equivalently, the number of upper sets) in this partial order.

More generally, we consider directed graphs with weights. Let A be a κ-by-κ matrix over $\mathbb{C}$. Given a directed graph $G = (V, E)$, the graph homomorphism problem is to compute

$$Z_A(G) = \sum_{\sigma : V \to [\kappa]} \prod_{(u,v) \in E} A_{\sigma(u),\sigma(v)}. \tag{2.3.4}$$

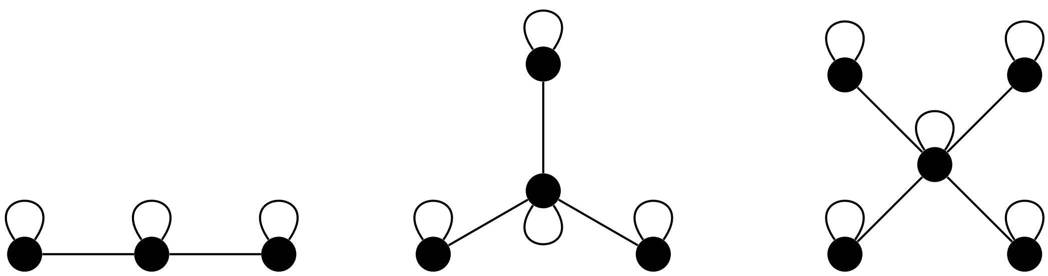
The target graph H is now implicitly defined by the matrix A, which is the weighted adjacency matrix of H. By convention, an ordered pair of vertices (u, v) is an edge in a directed graph G if there is an edge directed from u to v (i.e. the tail of the directed edge is at u and the head of the directed edge is at v).

In statistical physics, (2.3.4) is called the partition function. For various models of particle interactions, it represents the total energy of a system as one sums over every possible configuration



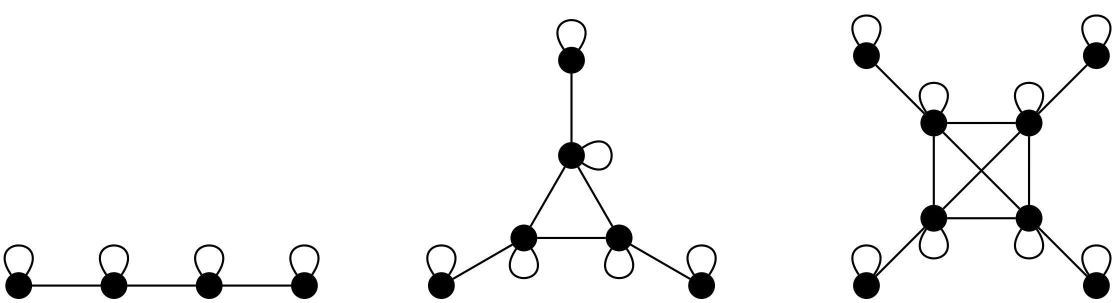
(a) $q = 2$ (b) $q = 3$ (c) $q = 4$

Figure 2.2: Target graph H for counting q-colorings for $q \in \{2, 3, 4\}$ as an $\#H$-coloring problem.



(a) $q = 2$ (b) $q = 3$ (c) $q = 4$

Figure 2.3: Target graph H for counting q-particle Widom-Rowlinson configurations for $q \in \{2, 3, 4\}$ as an $\#H$-coloring problem.



(a) $q = 2$ (b) $q = 3$ (c) $q = 4$

Figure 2.4: Target graph H for counting q-type Beach model configurations for $q \in \{2, 3, 4\}$ as an $\#H$-coloring problem.

of the particles. The Ising model [83] corresponds to the partition function with matrix $A = \left[\begin{smallmatrix} \lambda & 1 \\ 1 & \lambda \end{smallmatrix}\right]$ for a parameter λ that can take any positive real number. The Ashkin-Teller Model [2] corresponds to the partition function with matrix $A = \left[\begin{smallmatrix} a & b & c & d \\ b & a & d & c \\ c & d & a & b \\ d & c & b & a \end{smallmatrix}\right]$ for parameters a, b, c, d that can take any positive real numbers. The Potts model [112] corresponds to the partition function with matrix $A = J_n + (\lambda - 1)I_n$, where J_n is the n-by-n matrix of all 1's and I_n is the n-by-n identity matrix. Once again, λ is a parameter that can take any positive real number.

Counting configurations in the q-particle Widom-Rowlinson model [142] corresponds to the #H-coloring problem in which H is the star graph on $q+1$ vertices and all vertices have self loops (see Figure 2.3). Counting configurations in the q-type Beach model [18, 19] (cf. [75, Chapter 8]) corresponds to the #H-coloring problem in which H is the complete graph on q vertices, each of these q vertices has a pendant vertex, and all $2q$ vertices have a self loop (see Figure 2.4). For more on this connection with statistical physics, see [141, Chapter 4] or [121, Chapter 2] (as well as [5]).

We can express any #H-coloring problem as a counting constraint satisfaction problem. This, in turn, allows us to express any #H-coloring problem as a Holant problem by Lemma 2.2.1. Let A be a κ-by-κ matrix, which defines the target graph H with vertex set $V(H) = [\kappa]$. Let the input graph have n vertices. Then the exponential sum in (2.3.4) sums over all ways to assign one of κ possible values to these n vertices just as the exponential sum in (2.2.1) sums over all ways to assign one of κ possible values to the n variables. Then in the product of (2.3.4), each edge is playing the role of a clause with a binary constraint defined by the matrix A.

We summarize this in the following lemma.

Lemma 2.3.1. *Let A be a κ-by-κ matrix, and let f be a binary signature defined over the domain $[\kappa]$ such that $f(x, y) = A_{x,y}$. Then*

$$Z_A(\cdot) \equiv_T \#\mathrm{CSP}_\kappa(f).$$

If the input of $Z_A(\cdot)$ is restricted to be planar, then

$$Z_A(\cdot) \equiv_T \text{Pl-}\#\mathrm{CSP}_\kappa(f).$$

Chapter 3

Reductions

3.1 Gadget Constructions

3.1.1 Local Gadget Constructions via $\mathcal{F}$-gates

A basic type of reduction is what might be generally known as a gadget construction. In the context of Holant problems, we create "local" gadget constructions in order to realize a signature. Fix a set $\mathcal{F}$ of signature over a domain D of size κ. We say a signature f is *realizable* or *obtainable* from $\mathcal{F}$ if there is a gadget with some dangling edges such that each vertex is assigned a signature from $\mathcal{F}$, and the resulting graph, when viewed as a black-box signature with inputs on the dangling edges, is exactly f.

Formally, such a notion is defined by an $\mathcal{F}$-gate [46]. An $\mathcal{F}$-gate F is similar to a signature grid (G, π) for $\operatorname{Holant}(\mathcal{F})$ except that $G = (V, E, E')$ is a graph with regular edges in E and m dangling edges in E'. The dangling edges define external variables for the $\mathcal{F}$-gate. They are ordered by starting at the edge marked with a diamond and proceeding counterclockwise. (See Figure 3.1 for an example.) Then the $\mathcal{F}$-gate F defines the function

$$\Gamma(y_1, \ldots, y_m) = \operatorname{Holant}_\kappa(G', \pi'),$$

where $(y_1, \ldots, y_m) \in D^m$ is an assignment on the dangling edges, G' is the graph obtained from G

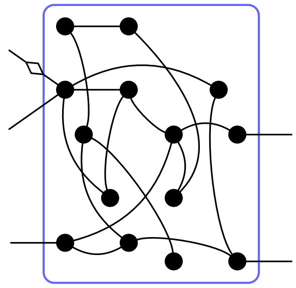


Figure 3.1: An $\mathcal{F}$-gate with 5 dangling edges.

after attaching the dangling end of a dangling edge $e_i \in E'$ (which is assigned y_i) to a new vertex v_{e_i}, and $\pi(v_{e_i}) = \delta_{y_i}$ is the function that outputs 1 when its input is y_i and outputs 0 otherwise. We call this function Γ the signature of the $\mathcal{F}$-gate. We also call an $\mathcal{F}$-gate a gadget. If the signature of an $\mathcal{F}$-gate is invariant under cyclic permutations of inputs, then we omit the diamond since it is unnecessary. We say that such signatures are *rotationally symmetric*.

An $\mathcal{F}$-gate is planar if the underlying graph can be embedded in the plane without edge crossings and the dangling edges are in the outer face. Now suppose we have two signature sets $\mathcal{F}$ and $\mathcal{G}$ in the context of a bipartite Holant problem $\text{Holant}_\kappa(\mathcal{F} \mid \mathcal{G})$. Then an $(\mathcal{F} \mid \mathcal{G})$-gate is an $(\mathcal{F} \cup \mathcal{G})$-gate such that the underlying graph is bipartite, the vertices in one part are assigned signatures from $\mathcal{F}$, and the vertices in the other part are assigned signatures from $\mathcal{G}$. Furthermore, we say that an $(\mathcal{F} \mid \mathcal{G})$-gate is *on the left* (resp. *on the right*) if each vertex incident to a dangling edge is assigned a signature from $\mathcal{F}$ (resp. $\mathcal{G}$). A planar $(\mathcal{F} \mid \mathcal{G})$-gate is both a planar $(\mathcal{F} \cup \mathcal{G})$-gate and an $(\mathcal{F} \mid \mathcal{G})$-gate.

Using $\mathcal{F}$-gates, we can reduce one Holant problem to another.

Lemma 3.1.1. *Let $\mathcal{F}$ be a set of signatures over a domain of size κ. If there exists an $\mathcal{F}$-gate with signature f, then*

$$\text{Holant}_\kappa(\mathcal{F} \cup \{f\}) \leq_T \text{Holant}_\kappa(\mathcal{F}).$$

Similar statements hold for gadgets that are planar, bipartite, or both for Holant problems defined over the same class of graphs.

Proof. Let F be an $\mathcal{F}$-gate with signature f. Given an instance Ω of $\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{f\})$, we replace every appearance of f by the $\mathcal{F}$-gate F to obtain an instance Ω' of $\text{Pl-Holant}_\kappa(\mathcal{F})$. Since

f is the signature of the $\mathcal{F}$-gate F, the Holant values for these two signature grids are identical. Furthermore, the size of F is a constant with respect to Ω, so the size Ω' is only a constant factor larger than Ω. □

Even for a very simple signature set $\mathcal{F}$, the signatures for all $\mathcal{F}$-gates can be quite complicated and expressive. Indeed, given a set $\mathcal{F}$ of signatures and another signature f, it is undecidable to decide if there exists an $\mathcal{F}$-gate with signature f [54, Theorem 2].

It is convenient to write a signature as a matrix. An immediate advantage is that a matrix is more of a pictorial representation than a vector, which aids understanding. However, the more important reason is to simplify the computation of a gadget's signature.

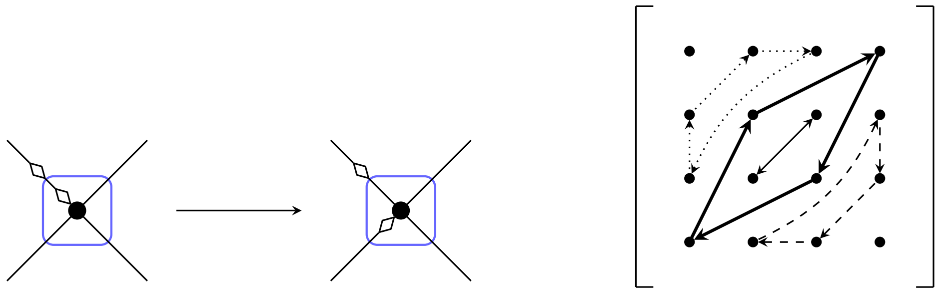
Definition 3.1.2. Let f be a signature of arity n over a domain of size κ. The *signature matrix of f with parameter ℓ* is an κ^ℓ-by-$\kappa^{n-\ell}$ matrix for some integer $0 \le \ell \le n$ in which the first ℓ inputs (in order) are the row index and the remaining $n-\ell$ inputs (in reverse order) are the column index.

If the arity of f is even, then the *signature matrix of f*, without specifying a parameter, is the signature matrix of f with parameter $\ell = \frac{n}{2}$ and is denoted by M_f.

The purpose of reversing the order of the column index is so that we can use matrix product in some of our gadget computations. If $f = (w, x, y, z)$ of arity 2 over the Boolean domain, then $M_f = \left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]$. If g is a signature of arity 4 over the Boolean domain with $g(w, x, y, z) = g^{wxyz}$, then

$$M_g = \begin{bmatrix} g^{0000} & g^{0010} & g^{0001} & g^{0011} \\ g^{0100} & g^{0110} & g^{0101} & g^{0111} \\ g^{1000} & g^{1010} & g^{1001} & g^{1011} \\ g^{1100} & g^{1110} & g^{1101} & g^{1111} \end{bmatrix}.$$

A signature matrix of a signature is also known as a Young flattening of a tensor [96, Section 3.4]. Let F be an $\mathcal{F}$-gate with signature f of arity n. We often depict F with ℓ dangling edges protruding to the left and $n - \ell$ dangling edges protruding to the right to aid in the mapping from F to the signature matrix of f with parameter ℓ. This is unnecessary is when f is symmetric, or more generally, when f is invariant under cyclic permutations of its inputs.



(a) A counterclockwise rotation (b) Movement of signature matrix entries

Figure 3.2: The movement of the entries in the signature matrix of an arity 4 signature over the Boolean domain under a counterclockwise rotation of the input edges. Entries of Hamming weight 1 are in the dotted cycle, entries of Hamming weight 2 are in the two solid cycles (one has length 4 and the other one is a swap), and entries of Hamming weight 3 are in the dashed cycle.

For a signature f, a simple $\{f\}$-gate is a single vertex assigned f but with its inputs cyclically permuted. Suppose f is of arity 4 and is defined over the Boolean domain. Consider an $\{f\}$-gate F with signature f. If we rotate F counterclockwise by a quarter turn (or equivalently, if we cyclically permute the inputs to f so that the first input becomes the last input), then we get a new $\{f\}$-gate F' with signature f', and the signature matrix of f' is easily determined from the signature matrix of f using Figure 3.2.

3.1.2 Domain Bundling

Consider the signature $f = (1, 2, 3, 4, 5, 6, 7, 8, 9)$. I did not specify the arity of f or the domain size over which it is defined. Is it possible to determine these two values given the fact that f has nine outputs? It is not; this situation is ambiguous. Let n be the arity of f, and let k be the size of the domain over which it is defined. What we know is that $k^n = 9$, but there are two (positive integer) solutions to this: either $\kappa = 9$ and $n = 1$ or $\kappa = 3$ and $n = 2$.

We can utilize this ambiguity to create a reduction that we call domain bundling.

Lemma 3.1.3. *Suppose f is a signature of arity n over a domain of size κ. If $\kappa^n = (\kappa')^{n'}$ and $n \mid n'$, then*

$$\operatorname{Holant}_\kappa(f) \leq_T \operatorname{Holant}_{\kappa'}(f'),$$

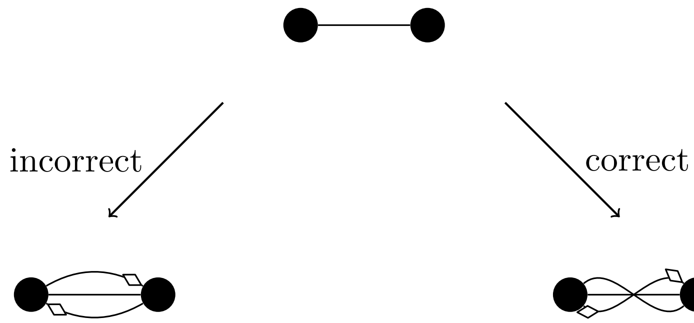


Figure 3.3: Simple example of domain bundling. Both the correct as well as the naive and incorrect way to connect the bundled edges are shown. As defined in the proof of Lemma 3.1.3, the graph G is above and the graph G' is below (and on the right).

where f' has the same output values as f (when sorted lexicographically by input) but is viewed as a signature of arity n' over a domain of size κ'.

Proof. If $(\kappa, n) = (\kappa', n')$, then there is nothing to prove, so assume otherwise. Then $n < n'$ and $\kappa > \kappa'$ since $n \mid n'$.

Let G be an instance of $\text{Holant}_\kappa(f)$, which must be an n-regular graph. We construct an instance G' of $\text{Holant}_{\kappa'}(f')$, which must be an n'-regular graph. The vertices in G' are same as the vertices in G. Each edge in G corresponds to $\frac{n'}{n}$ edges in G' between the same pair of vertices. These $\frac{n'}{n}$ edges, which each take one of κ' possible assignments, simulate the κ possible assignments in G since $\kappa = (\kappa')^{\frac{n'}{n}}$. It remains to ensure that the inputs of f' at each vertex map correctly to the variables on its incident edges.

By convention, the inputs of a signature map to the variables on its incident edges from first to last (or equivalently, from most to least significant) as one traverses the edges counterclockwise in some (not necessarily planar) embedding in the plane from specified initial edge. Consider two incident vertices in G by an edge e. The variable on this edge takes κ possible values and of course the two copies of f assigned to the two incident vertices agree on the meaning of these κ values. But in G', care must be taken to ensure that the two copies of f' agree on the meaning of assignments to the corresponding $\frac{n'}{n}$ incident edges. If they were connected in parallel in a planar way, then the order that one copy of f' would interpret the $\kappa = (\kappa')^{\frac{n'}{n}}$ assignments to these $\frac{n'}{n}$ edges would be exactly opposite to the order that the other copy of f' would interpret them. To fix this, we connect the edges to these vertices in the opposite order. At each vertex, the (cyclic) order is determined by a counterclockwise traversal of the vertex. □

Example 3.1.4. *Consider the graph G at the top of Figure 3.3. Suppose the signature assigned to both vertices is $f = (f_0, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$ of arity 1 over a domain of size 8. The Holant of this graph is*

$$f_0^2 + f_1^2 + f_2^2 + f_3^2 + f_4^2 + f_5^2 + f_6^2 + f_7^2.$$

We can also view $(f_0, f_1, f_2, f_3, f_4, f_5, f_6, f_7)$ as a signature f' of arity 3 over a domain of size 2. If we assign f' to both vertices in the graph in the lower left of Figure 3.3, then the Holant is

$$f_0^2 + f_1 f_4 + f_2^2 + f_3 f_6 + f_4 f_1 + f_5^2 + f_6 f_3 + f_7^2.$$

This Holant differs from the previous Holant because the three bundled edges were connected in the "wrong" order. In contrast, if we assign f' to both vertices in the graph in the lower right of Figure 3.3, then the Holant is

$$f_0^2 + f_1^2 + f_2^2 + f_3^2 + f_4^2 + f_5^2 + f_6^2 + f_7^2,$$

just as it was before applying the domain bundling reduction.

If f' were a symmetric signature, then connecting the $\frac{n'}{n}$ edges in any order yields the same contribution from f' and thus the same Holant value. In particular, one can put the edges in parallel in a planar way, which gives a planar reduction.

Corollary 3.1.5. *Let the situation be as in Lemma 3.1.3. If f' is symmetric, then*

$$\text{Pl-Holant}_\kappa(f) \le_T \text{Pl-Holant}_{\kappa'}(f').$$

I introduced the idea of domain bundling in the context of a single signature. However, we typically apply this argument to multiple signatures in a bipartite graph. The proof for the bipartite case is essentially the same as the proof of Lemma 3.1.3 when the bipartite restriction is absent.

Lemma 3.1.6. *Suppose f (resp. g) is a signature of arity n (resp. m) over a domain of size κ. If $\kappa^n = (\kappa')^{n'}$ and $n \mid n'$ as well as $\kappa^m = (\kappa')^{m'}$ and $m \mid m'$, then*

$$\operatorname{Holant}_\kappa(f \mid g) \le_T \operatorname{Holant}_\kappa(f' \mid g'),$$

where f' (resp. g) has the same output values as f (resp. g) (when sorted lexicographically by input) but is viewed as a signature of arity n' (resp. m') over a domain of size κ'.

3.2 Equivalent Expressions of the Same Problem

Some counting problems can be expressed as a Holant problem in more than one way. The primary way of mapping between these expressions is with a holographic transformation. Recently, another mapping was found, which is given in Subsection 3.2.2.

3.2.1 Holographic Transformation

A holographic transformation is the primary way of showing that two Holant problems with different expressions are actually the same. Surely you already know the simplest example of this: in a graph, the number of vertex covers is equal to the number of independent sets. The proof is that the complement of one type of set is the other. At the vertex level, the complement exchanges the assignments of 0 and 1.

A holographic transformation puts the assignments of 0 and 1 into a superposition much like the states of a qubit in quantum computing. However, quantum computation is not required—or even any computation at all. Just like the example with vertex covers and independent sets, we are merely describing a mathematical proof that two different looking problems are actually the same. We are undergoing a change of basis and viewing the problem from this new perspective.

To formally introduce the idea of holographic reductions, it is convenient to consider bipartite graphs. For a general graph, we can always transform it into a bipartite graph while preserving the Holant value as follows. For each edge in the graph, we replace it by a path of length two. (This operation is called the *2-stretch* of the graph and yields the edge-vertex incidence graph.) Each new vertex is assigned the binary EQUALITY signature $(=_2) = [1, 0, 1]$.

For a κ-by-κ matrix T and a signature set $\mathcal{F}$, define $T\mathcal{F} = \{g \mid \exists f \in \mathcal{F} \text{ of arity } n,\ g = T^{\otimes n} f\}$, and similarly for $\mathcal{F}T$, where the tensor product denoted by $\otimes$ is the Kronecker product. Whenever we write $T^{\otimes n} f$ or $T\mathcal{F}$, we view the signatures as column vectors; similarly for $fT^{\otimes n}$ or $\mathcal{F}T$ as row vectors.

Let T be an invertible κ-by-κ matrix. The holographic transformation defined by T is the following operation: given a signature grid $\Omega = (H, \pi)$ of $\text{Holant}(\mathcal{F} \mid \mathcal{G})$, for the same bipartite graph H, we get a new grid $\Omega' = (H, \pi')$ of $\text{Holant}(\mathcal{F}T \mid T^{-1}\mathcal{G})$ by replacing each signature in $\mathcal{F}$ or $\mathcal{G}$ with the corresponding signature in $\mathcal{F}T$ or $T^{-1}\mathcal{G}$. Then by a result typically called Valiant's Holant Theorem [132] (see also [25]), the Holant value has not changed. His proof is for domain size $\kappa = 2$, but also holds for any domain size. We state this as a lemma and provide a proof.

Lemma 3.2.1. *Let $\mathcal{F}$ and $\mathcal{G}$ be sets of complex-valued signatures over a domain of size κ. Suppose Ω is a bipartite signature grid over $(\mathcal{F} \mid \mathcal{G})$. If $T \in \mathbf{GL}_\kappa(\mathbb{C})$, then*

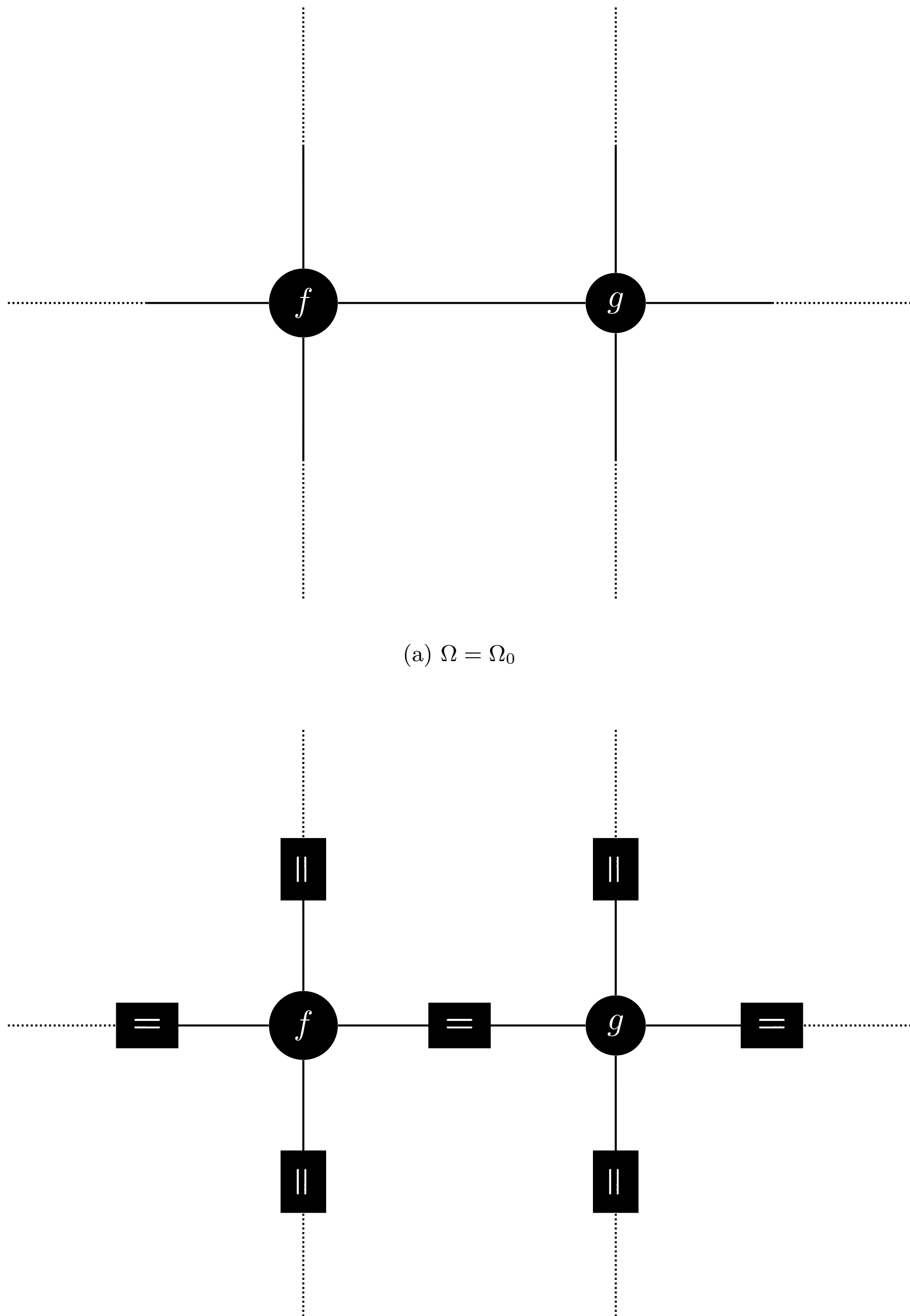
$$\text{Holant}_\kappa(\Omega; \mathcal{F} \mid \mathcal{G}) = \text{Holant}_\kappa(\Omega'; \mathcal{F}T \mid T^{-1}\mathcal{G}),$$

where Ω' is the corresponding signature grid over $(\mathcal{F}T \mid T^{-1}\mathcal{G})$.

Proof. We modify $\Omega = \Omega_0$ in several steps until it becomes $\Omega' = \Omega_3$. In each step, the Holant value is unchanged. We illustrate our proof in Figure 3.4.

Let $G = (U, V, E)$ be the graph underlying $\Omega = \Omega_0$. Vertices in U are assigned signatures in $\mathcal{F}$ by π while vertices in V are assigned signatures in $\mathcal{G}$ by π. We do the following operations for each edge in E. Let $e = (u, v) \in E$ be an edge with endpoints $u \in U$ and $V \in V$. Initially, Figure 3.4a depicts the neighborhood of u and v in $\Omega = \Omega_0$. In this example, both u and v are incident to three other edges but the vertices incident to the other ends of these edges are not shown.

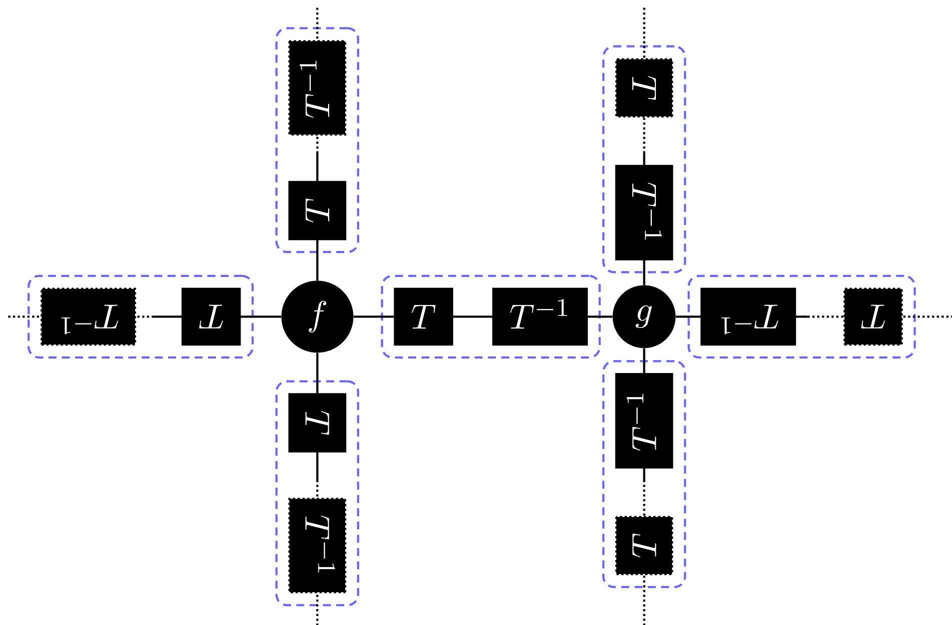
We do the following operations for each edge in E. Let $e = \{u, v\} \in E$ be an edge with endpoints $u \in U$ and $V \in V$. We subdivide e and assign $=_2$ to the new vertex w. This increases the number of terms in the Holant sum by a factor of 3. All original terms still appear and all new terms are 0. Let the resulting signature grid be Ω_1. See Figure 3.4b.



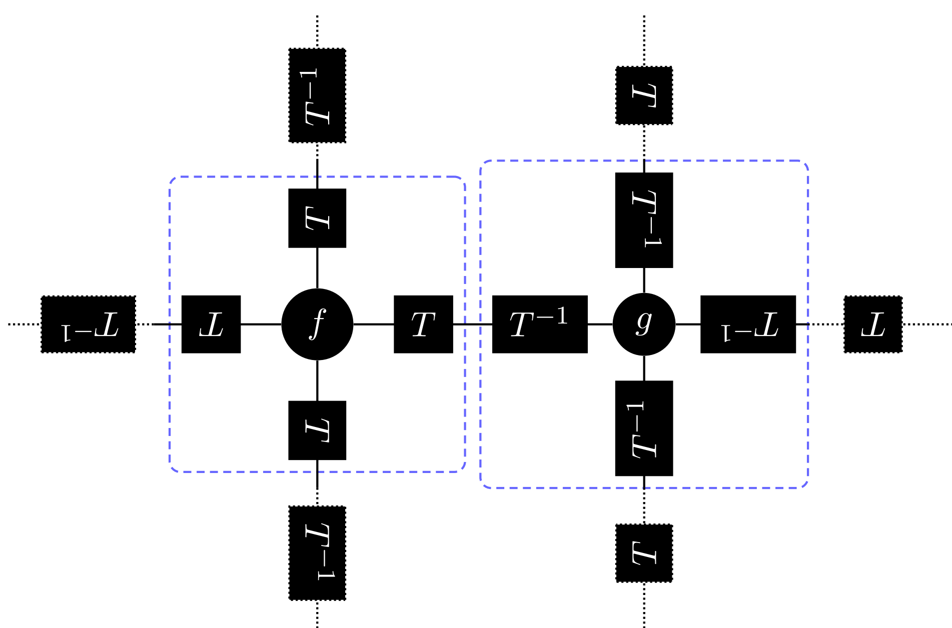
(a) $\Omega = \Omega_0$

(b) Ω_1

Figure 3.4: Neighborhood around two adjacent vertices.



(c) Ω_2



(d) Ω_3

Figure 3.4: Neighborhood around two adjacent vertices.

Then we subdivide w to get two adjacent vertices w_u and w_v so that we now have a path (u, w_u, w_v, v). Let f_u (resp f_v) be the signature whose signature matrix is T (resp. T^{-1}). We assign f_u (resp. f_v) to w_u (resp. w_v). If T is not a symmetric matrix, then f_u and f_v are not symmetric and it matters which edge corresponds to which input. The first input for f_u (resp. f_v) corresponds to the edge $\{u, w_u\}$ (resp. $\{w_u, w_v\}$). Let the resulting signature grid be Ω_2. See Figure 3.4c, which indicates the first inputs of f_u and f_v by the rotation of T and T^{-1} (instead of the standard notation of putting a diamond on the edge corresponding to the first input). Their first input is to their left. If we contract the edge $\{w_u, w_v\}$ within the dashed box, then we get back Ω_1. Thus, the Holant value has not changed.

Now Ω_3 in Figure 3.4d is actually the same as Ω_2. To get $\Omega_4 = \Omega'$, we contract $\{u, w_u\}$ and $\{w_v, v\}$. After doing so, we once again have the graph G. What has changed is the assignment to each vertex. The vertex u is now assigned $fT^{\otimes 4}$, and the vertex v is now assigned $(T^{-1})^{\otimes 4} g$.

In general, the new assignment to each vertex is the transformed signature in $\mathcal{F}T$ or $T^{-1}\mathcal{G}$ respectively, as claimed. □

Therefore, an invertible holographic transformation does not change the complexity of the Holant problem in the bipartite setting. Furthermore, there is a special kind of holographic transformation, the orthogonal transformation, that preserves the binary equality and thus can be used freely in the standard setting. For $\kappa = 2$, this first appeared in [45] as Theorem 2.2. We also state it as a lemma.

Lemma 3.2.2. *Let $\mathcal{F}$ be a set of complex-valued signatures over a domain of size κ. Suppose Ω is a signature grid over $\mathcal{F}$. If $H \in \mathbf{O}_\kappa(\mathbb{C})$, then*

$$\operatorname{Holant}_\kappa(\Omega; \mathcal{F}) = \operatorname{Holant}_\kappa(\Omega'; H\mathcal{F}),$$

where Ω' is the corresponding signature grid over $H\mathcal{F}$.

We use Lemma 3.2.1 and Lemma 3.2.2 to reduce between both hard problems and easy problems. Some of our reductions between easy problems use the following definition.

Definition 3.2.3. Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures over a domain of size κ. We say $\mathcal{F}$ is $\mathscr{C}$-transformable if there exists a $T \in \mathbf{GL}_\kappa(\mathbb{C})$ such that $(=_2)T^{\otimes 2} \in \mathscr{C}$ and $\mathcal{F} \subseteq T\mathscr{C}$.

This definition is important because if $\operatorname{Holant}_\kappa(\mathscr{C})$ is tractable over any set of graphs, then $\operatorname{Holant}_\kappa(\mathcal{F})$ is tractable over the same set of graphs for any $\mathscr{C}$-transformable set $\mathcal{F}$.

Lemma 3.2.4. *Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures over a domain of size κ. If $\mathcal{F}$ is $\mathscr{C}$-transformable, then*

$$\operatorname{Holant}_\kappa(\mathcal{F}) \leq_T \operatorname{Holant}_\kappa(\mathscr{C}),$$

where both problems are defined over the same set of graphs.

3.2.2 Another Identity

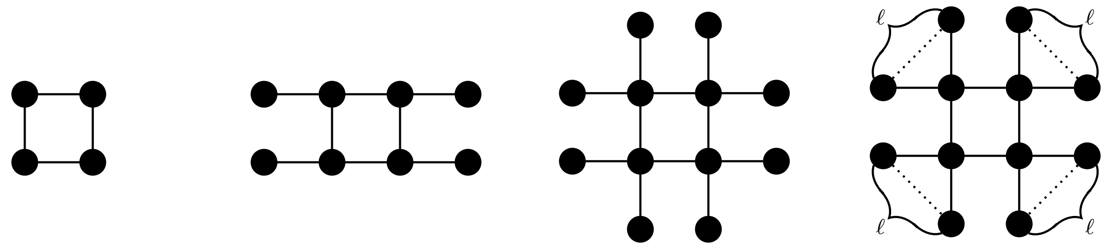
Until recently, every pair of Holant problems with the same output for every input were known to have a holographic transformation between them (except for some trivial examples). Then in [29, Subsection 4.3], we made the following observation.

Lemma 3.2.5. *Let x and y be indeterminates. Then for every $(2,4)$-regular bipartite graph G,*

$$\operatorname{Holant}_2(G; [0,1,0] \mid [x,y,1,0,0]) = \operatorname{Holant}_2(G; [0,1,0] \mid [0,0,1,0,0])$$

as polynomials in x and y.

Proof. Consider $\operatorname{Holant}_2(G; [0,1,0] \mid [x,y,1,0,0])$ for any $(2,4)$-regular bipartite graph G, which is a polynomial $p(x,y)$ in the indeterminates x and y. Because $[0,1,0]$ is the only signature on the left, any nonzero term in the Holant sum must assign 1 to exactly half of the edges in G. On the right side, if some copy of $[x,y,1,0,0]$ contributes an x or y in some assignment, then less than half of its incident edges are assigned 1. To compensate, some other copy of $[x,y,1,0,0]$ must have more than half of its incident edges assigned 0, so it contributes a factor of 0. Thus $p(x,y)$ is a constant, so it is equal to any evaluation, including $(x,y) = (0,0)$ as claimed. ∎



$G = G_0$ G_1 G_2 G_ℓ

Figure 3.5: Some graphs obtained from an initial graph G in the proof of Lemma 3.3.1.

For all $x, y \in \mathbb{C}$, there is no holographic transformation between these two Holant problems. This is the first counterexample involving non-unary signatures in the Boolean domain (i.e. $\kappa = 2$) to the converse of Lemma 3.2.2, which provides a negative answer to a conjecture made by Xia in [143, Conjecture 4.1]. This result clearly generalizes to similarly defined signatures of even arity on the right.

3.3 Polynomial Interpolation

Valiant [125] initiated the study of counting problems by defining the class #P. Then to explain the apparent intractability of counting perfect matchings, he proved that this problem is #P-hard (under polynomial-time Turing reductions). Immediately after, he initiated the use of polynomial interpolation as a technique to obtain reductions between counting problems. This is a powerful tool in the study of counting problems that makes it possible to prove complexity dichotomy theorems.

Let me begin with an example from Valiant's paper [125, reduction 6 in the proof of Theorem 1].

Lemma 3.3.1. #PerfectMatching $\le_T$ #Matching

Proof. Let $G = (V, E)$ be a graph. We want to determine the number of perfect matchings in G assuming that we have an oracle to count matchings.

For integers $0 \le \ell \le n$, let G_ℓ be the following modification of G. For each vertex $v \in V$, we add a new vertex v_k for all $1 \le k \le \ell$. Then we add an edge between v and v_k for all $1 \le k \le \ell$. See Figure 3.5 for some examples of these graphs beginning with a specific graph G.

Let m_k be the number of matchings in G that omit k vertices. Then we can express the number

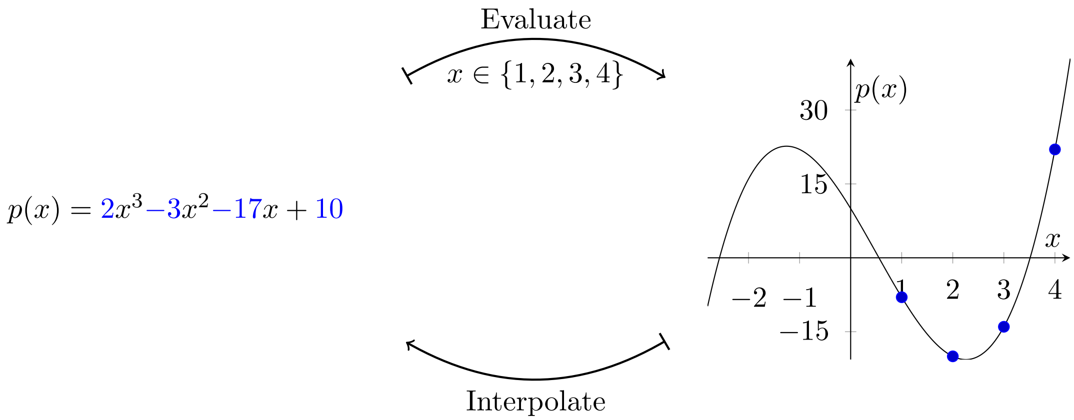


Figure 3.6: Interpolation is the inverse of evaluation.

of matchings in G_ℓ as

$$\sum_{k=0}^{n} (1+\ell)^k m_k = \#\textsc{Matching}(G_\ell).$$

This is because a matching M of size k in G can be extended to $(1+\ell)^k$ matchings in G_ℓ. Each vertex that is not matched by M has $1+\ell$ possibilities: it can remain unmatched or it can be matched with any of its new ℓ neighbors in G_ℓ. These choices are independent for each vertex, hence the exponent of k.

We collect these equations to form the linear system

$$\begin{bmatrix} (1+0)^0 & (1+0)^1 & \cdots & (1+0)^n \\ (1+1)^0 & (1+1)^1 & \cdots & (1+1)^n \\ \vdots & \vdots & \ddots & \vdots \\ (1+n)^0 & (1+n)^1 & \cdots & (1+n)^n \end{bmatrix} \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ m_n \end{bmatrix} = \begin{bmatrix} \#\textsc{Matching}(G_0) \\ \#\textsc{Matching}(G_1) \\ \vdots \\ \#\textsc{Matching}(G_n) \end{bmatrix}.$$

Using our oracle, we know the right side. On the left, the coefficient matrix is Vandermonde. It is invertible because the entries in the second column are distinct. Therefore, we can invert this matrix and solve for the unknown m_k's. Then m_0, the number of matchings in G that omit no vertices, is the number of perfect matchings G as desired. □

The word "polynomial" did not appear in this proof, so what makes it an example of polynomial interpolation? The polynomial is implicit; it is $p(x) = \sum_{k=0}^{n} m_k x^k$. Asking our oracle for the number of matchings in G_ℓ is like evaluating $p(x)$ at $1 + \ell$. Polynomial interpolation is the process of converting from points and their evaluations to the coefficients of the polynomial being evaluated (see Figure 3.6), which is what this proof did.

Since our $n+1$ evaluation points are distinct, we can recover the coefficients of $p(x)$. Given these coefficients, our reduction can proceed by computing any polynomial-time computable function of them. However, it is often the case that we are interested in some evaluation of the interpolated polynomial. In the proof of Lemma 3.3.1, we evaluated the interpolated polynomial at 0 to obtain $p(0) = m_0$.

Now for an example where the polynomial is quite explicit. The chromatic polynomial, which is denoted by $\chi(G; \lambda)$, is the unique polynomial satisfying $\chi(G; \lambda) = \#\lambda\text{-}\textsc{VertexColoring}(G)$ for all $\lambda \in \mathbb{N}$. That is, it is the unique polynomial that evaluates to the number of vertex colorings of G using at most λ colors when λ is a natural number. A nice exposition of the chromatic polynomial can be found in [82].

The following example of polynomial interpolation is a dichotomy theorem for the chromatic polynomial. The reduction we use comes from Linial [103] and the dichotomy was first explicitly stated in [86]. Let $\chi(\lambda)$ be the problem of evaluating $\chi(G; \lambda)$ on an input graph G.

Theorem 3.3.2. *Let $\lambda \in \mathbb{C}$. Then $\chi(\lambda)$ is #P-hard unless $\lambda \in \{0, 1, 2\}$, in which case, the problem is computable in polynomial time.*

Proof. If $\lambda = 0$, then $\chi(G, \lambda) = 0^0 = 1$ if G has no vertices and is 0 otherwise. If $\lambda = 1$, then $\chi(G, \lambda) = 1$ if G has no edges and is 0 otherwise. If $\lambda = 2$, then $\chi(G, \lambda) = 2^k$ if G is bipartite with k connected components and is 0 otherwise.

Now suppose $\lambda \notin \{0, 1, 2\}$. We reduce from $\chi(3)$, which is #P-hard (see, for example [103, Main Theorem, Case (6)] or [4, Proposition 5]). Let G be a graph with n vertices, and let K_t be the complete graph on t vertices. We use $G + K_t$ to denote the graph obtained from G by adding

K_t and all possible edges between the vertices of G and the vertices of K_t. Then clearly

$$\lambda(\lambda-1)\cdots(\lambda-t+1)\chi(G;\lambda-t)=\chi(G+K_t;\lambda).$$

when $\lambda \in \mathbb{N}$. Thus, it must also hold as an equation of polynomials when considering λ as an indeterminate. After rearranging, we have

$$\chi(G;\lambda-t)=\frac{1}{\lambda(\lambda-1)\cdots(\lambda-t+1)}\chi(G+K_t;\lambda). \tag{3.3.1}$$

If $\lambda \geq 3$ is an integer, then by setting $t=\lambda-3$, we can directly solve for $\chi(G;3)$ via (3.3.1). Otherwise, $\lambda \notin \mathbb{N}$. Then using (3.3.1), we can compute $\chi(G;\lambda-t)$ for all $0 \leq t \leq n$. From these evaluations, we can interpolate the coefficients of $\chi(G;\Lambda)$ and evaluate it at $\Lambda=3$. □

The heart of polynomial interpolation as a reduction technique is finding an equation like (3.3.1). On the left, we have the original graph G with some different evaluation point λ'. On the right, we have the original evaluation point λ with some different graph G'. The main question is what to pick for G'? It must be part of an infinite family because the degree of the polynomial being interpolated grows with the size of G. The sizes of the graphs in this family must not grow too quickly so that we can construct them in polynomial time. And finally, each graph in the family must be expressible using the original graph G for some distinct evaluation point λ'.

To this end, it is typically best to minimize the sizes of these graphs. Additional vertices and edges contribute terms to an equation like (3.3.1) that restrict the usefulness of the reduction. For example, one could say that the kth vertex and its incident edges added to G in the construction of $G+K_t$ (for $0 \leq k \leq t$) contribute a factor of $\frac{1}{\lambda-k+1}$ to the right side of (3.3.1). Because of these factors, the interpolation fails when $\lambda \in \mathbb{N}$. After adding the $(\lambda+1)$th vertex and its edges, the equation breaks down. The right side is no longer well defined because of a division by 0.

The proofs of both Lemma 3.3.1 and Theorem 3.3.2 involve interpolation of a single variable polynomial. After homogenizing, they become homogeneous polynomials in two variables. For such homogeneous polynomials of degree d, interpolation requires at least $d+1$ evaluations, and $d+1$ evaluations suffice iff, when viewed as length-two vectors, these $d+1$ points are pairwise linearly

Figure 3.7: Example construction from the proof of Lemma 3.3.3 with $\ell = 3$.

independent.

Of course a univariate polynomial defines a homogeneous polynomial in two variables. Some reductions between counting problems are accomplished via interpolation of homogeneous polynomials in more than two variables. An early example of this occurs in [113, Main Theorem, Case 1], which interpolates a homogeneous polynomial in three variables.

Lemma 3.3.3. #VertexCover $\leq_T$ #BipartiteVertexCover

Proof. Given a graph G with n vertices, we create a graph G_ℓ for every $1 \leq \ell \leq N = \binom{n+2}{2}$ in two steps as follows. First we perform ℓ-thickening on G to obtain a graph G'_ℓ, which replaces every edge of G with ℓ parallel copies. When we perform 4-stretching on G'_ℓ to obtain G_ℓ, which replaces each edge with a path of length 4. Since this stretch is by an even amount, the resulting graph is bipartite. See Figure 3.7 for an example of with $\ell = 3$.

Let c_{ijk} be the number of $S \subseteq V$ such that

- i edges have neither endpoint in S,
- j edges have exactly endpoint in S, and
- i edges have both endpoints in S.

Also let

$$p(x, y, z) = \sum_{\substack{i+j+k=n \\ i,j,k \geq 0}} x^i y^j z^k c_{ijk}.$$

Then #VertexCover$(G) = p(0, 1, 1)$.

A path of length 4 has ten vertex covers. We partition them based on the inclusion or exclusion of the endpoints of the path.

- There are 2 vertex covers when neither endpoint is in the vertex cover,
- there are 3 vertex covers when exactly one endpoint is in the vertex cover, and

- there are 5 vertex covers when both endpoints are in the vertex cover.

Thus,

$$\#\text{VertexCover}(G_\ell) = p(2^\ell, 3^\ell, 5^\ell). \tag{3.3.2}$$

This gives us a Vandermonde system that has full rank iff $2^i 3^j 5^k \neq 2^{i'} 3^{j'} 5^{k'}$ when $i + j + k = i' + j' + k' = n$ but $(i, j, k) \neq (i', j', k')$, which is clearly the case. Therefore, we can interpolate $p(x, y, z)$ and evaluate it at $p(0, 1, 1)$ to obtain the number of vertex covers of G. □

Remark. Let P_n be the path graph of length n. It is straightforward to verify that the vertex covers of P_4 can be partitioned into the triple $(2, 3, 5)$ as claimed in the proof. One can derive this triple by observing that the vertex covers of path graphs satisfy the recurrence relation

$$\#\text{VertexCover}(P_n) = \#\text{VertexCover}(P_{n-1}) + \#\text{VertexCover}(P_{n-2}),$$

the same recurrence relation as the Fibonacci numbers. Let F_n be the nth Fibonacci number. The (initial) triple of vertex covers for P_1 is $(0, 1, 1) = (F_0, F_1, F_2)$. Then the triple of vertex covers for P_k is (F_{k-1}, F_k, F_{k+1}). The proof uses $k = 4$ since this is the smallest positive even number for which the triple (F_{k-1}, F_k, F_{k+1}) gives a Vandermonde system of full rank.

For homogeneous polynomials of degree d in n variables, there is no known useful characterization of when some minimum number of points (namely $\binom{d+n-1}{n-1}$ points) can be used to interpolate such polynomials. However, this does not prevent us from using special collections of points for which we can prove that interpolation does succeed.

For reductions between Holant problems, we often use interpolation of homogeneous polynomials in more than two variables. See Section 11.3 for an explanation of how this is done. In particular, we say that a three numbers like 2, 3, and 5 satisfy the *lattice condition* (cf. Definition 11.3.3) if a Vandermonde system like that in (3.3.2) is full rank.

Chapter 4

Tractable Signatures

This chapter contains most of what is known about tractable Holant problems and the signatures that define them. Although I would say that we have a good understanding of the tractable cases, many questions still remain. I pose them throughout the chapter as they arise.

Given a problem P covered by a complexity dichotomy, it is natural to ask if P is easy. This question is called the *meta question* or the *decidability problem* of the dichotomy. If P is easy, a second question arises: what is the efficient algorithm that solves P? Some of the open problems in this chapter are about these meta questions. In this context, we only consider finite $\mathcal{F}$ since it is the input to the problem.

4.1 Product Type

As a symmetric signature in the Boolean domain, the Equality signature $=_n$ of arity n is denoted by $[1, 0, \ldots, 0, 1]$ with $n-1$ zeros. The Holant of a signature grid that only uses Equality signatures is easy to compute. Each connected component contributes a factor of 2. Let's generalize these signatures while still being able to easily compute the Holant.

One way to generalize them is to allow weights. For $a, b \in \mathbb{C}$, we call the signature $[a, 0, \ldots, 0, b]$ a generalized Equality signature. Consider a signature grid that only uses such signatures. Each connected component of size k contributes a factor of $a^k + b^k$ if each Equality signature had the same weights. A similar (but more complicated) expression exists if the weights were to vary.

Consider a different generalization. In addition to all of the EQUALITY signatures, we also allow the binary DISEQUALITY signature $\neq_2$. Now what does each connected component contribute? Well, if there is any cycle containing an odd number of $\neq_2$, then the contribution is 0. Otherwise, the contribution is again 2. Does this mean that we must inspect every cycle? No, and its a good thing too, because that would take exponential time.

The key observation is that among the exponentially many edge assignments, there are at most two that could contribute a nonzero value. Furthermore, we can use a propagation algorithm to efficiently determine if these two assignments exists and find them when they do. Pick any edge and fix its assignment to 0. Then in order for the incident signatures to contribute nonzero values, the assignments to all adjacent edges is also fixed. We recurse on each fixed edge. If we encounter an edge that is fixed to different assignments, then we have a contradiction and the Holant is 0. Otherwise, we have determined a consistent assignment for the connected component containing the edge we initially picked. The other consistent assignment is the complement of this one, but we could also find it from first principles by fixing the assignment of the initial edge to 1 and running the propagation algorithm again.

There are three other symmetric signatures that we could add: $[0, 0]$, $[1, 0]$, and $[0, 1]$. If $[0, 0]$ is used, then there are no assignments and the Holant is 0. If $[1, 0]$ or $[0, 1]$ is used, then any connected component in which they appear has at most one assignment that could contribute a nonzero value. This only makes our life easier.

The propagation algorithm also works for the following types of signatures that may have weights and may not be symmetric.

Definition 4.1.1. Let f be a signature of arity n over the Boolean domain. Then f is a *generalized equality* signature if $n = 1$ and f is $[0, 0]$, $[1, 0]$, $[0, 1]$ (up to scale), or $n \geq 2$ and

$$\exists \mathbf{x} \in \{0,1\}^n, \quad \forall \mathbf{y} \in \{0,1\}^n, \quad f(\mathbf{y}) = 0 \iff \mathbf{y} \notin \{\mathbf{x}, \overline{\mathbf{x}}\}.$$

The last condition says that the support set of f contains precisely two complementary indices. We generalize this once further to signatures over larger domain sizes.

Definition 4.1.2. For any integer $\kappa \geq 2$, let f be a signature of arity n over the domain $[\kappa]$. Then f is a *generalized equality* signature if $n = 1$ and $|\operatorname{support}(f)| \leq 1$, or $n \geq 2$ and

$$\forall i \in [n], \quad \forall x \in [\kappa], \quad |\operatorname{support}(f_{x_i=x})| = 1,$$

where $f_{x_i=x}$ is the restriction of f to inputs with its ith input x_i fixed to x. We use $\mathscr{E}$ to denote the set of all generalized EQUALITY signatures.

Remark. Definition 4.1.1 and Definition 4.1.2 are written so that all signatures in $\mathscr{E}$ are irreducible according to Definition 5.3 in [31].

The last condition says that the support set of f contains precisely κ "disjoint" indices. They are disjoint in the sense that each index differs from all others in every position. It is easy to see that the tensor rank of each signature in $\mathscr{E}$ is at most the domain size κ over which it is defined.

For these signatures with a larger domain, we only need to modify the propagation algorithm slightly. After picking the first edge, we consider fixing it each element in $[\kappa]$. In each connected component, there are at most κ assignments that could contribute a nonzero value, and the propagation algorithm will find any such assignment that exist. This gives the following result.

Theorem 4.1.3. *Suppose $\kappa \geq 2$ is the domain size. Then* $\operatorname{Holant}_k(\mathscr{E})$ *is computable in polynomial time.*

We are supposed to be defining the product-type signatures, but the word "product" has yet to occur. There is one more generalization, and it uses the term "product".

Definition 4.1.4. A signature is of *product type* if it is the tensor product of signatures from $\mathscr{E}$ with any ordering of inputs. We use $\mathscr{P}$ to denote the set of all product-type signatures.

In other words, $\mathscr{P}$ is the closure of $\mathscr{E}$ under tensor products and reordering of inputs. Tensor products and reordering of inputs using signatures from a set $\mathcal{F}$ is a special case of an $\mathcal{F}$-gate. Specifically, it is the special case when there are no internal edges. it is the special case with no internal edges. Thus, we obtain the following corollary by combining Theorem 4.1.3 and Lemma 3.1.1.

Corollary 4.1.5. *Suppose $\kappa \geq 2$ is the domain size. Then* $\operatorname{Holant}_k(\mathscr{P})$ *is computable in polynomial time.*

By combining Corollary 4.1.5 and Lemma 3.2.4, we obtain another corollary.

Corollary 4.1.6. *Let $\mathcal{F}$ be any set of complex-valued signatures over a domain of size κ. If $\mathcal{F}$ is $\mathscr{P}$-transformable, then* $\operatorname{Holant}_k(\mathcal{F})$ *is computable in polynomial time.*

Definition 3.3 in [53] contains the original definition of $\mathscr{P}$ (over the Boolean domain) and uses a slightly different notion of the term "product". Now we list the symmetric signatures in $\mathscr{P}$ over the Boolean domain.

Proposition 4.1.7 (Lemma A.1 in [81])**.** *Let $f \in \mathscr{P}$ be a symmetric signature. Then there exists $a, b \in \mathbb{C}$ and $n \in \mathbb{Z}^+$ such that f takes one of the following forms:*

1. $[a, b]^{\otimes n}$;
2. $[0, a, 0]$;
3. $[a, 0, \ldots, 0, b] = a[1, 0]^{\otimes n} + b[0, 1]^{\otimes n}$.

Let $\mathcal{F}$ be a finite set of complex-valued signatures over the Boolean domain. Then [34] gave a polynomial-time algorithm to decide if $\mathcal{F}$ is $\mathscr{P}$-transformable. When it is, the algorithm also finds a corresponding transformation. Further suppose that $\mathcal{F}$ only contains symmetric signatures, and these signatures are given in their exponentially more succinct representation, Then [34] also gave a polynomial-time algorithm to decide if $\mathcal{F}$ is $\mathscr{P}$-transformable. When it is, the algorithm also finds a corresponding transformation. Many questions remain about deciding $\mathcal{P}$-transformability. I state the central ones as open problems.

Open Problem 4.1.8. Let $\mathcal{F}$ and $\mathcal{G}$ be finite sets of complex-valued signatures over a domain of size $\kappa \geq 2$.

1. Is there an algorithm to decide if there exists a $T \in \mathbf{GL}_\kappa(\mathbb{C})$ such that $\mathcal{G}T \subseteq \mathscr{P}$ and $\mathcal{F} \subseteq T\mathscr{P}$?
2. If so, is there an algorithm to find such a T?
3. In either case, if an algorithm exists, then does a polynomial-time algorithm exist?

The same questions when $\mathcal{F}$ and $\mathcal{G}$ contain only symmetric signatures, and these signatures are given in their exponentially more succinct representation.

To repeat, the results in [34] answer these questions in the affirmative in the special case that $\kappa = 2$ and $\mathcal{G} = \{=_2\}$.

4.2 Affine

Definition 4.2.1. Let f by a signature of arity n with inputs $x_1, \ldots, x_n$ over the Boolean domain. Then f is *affine* if it has the form

$$\lambda \cdot \chi_{A\mathbf{x}=\mathbf{0}} \cdot i^{q(\mathbf{x})},$$

where $\lambda \in \mathbb{C}$, $\mathbf{x} = (x_1, x_2, \ldots, x_n, 1)^\intercal$, A is a matrix over $\mathbb{Z}_2$, χ is a 0-1 indicator function such that $\chi_{A\mathbf{x}=\mathbf{0}}$ is 1 iff $A\mathbf{x} = \mathbf{0}$, and $q \in \mathbb{Z}_4[\mathbf{x}]$ is a quadratic polynomial with even coefficients on cross terms. We use $\mathscr{A}$ to denote the set of all affine signatures.

Of course the number of columns in the matrix A must be $n + 1$ but there is no restriction on the number of rows. It is permissible that A is the all-zero matrix so that $\chi_{A\mathbf{x}=\mathbf{0}} = 1$ holds for all $\mathbf{x}$. The name affine comes from the fact that solutions to $A\mathbf{x} = \mathbf{0}$ (and thus the support of f) form an affine subspace.

The next result shows how the nontrivial symmetric affine signatures have a compact expression. When viewed as tensors, this shows that they have tensor rank (at most) 2.

Proposition 4.2.2. *Let $f \in \mathscr{A}$ be a unary signature or a non-degenerate symmetric signature. Then f is an element in one (or possibly more if* $\operatorname{arity}(f) \le 2$*) of the following three sets:*

$$\begin{aligned}
\mathscr{F}_1 &= \left\{\lambda\left([1,0]^{\otimes k} + i^r[0,\quad 1]^{\otimes k}\right) \mid \lambda \in \mathbb{C},\ \ k \in \mathbb{Z}^+,\ \ r \in \{0,1,2,3\}\right\};\\
\mathscr{F}_2 &= \left\{\lambda\left([1,1]^{\otimes k} + i^r[1,-1]^{\otimes k}\right) \mid \lambda \in \mathbb{C},\ \ k \in \mathbb{Z}^+,\ \ r \in \{0,1,2,3\}\right\};\\
\mathscr{F}_3 &= \left\{\lambda\left([1,\ i]^{\otimes k} + i^r[1,-i]^{\otimes k}\right) \mid \lambda \in \mathbb{C},\ \ k \in \mathbb{Z}^+,\ \ r \in \{0,1,2,3\}\right\}.
\end{aligned}$$

Proposition 4.2.3. *Let $f \in \mathscr{A}$ be a unary signature different from* $[0,0]$ *or a non-degenerate symmetric signature. If the first nonzero entry in f is* 1*, then f takes one of the following forms:*

1. $[1, 0, \ldots, 0, \pm 1]$; $(\mathscr{F}_1, r = 0, 2)$
2. $[1, 0, \ldots, 0, \pm i]$; $(\mathscr{F}_1, r = 1, 3)$
3. $[1, 0, 1, 0, \ldots, 0 \text{ or } 1]$; $(\mathscr{F}_2, r = 0)$
4. $[1, -i, 1, -i, \ldots, 1 \text{ or } -i]$; $(\mathscr{F}_2, r = 1)$
5. $[0, 1, 0, 1, \ldots, 0 \text{ or } 1]$; $(\mathscr{F}_2, r = 2)$
6. $[1, i, 1, i, \ldots, i \text{ or } 1]$; $(\mathscr{F}_2, r = 3)$
7. $[1, 0, -1, 0, 1, 0, -1, 0, \ldots, 0 \text{ or } 1 \text{ or } -1]$; $(\mathscr{F}_3, r = 0)$
8. $[1, 1, -1, -1, 1, 1, -1, -1, \ldots, 1 \text{ or } -1]$; $(\mathscr{F}_3, r = 1)$
9. $[0, 1, 0, -1, 0, 1, 0, -1, \ldots, 0 \text{ or } 1 \text{ or } -1]$; $(\mathscr{F}_3, r = 2)$
10. $[1, -1, -1, 1, 1, -1, -1, 1, \ldots, 1 \text{ or } -1]$. $(\mathscr{F}_3, r = 3)$

Because of Proposition 4.2.2, I used to wonder if all affine signatures had tensor rank (at most) 2. This is not the case. The affine signature f' introduced below (before Lemma 4.2.14) has tensor rank at least 4. This follows from [96, Exercise 2.6.6.3] since its signature matrix has full rank. I still wonder though if there is some connection between the complexity of a signature f (equivalently, the complexity of contracting tensor networks only involvoing f) and the tensor rank of f. For example, Håstad [76] proved that determining the rank of an aribtray tensor is NP-hard. Maybe determining the rank of a tensor in $\mathscr{A}$ is easier.

Open Problem 4.2.4. Given an element of $\mathscr{A}$, what is the complexity of determining its tensor rank? Can this be done in polynomial time? If so, can one do better and give a classification in terms of tensor rank or border rank?

After normalizing, it is easy to count the symmetric affine signatures that are not identically zero. For arity 1, there are six such signatures ($[1, 0], [0, 1], [1, \pm 1], [1, \pm i]$). For larger arities, a degenerate symmetric affine signature that is not identically zero is a tensor power of one of these six. Then for arity 2, there are nine non-degenerate symmetric affine signatures up to scale ($[1, 0, \pm 1], [1, 0, \pm i], [1, \pm i, 1], [1, 1, \pm 1], [0, 1, 0]$). For arity at least 3, there are twelve non-degenerate symmetric affine signatures up to scale, which are listed above.

For affine signatures that are not necessary symmetric, the situation is more complicated. It is easy to count the possible quadratic polynomials with constant term 0 (there are $4^n 2^{\binom{n}{2}}$ possibilities), and the number possible affine supports is also known [98] (there are $\sum_{k=0}^{n} 2^{n-k} \prod_{i=0}^{k-1} \frac{2^n-2^i}{2^k-2^i}$ possibilities), but I don't know how these two quantities "intersect" in Definition 4.2.1.

Open Problem 4.2.5. Determine the number of affine signatures up to scale over the Boolean domain as a function of the arity.

Table 4.1: The number of affine signatures over the Boolean domain that are not identically zero as a function of arity (according to my computer calculations). With an AMD Phenom II processor and 3 GB of RAM, my computer outputs the last entry in just under 10 minutes. I am unable to compute the next entry due to lack of memory (since it is easier to write the algorithm that keeps everything in memory at once).

arity	number
1	6
2	60
3	1080
4	36720

If it helps, one can restrict to signatures that are irreducible according to Definition 5.3 in [31]. I have written code in Mathematica to compute the affine signatures over the Boolean domain that are not identically zero and with the first nonzero entry normalized to 1. The numbers I get are given in Table 4.1. Neither this sequence nor this sequence with each number increased by 1 (to account for the identically-zero signature) are in The On-Line Encyclopedia of Integer Sequences.

The affine signatures over the Boolean domain were originally defined in Definition 3.1 of [53]. This paper also proves their tractability.

Theorem 4.2.6 (Theorem 4.1 in [53])**.** Holant$_2(\mathscr{A})$ *is computable in polynomial time.*

By combining theorem 4.2.6 and Lemma 3.2.4, we obtain the following corollary.

Corollary 4.2.7. *Let $\mathcal{F}$ be any set of complex-valued signatures over the Boolean domain. If $\mathcal{F}$ is $\mathscr{A}$-transformable, then* Holant$_k(\mathcal{F})$ *is computable in polynomial time.*

Definition 3.1 of [53] (for the affine signatures over the Boolean domain) is different from but equivalent to Definition 4.2.1. My reason for using this alternative definition is to highlight the similarity with the tractability result in [21]. After giving a definition for affine signatures over a domain of size 3, I show how to use [21] to prove their tractability.

Definition 4.2.8. Let f by a signature of arity n with inputs $x_1, \ldots, x_n$ over the domain $\mathbb{Z}_3$. Then f is *affine* if it has the form

$$\lambda \cdot \chi_{A\mathbf{x}=\mathbf{0}} \cdot e^{\frac{2\pi i}{3} q(x)},$$

where $\lambda \in \mathbb{C}$, $\mathbf{x} = (x_1, x_2, \ldots, x_n, 1)^{\intercal}$, A is a matrix over $\mathbb{Z}_3$, χ is a 0-1 indicator function such that $\chi_{A\mathbf{x}=\mathbf{0}}$ is 1 iff $A\mathbf{x} = \mathbf{0}$, and $q \in \mathbb{Z}_3[\mathbf{x}]$ is a quadratic polynomial. We still use $\mathscr{A}$ to denote the set of all affine signatures.

Lemma 4.2.9. $\operatorname{Holant}_3(\mathscr{A})$ *is computable in polynomial time.*

Proof. Given an instance of $\operatorname{Holant}_3(\mathscr{A})$ with m edges, the output can be expressed as the summation of a single function $F(\mathbf{x}) = \chi_{A\mathbf{x}=\mathbf{0}} \cdot e^{\frac{2\pi i}{3} q(\mathbf{x})}$, where $\mathbf{x} = (x_1, \ldots, x_m)$. This is because $\mathscr{A}$ is closed under multiplication. In polynomial time, we can solve the linear system $A\mathbf{x} = \mathbf{0}$ over $\mathbb{Z}_3$ to determine if it is feasible. If it is infeasible, then F is identically zero, so the output is simply 0.

Otherwise, the linear system is feasible (including possibly vacuous). Without loss of generality, we can assume that $y_1, \ldots, y_n \in \{x_1, \ldots, x_m\}$ are independent variables over $\mathbb{Z}_3$ while all others are dependent variables, where $0 \le n \le m$. Each dependent variable can be expressed by an affine linear form of $y_1, \ldots, y_n$. We can substitute for all dependent variables in $q(\mathbf{x})$, which gives a new quadratic polynomial $q'(\mathbf{y})$, where $\mathbf{y} = (y_1, \ldots, y_n)$. Thus, we have

$$\sum_{x_1,\ldots,x_m \in \mathbb{Z}_3} \chi_{A\mathbf{x}=\mathbf{0}} \cdot e^{\frac{2\pi i}{3} q(x_1,\ldots,x_m)} = \sum_{y_1,\ldots,y_n \in \mathbb{Z}_3} e^{\frac{2\pi i}{3} q'(y_1,\ldots,y_n)}. \tag{4.2.1}$$

Then the right side of (4.2.1) is computable in polynomial time by Theorem 1 in [53]. □

The Equality signatures over the $\mathbb{Z}_3$ domain are examples of affine signatures. Their support is defined by equating all variables, and their quadratic polynomial is the constant polynomial 0. We give another, less-trivial example. Consider the signature

$$f(x, y, z) = \begin{cases} a & x = y = z \\ c & x \neq y \neq z \neq x \\ 0 & \text{otherwise,} \end{cases} \tag{4.2.2}$$

where $a^3 = c^3$. If $a = c = 0$, then this example is trivial, so assume otherwise. Then the support of f is the affine subspace of $\mathbb{Z}_3$ defined by $x + y + z = 0$. Let $\omega = e^{\frac{2\pi i}{3}}$ and let $q_{\frac{c}{a}}(x, y, z) = \lambda_{\frac{c}{a}}(xy + xz + yz)$ be a quadratic polynomial, where $\lambda_1 = 0$, $\lambda_\omega = 2$, and $\lambda_{\omega^2} = 1$. Then $a\omega^{q_{\frac{c}{a}}(x,y,z)}$ agrees with f when $x + y + z = 0$. This gives the following corollary.

Corollary 4.2.10. *Let $a, c \in \mathbb{C}$. If f is defined as in (4.2.2), then* $\operatorname{Holant}_3(f)$ *is computable in polynomial time.*

By combining Corollary 4.2.10 with Lemma 3.2.2, we have another corollary.

Corollary 4.2.11. *Let $a, c \in \mathbb{C}$. Suppose $T \in \mathbf{O}_3(\mathbb{C})$. If f is defined as in (4.2.2), then* $\operatorname{Holant}_3(T^{\otimes 3} f)$ *is computable in polynomial time.*

Even though I have given a definition for affine signatures over a domain of size 3, this definition has yet to be thoroughly tested.

Open Problem 4.2.12. Over a domain of size $\kappa \geq 3$, determine the "right" definition of the affine signatures.

The "right" definitions are the ones that fit like a glove in the statement of a powerful dichotomy theorem. For the Boolean domain, there are plenty of dichotomy theorems to believe that Definition 4.2.1 is the "right" one. The dichotomy in [53] is probably the best example. Definition 4.2.8 is used to state a dichotomy theorem in Chapter 11. However, only the special case given in Corollary 4.2.11 appears in the statement of that dichotomy theorem because only signatures of a rather restricted form are considered.

Here is my guess as to the "right" definition of the affine signatures over a domain of size κ. If κ is not a power of 2, then affine signatures are of the form $\lambda \cdot \chi_{A\mathbf{x}=\mathbf{0}} \cdot e^{\frac{2\pi i}{\kappa} q(x)}$, where A is a matrix over $\mathbb{Z}_\kappa$ and $q \in \mathbb{Z}_\kappa[\mathbf{x}]$ is a quadratic polynomial. If κ is a power of 2, then affine signatures are of the form $\lambda \cdot \chi_{A\mathbf{x}=\mathbf{0}} \cdot e^{\frac{2\pi i}{2\kappa} q(x)}$, where A is a matrix over $\mathbb{Z}_\kappa$ and $q \in \mathbb{Z}_\kappa[\mathbf{x}]$ is a quadratic polynomial with even coefficients on cross terms.

Open Problem 4.2.13. Given some definition of the affine signatures over a domain of size $\kappa \geq 3$ (preferably the "right" one), determine the number of affine signatures up to scale over this domain size as a function of the arity.

Table 4.2: The number of affine signatures over a domain of size 3 (as defined by Definition 4.2.8) that are not identically zero as a function of arity (according to my computer calculations). The last line took just over 50 minutes to compute. Also unable to compute the next entry due to lack of memory.

arity	number
1	6
2	120
3	8400

I am not positive that Definition 4.2.8 is the "right" definition of the affine signatures over a domain of size 3. Nonetheless, I have written code in Mathematica to compute the ones that are not identically zero and with the first nonzero entry normalized to 1. The numbers I get are given in Table 4.2. Again, neither this sequence nor this sequence with each number increased by 1 (to account for the identically-zero signature) are in The On-Line Encyclopedia of Integer Sequences.

Now we give two examples of asymmetric affine signatures over the Boolean domain. Let

$$f'(x_1, x_2, y_1, y_2) = \begin{cases} 1 & (x_1, x_2) = (y_1, y_2), \\ -1 & (x_1, x_2) \neq (y_1, y_2). \end{cases} \tag{4.2.3}$$

This signature is affine over the Boolean domain since $f'(x_1, x_2, y_1, y_2) = (-1)^{q(x_1,x_2,y_1,y_2)}$, where q is the quadratic polynomial

$$q(x_1, x_2, y_1, y_2) = x_1 + x_2 + y_1 + y_2 + x_1x_2 + y_1y_2 + x_1y_2 + x_2y_1. \tag{4.2.4}$$

Let

$$g'(x_1, x_2, y_1, y_2, z_1, z_2) = \begin{cases} 1 & (x_1, x_2) = (y_1, y_2) = (z_1, z_2), \\ 0 & \text{otherwise.} \end{cases} \tag{4.2.5}$$

This signature is affine over the Boolean domain since its support is defined by the affine equations $(x_1, x_2) = (y_1, y_2) = (z_1, z_2)$ and its quadratic polynomial is the constant polynomial 0.

Alternatively, we can view $f'(x_1, x_2, y_1, y_2)$ as a signature $f(x, y)$ of arity 2 over a domain of size 4, where $x = (x_1, x_2)$ and $y = (y_1, y_2)$. Then

$$f(x, y) = \begin{cases} 1 & x = y, \\ -1 & x \neq y. \end{cases} \tag{4.2.6}$$

Similarly, we can view $g'(x_1, x_2, y_1, y_2, z_1, z_2)$ as a signature $g(x, y, z)$ of arity 3 over a domain of size 4, where $x = (x_1, x_2)$, $y = (y_1, y_2)$, and $z = (z_1, z_2)$. Then $g = (=_3)$. This gives the following tractability result, which corresponds to a tractable case in the dichotomy theorem presented in Chapter 11.

Lemma 4.2.14. *If f is defined as in (4.2.6), then* $\operatorname{Holant}_4(f \mid =_3)$ *is computable in polynomial time.*

Proof. Let f' and g' be defined as in (4.2.3) and (4.2.5) respectively. Then by Lemma 3.1.6, we have

$$\operatorname{Holant}_4(f \mid =_3) \leq_T \operatorname{Holant}_2(f' \mid g').$$

By ignoring the bipartite restriction, we further have

$$\operatorname{Holant}_2(f' \mid g') \leq_T \operatorname{Holant}_2(f', g').$$

As demonstrated above, f' and g' are both affine signatures over the Boolean domain, so we are done by Theorem 4.2.6. □

This highlights another difficulty with counting affine signatures over higher domains. I don't think that one should count f and $g = (=_3)$ as affine signatures (over a domain of size 4) when trying to answer Open Problem 4.2.13. Instead, I would say that they are being simulated by affine signatures over a domain of a different size (namely a domain of size 2). However, the Equality signatures of all arities should be affine signatures over domains of all sizes, so $g = (=_3)$ is certainly an affine signature over a domain of size 4 for the "right" definition of affine signatures.

Let $\mathcal{F}$ be a finite set of complex-valued signatures over the Boolean domain. Then [34] gave a polynomial-time algorithm to decide if $\mathcal{F}$ is $\mathscr{A}$-transformable. When it is, the algorithm also finds a corresponding transformation. Further suppose that $\mathcal{F}$ only contains symmetric signatures, and these signatures are given in their exponentially more succinct representation, Then [34] also gave a polynomial-time algorithm to decide if $\mathcal{F}$ is $\mathscr{A}$-transformable. When it is, the algorithm also finds a corresponding transformation. Many questions remain about deciding $\mathcal{A}$-transformability. I state the central ones as open problems.

Open Problem 4.2.15. Let $\mathcal{F}$ and $\mathcal{G}$ be finite sets of complex-valued signatures over a domain of size $\kappa \geq 2$.

1. Is there an algorithm to decide if there exists a $T \in \mathbf{GL}_\kappa(\mathbb{C})$ such that $\mathcal{G}T \subseteq \mathscr{A}$ and $\mathcal{F} \subseteq T\mathscr{A}$?
2. If so, is there an algorithm to find such a T?
3. In either case, if an algorithm exists, then does a polynomial-time algorithm exist?

The same questions when $\mathcal{F}$ and $\mathcal{G}$ contain only symmetric signatures, and these signatures are given in their exponentially more succinct representation.

To repeat, the results in [34] answer these questions in the affirmative in the special case that $\kappa = 2$ and $\mathcal{G} = \{=_2\}$.

4.3 Matchgate

Matchgates were introduced by Valiant [128, 127, 129] to give polynomial-time algorithms for a number of counting problems over planar graphs.[1] A matchgate is simply a gadget used to reduce some counting problem to that of counting weighted perfect matchings. Over planar graphs, the latter problem is computable in polynomial by Kasteleyn's algorithm [89, 88].[2]

Let me give a simple example to explain this idea. The unweighted perfect matching signature

[1]Valiant's matchgates in [128] sometimes contained edge crossings. However, this doesn't mean that Valiant's argument in [128] was able to simulate quantum circuits with edge crossings. Indeed, Valiant required that the quantum circuits only had "nearest neighbor interactions", which means that there are no edge crossings. To allow edge crossings in matchgates is an attempt to increase their expressiveness. However, edge crossings do not increase the expressiveness of matchgates. See [28], especially Section 5, for a proof of this.

[2]This algorithm is also called the FKT algorithm because it generalizes previous work by both Kasteleyn [87] and as well as Temperley and Fisher [120]. These two works both proved that the dimer model (i.e. counting perfect matchings) over two-dimensional lattices is "exactly solvable", which essentially means that the number of perfect matchings over such graphs has a mathematical expression that only involves polynomially many fundamental operations. The motivation for considering the two-dimensional (weighted) dimer model is that it expresses, after a holographic transformation, the Ising model as a special case.

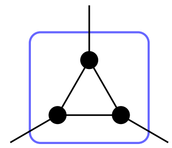


Figure 4.1: Triangle gadget.

of arity 3 is $[0,1,0,0]$. The problem Pl-Holant($[0,1,0,0]$) is to count perfect matchings over planar 3-regular graphs. Of course this problem is computable in polynomial time. Then by Lemma 3.1.1, Pl-Holant(f) is also computable in polynomial time for any $\{[0,1,0,0]\}$-gate f.

Consider the gadget in Figure 4.1. We assign $[0,1,0,0]$ to all three vertices. Let g be the signature of the resulting gadget. Then what is g? By the symmetry of the gadget, it is easy to see that g is a symmetric signature, so we can write $g = [g_0, g_1, g_2, g_3]$. Because there is an odd number of vertices in this gadget, we have $g_0 = g_2 = 0$. If exactly one vertex is matched externally (which means that one dangling edge is assigned 1 and the other two dangling edges are assigned 0), then there is one internal perfect matching (of size 1), so $g_1 = 1$. If all three vertices are matched externally (which means that all three dangling edges are assigned 1), then there is one internal perfect matching (the empty matching of size 0), so $g_3 = 1$ as well. Alternatively, assigning 0's to all internal edges is the only internal assignment such that the signature $[0,1,0,0]$ assigned to each of the three vertices does not tribute a factor of 0. For this internal assignment, the three vertices each contributes a factor of 1. Thus, $g = [0,1,0,1]$. For this reason, g is called a matchgate signature.

Definition 4.3.1. Let $\mathscr{W}$ be the set of all weighted matching signatures (that is, signatures with output 0 on any input with Hamming weight *different from* 1). Then f is a *matchgate signature* if it is the signature of some planar $\mathscr{W}$-gate. We use $\mathscr{M}$ to denote the set of all matchgate signatures.

Said another way, the set $\mathscr{M}$ of matchgate signatures is the closure of $\mathscr{W}$ under the construction of all possible planar $\mathscr{W}$-gates. Therefore, by Lemma 3.1.1 and the FKT algorithm, we have the following result.

Theorem 4.3.2. Pl-Holant($\mathscr{M}$) *is computable in polynomial time.*

However, Definition 4.3.1 is not useful in checking if a given signature is a matchgate signature. The work of Cai et al. [25, 26, 40] developed a theory to give a useful characterization of matchgate signatures. See [28] for a self-contained account of this theory.

Here is an explicit list of the symmetric matchgate signatures.

Proposition 4.3.3. *Let $f \in \mathscr{M}$ be a symmetric signature. Then there exists $a, b \in \mathbb{C}$ and $n \in \mathbb{N}$ such that f takes one of the following forms:*

1. $[a^n, 0, a^{n-1}b, 0, \ldots, 0, ab^{n-1}, 0, b^n]$ *(of arity $2n \geq 2$);*
2. $[a^n, 0, a^{n-1}b, 0, \ldots, 0, ab^{n-1}, 0, b^n, 0]$ *(of arity $2n+1 \geq 1$);*
3. $[0, a^n, 0, a^{n-1}b, 0, \ldots, 0, ab^{n-1}, 0, b^n]$ *(of arity $2n+1 \geq 1$);*
4. $[0, a^n, 0, a^{n-1}b, 0, \ldots, 0, ab^{n-1}, 0, b^n, 0]$ *(of arity $2n+2 \geq 2$).*

In the last three cases with $n = 0$, the signatures are $[1, 0]$, $[0, 1]$, and $[0, 1, 0]$. Any multiple of these is also a matchgate signature.

Notice that every other entry is 0. This is a simple consequence of the fact that perfect matchings contain an even number of vertices. The parity of a matchgate signature is even (resp. odd) if its support is on entries of even (resp. odd) Hamming weight. More generally, a signature is said to satisfy the *parity condition* if all entries of even Hamming weight are 0 or if all entries of odd Hamming weight are 0. The nonzero entries of the symmetric matchgate signatures form a geometric progression.

Another useful way to view the symmetric signature in $\mathscr{M}$ is via a low tensor rank decomposition. Such expressions make it easier to apply a holographic transformation. These expressions are similar to those in Proposition 4.2.2 for the affine signatures over the Boolean domain. To state these low rank decompositions, we use the following definition.

Definition 4.3.4. Let S_n be the symmetric group of degree n. Then for positive integers t and n with $t \leq n$ and unary signatures $v, v_1, \ldots, v_{n-t}$, we define

$$\operatorname{Sym}_n^t(v; v_1, \ldots, v_{n-t}) = \sum_{\pi \in S_n} \bigotimes_{k=1}^{n} u_{\pi(k)},$$

where the ordered sequence $(u_1, u_2, \ldots, u_n) = (\underbrace{v, \ldots, v}_{t \text{ copies}}, v_1, \ldots, v_{n-t})$.

Proposition 4.3.5. *Let $f \in \mathscr{M}$ be a symmetric signature of arity n. Then there exist $a, b, \lambda \in \mathbb{C}$ such that f takes one of the following forms:*

1. $[a,b]^{\otimes n} + [a,-b]^{\otimes n} = \begin{cases} 2[a^n, 0, a^{n-2}b^2, 0, \ldots, 0, b^n] & n \text{ is even,} \\ 2[a^n, 0, a^{n-2}b^2, 0, \ldots, 0, ab^{n-1}, 0] & n \text{ is odd;} \end{cases}$
2. $[a,b]^{\otimes n} - [a,-b]^{\otimes n} = \begin{cases} 2[0, a^{n-1}b, 0, a^{n-3}b^3, 0, \ldots, 0, ab^{n-1}, 0] & n \text{ is even,} \\ 2[0, a^{n-1}b, 0, a^{n-3}b^3, 0, \ldots, 0, b^n] & n \text{ is odd;} \end{cases}$
3. $\lambda \operatorname{Sym}_n^{n-1}([1,0];[0,1]) = [0, \lambda, 0, \ldots, 0]$;
4. $\lambda \operatorname{Sym}_n^{n-1}([0,1];[1,0]) = [0, \ldots, 0, \lambda, 0]$.

The first two items are decompositions of tensor rank (at most) 2. The last two items are decompositions of tensor rank n. Furthermore, one can check that these signatures do indeed have tensor rank n. However, they have border rank 2. (These facts follow from slight modifications to the example and the reasoning for it in [96, Subsection 2.4.5]). I pose the corresponding question for $\mathscr{M}$ as in Open Problem 4.2.4.

Open Problem 4.3.6. Given an element of $\mathscr{M}$, what is the complexity of determining its tensor rank? Can this be done in polynomial time? What about the same questions for border rank? If so, can one do better and give a classification in terms of tensor rank or border rank?

By combining Theorem 4.3.2 with Lemma 3.2.4, we have the following corollary.

Corollary 4.3.7. *Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures in Boolean variables. If $\mathcal{F}$ is $\mathscr{M}$-transformable, then* Pl-Holant$(\mathcal{F})$ *is computable in polynomial time.*

In contrast with the product-type and affine signatures, there doesn't seem to be a generalization of $\mathscr{M}$ to larger domain sizes. Of course the signatures in $\mathscr{M}$ can simulate others defined over a larger domain size by a domain bundling argument like Corollary 3.1.5. But the following questions remain.

Open Problem 4.3.8. For a domain of size $\kappa \geq 3$, what #CSPs are hard over general graphs but tractable over planar graphs? Are any of these problems tractable for some reason other than by reduction to Boolean domain Holant problems defined by signatures in $\mathscr{M}$?

Matchgates were not studied in isolation. The theory of matchgates was developed in conjunction with the theory of holographic reductions (cf. Subsection 3.2.1). The initial holographic reductions reduced to Holant problems using certain matchgate signatures over planar graphs (either by Valiant [130, 131, 133] or Cai et al. [42, 43, 44, 41]). This resulted in an abundance of counting problems over planar graphs with new polynomial-time algorithms. Although the phrase "holographic algorithm" is meant to refer to any polynomial-time algorithm that uses a holographic reduction, it became synonymous with polynomial-time algorithms with a holographic reduction to matchgates (or more preciously, to Holant problems using matchgate signatures over planar graphs).

Holographic reductions to Holant problems with product-type or affine signatures also exist [52, 47, 35, 80, 30, 34, 27]. To prove the dichotomy theorems just referenced, it is not necessary to obtain characterizations of the $\mathscr{A}$-, $\mathscr{P}$-, or $\mathscr{M}$-transformable signatures. The condition of transformability arises naturally in the proof. However, the characterizations of transformability are useful in answering the decidability question of these dichotomy theorems. The current understanding of $\mathscr{A}$- and $\mathscr{P}$-transformability are discussed in the text surrounding Open Problem 4.1.8 and Open Problem 4.2.15 respectively.

The corresponding questions for $\mathscr{M}$ with two sets of symmetric signatures were answered in the affirmative by Theorem 4.1 in [43]. What remains open is the case in which the signatures might not be symmetric (which tends to be easier since the input is exponentially larger than in the symmetric case).

Open Problem 4.3.9. Let $\mathcal{F}$ and $\mathcal{G}$ be finite sets of complex-valued signatures over a domain of size $\kappa \geq 2$.

1. Is there an algorithm to decide if there exists a $T \in \mathbf{GL}_2(\mathbb{C})$ such that $\mathcal{G}T \subseteq \mathscr{M}$ and $\mathcal{F} \subseteq T\mathscr{M}$?
2. If so, is there an algorithm to find such a T?
3. In either case, if an algorithm exists, then does a polynomial-time algorithm exist?

The heart of the FKT algorithm is to find a Pfaffian orientation. For such an orientation, the number of perfect matchings is the Pfaffian of the corresponding signed adjacency matrix. Furthermore, The Pfaffian of such a matrix can be computed in polynomial time using efficient algorithms for the determinant. What Kasteleyn proved in [89, 88] is that planar graphs have a Pfaffian orientation and that one can be found in polynomial time.

Is there a larger class of graphs with a Pfaffian orientation that count be found in polynomial time? The answer is yes! Little [104] proved that $K_{3,3}$-free graphs have a Pfaffian orientation, and Vazirani [134] showed how to find such an orientation in polynomial time. More recently, it was shown how to count perfect matchings over K_5-free graphs [118], and then over H-free graphs for any graph H with a planar embedding that has at most one crossing [56], both in polynomial time. However, these new results are not based on finding a Pfaffian orientation.

4.4 Vanishing

This section is about a tractable case called vanishing signatures. It was published in [29, 30].

Vanishing signatures were first introduced in [72] in the parity setting to denote signatures for which the Holant value is always 0 modulo 2.

Definition 4.4.1. A set of signatures $\mathcal{F}$ is called *vanishing* if the value $\operatorname{Holant}_\Omega(\mathcal{F})$ is 0 for every signature grid Ω. A signature f is called *vanishing* if the singleton set $\{f\}$ is vanishing.

A simple lemma (Lemma 6.2 in [72]) from the parity setting also holds over any field $\mathbb{F}$ with the same proof. It also works for signatures that may not be symmetric. Let $f + g$ denote the entry-wise addition of two signatures f and g with the same arity, i.e. $(f + g)_\ell = f_\ell + g_\ell$ for any index ℓ.

Lemma 4.4.2. *Let $\mathcal{F}$ be a vanishing signature set. If f is the signature of an $\mathcal{F}$-gate, then $\mathcal{F} \cup \{f\}$ is also vanishing. If $f, g \in \mathcal{F}$ of the same arity, then $\mathcal{F} \cup \{f + g\}$ is vanishing as well.*

Obviously, any signature that is identically zero is vanishing. This is trivial. However, the concept of vanishing signatures is not trivial. The unary signature $[1, i]$ connected to another $[1, i]$ has a Holant value of 0. Consider a signature set $\mathcal{F}$ in which every signature of arity n is degenerate. That is, every signature of arity n is a tensor product of unary signatures. Moreover, for each signature, suppose that more than half of the unary signatures in the tensor product are

$[1, i]$. For any signature grid Ω with signatures from $\mathcal{F}$, it can be decomposed into many pairs of unary signatures. The total Holant value is the product of the Holant on each pair. Since more than half of the unaries in each signature are $[1, i]$, more than half of the unaries in Ω are $[1, i]$. Then two $[1, i]$'s must be paired up. Hence $\text{Holant}(\Omega) = 0$. Thus, all such signatures form a vanishing set. We also observe that this argument holds when $[1, i]$ is replaced by $[1, -i]$.

These observations were the genesis for the complex-weighted Holant dichotomy in [30]. A dichotomy for real-weighted Holant had recently been proved by Huang and Lu [80]. The tractable cases for their dichotomy were same as the complex-weighted #CSP dichotomy [53] after one accounts for additional tractability from holographic transformations (cf. Corollary 4.1.6 and Corollary 4.2.7). However, we knew that a new vanishing tractable case would appear when considering complex weights.

Thus, we began the work in [30] by characterizing all sets of symmetric vanishing signatures. The signatures in the example described above are generally not symmetric, so we use symmetrization operation given in Definition 4.3.4. Note that we include redundant permutations of v in the definition. Equivalent v_i's also induce redundant permutations. These redundant permutations simply introduce a nonzero constant factor, which does not change the complexity. However, the allowance of redundant permutations simplifies our calculations. An illustrative example of Definition 4.3.4 is

$$\begin{aligned}\text{Sym}_3^2([1, i]; [a, b]) &= 2[a, b] \otimes [1, i] \otimes [1, i] + 2[1, i] \otimes [a, b] \otimes [1, i] + 2[1, i] \otimes [1, i] \otimes [a, b] \\ &= 2[3a, 2ia + b, -a + 2ib, -3b]. \qquad (4.4.7)\end{aligned}$$

Definition 4.4.3. A nonzero symmetric signature f of arity n has *positive vanishing degree* $k \geq 1$, which is denoted by $\text{vd}^+(f) = k$, if $k \leq n$ is the largest positive integer such that there exists $n - k$ unary signatures $v_1, \ldots, v_{n-k}$ satisfying

$$f = \text{Sym}_n^k([1, i]; v_1, \ldots, v_{n-k}).$$

If f cannot be expressed as such a symmetrization form, we define $\text{vd}^+(f) = 0$. If f is the all zero signature, define $\mathrm{vd}^+(f) = n + 1$.

We define *negative vanishing degree* vd^- similarly, using $-i$ instead of i.

Notice that it is possible for a signature f to have both $\mathrm{vd}^+(f)$ and $\mathrm{vd}^-(f)$ nonzero. For example, $f = [1, 0, 1]$ has $\mathrm{vd}^+(f) = \mathrm{vd}^-(f) = 1$.

By the discussion above and Lemma 4.4.2, we know that for a signature f of arity n, if $\mathrm{vd}^\sigma(f) > \frac{n}{2}$ for some $\sigma \in \{+, -\}$, then f is a vanishing signature. This argument is easily generalized to a set of signatures. In particular, the signature in (4.4.7) is vanishing for any $a, b \in \mathbb{C}$.

Definition 4.4.4. For $\sigma \in \{+, -\}$, we define $\mathscr{V}^\sigma = \{f \mid 2\,\mathrm{vd}^\sigma(f) > \mathrm{arity}(f)\}$.

Lemma 4.4.5. *Let $\mathcal{F}$ be a set of symmetric signatures. If $\mathcal{F} \subseteq \mathscr{V}^+$ or $\mathcal{F} \subseteq \mathscr{V}^-$, then $\mathcal{F}$ is vanishing.*

In Theorem 4.4.12, we show that these two sets capture all symmetric vanishing signature sets.

4.4.1 Characterizing Vanishing Signatures using Recurrence Relations

Now we give an equivalent characterization of vanishing signatures.

Definition 4.4.6. A symmetric signature $f = [f_0, f_1, \ldots, f_n]$ of arity n is in $\mathscr{R}_t^+$ for a nonnegative integer $t \geq 0$ if $t > n$ or for any $0 \leq k \leq n - t$, $f_k, \ldots, f_{k+t}$ satisfy the recurrence relation

$$\binom{t}{t} i^t f_{k+t} + \binom{t}{t-1} i^{t-1} f_{k+t-1} + \cdots + \binom{t}{0} i^0 f_k = 0. \tag{4.4.8}$$

We define $\mathscr{R}_t^-$ similarly but with $-i$ in place of i in (4.4.8).

It is easy to see that $\mathscr{R}_0^+ = \mathscr{R}_0^-$ is the set of all zero signatures. Also, for $\sigma \in \{+, -\}$, we have $\mathscr{R}_t^\sigma \subseteq \mathscr{R}_{t'}^\sigma$ when $t \leq t'$. By definition, if $\mathrm{arity}(f) = n$ then $f \in \mathscr{R}_{n+1}^\sigma$.

Let $f = [f_0, f_1, \ldots, f_n] \in \mathscr{R}_t^+$ with $0 < t \leq n$. Then the characteristic polynomial of its recurrence relation is $(1 + xi)^t$. Thus there exists a polynomial $p(x)$ of degree at most $t - 1$ such that $f_k = i^k p(k)$, for $0 \leq k \leq n$. This statement extends to $\mathscr{R}_{n+1}^+$ since a polynomial of degree n can interpolate any set of $n + 1$ values. Furthermore, such an expression is unique. If there are two polynomials $p(x)$ and $q(x)$, both of degree at most n, such that $f_k = i^k p(k) = i^k q(k)$ for $0 \leq k \leq n$, then $p(x)$ and $q(x)$ must be the same polynomial. Now suppose $f_k = i^k p(k)$ $(0 \le k \le n)$ for some polynomial p of degree at most $t-1$, where $0 < t \le n$. Then f satisfies the recurrence (4.4.8) of order t. Hence $f \in \mathscr{R}_t^+$.

Thus, $f \in \mathscr{R}_{t+1}^+$ iff there exists a polynomials $p(x)$ of degree at most t such that $f_k = i^k p(k)$ $(0 \le k \le n)$, for all $0 \le t \le n$. For $\mathscr{R}_{t+1}^-$, just replace i by $-i$.

Definition 4.4.7. For a nonzero symmetric signature f of arity n, it is of *positive* (resp. *negative*) *recurrence degree* $t \le n$, denoted by $\mathrm{rd}^+(f) = t$ (resp. $\mathrm{rd}^-(f) = t$), if and only if $f \in \mathscr{R}_{t+1}^+ - \mathscr{R}_t^+$ (resp. $f \in \mathscr{R}_{t+1}^- - \mathscr{R}_t^-$). If f is the all zero signature, we define $\mathrm{rd}^+(f) = \mathrm{rd}^-(f) = -1$.

Note that although we call it the recurrence degree, it refers to a special kind of recurrence relation. For any nonzero symmetric signature f, by the uniqueness of the representing polynomial $p(x)$, it follows that $\mathrm{rd}^\sigma(f) = t$ iff $\deg(p) = t$, where $0 \le t \le n$. We remark that $\mathrm{rd}^\sigma(f)$ is the maximum integer t such that f does *not* belong to $\mathscr{R}_t^\sigma$. Also, for an arity n signature f, $\mathrm{rd}^\sigma(f) = n$ if and only if f does not satisfy any such recurrence relation (4.4.8) of order $t \le n$ for $\sigma \in \{+, -\}$.

Lemma 4.4.8. *Let $f = [f_0, \dots, f_n]$ be a symmetric signature of arity n, not identically zero. Then for any nonnegative integer $0 \le t < n$ and $\sigma \in \{+, -\}$, the following are equivalent:*

(i) There exist t unary signatures $v_1, \dots, v_t$, such that

$$f = \mathrm{Sym}_n^{n-t}([1, \sigma i]; v_1, \dots, v_t). \tag{4.4.9}$$

(ii) $f \in \mathscr{R}_{t+1}^\sigma$.

Proof. We consider $\sigma = +$ since the other case is similar, so let $v = [1, i]$.

We start with $(i) \implies (ii)$ and proceed via induction on both t and n. For the first base case of $t = 0$, $\mathrm{Sym}_n^n(v) = [1, i]^{\otimes n} = [1, i, -1, -i, \dots, i^n]$, so $f_{k+1} = i f_k$ for all $0 \le k \le n-1$ and $f \in \mathscr{R}_1^+$.

The other base case is that $t = n-1$. Let $\mathrm{Sym}_n^1(v; v_1, \dots, v_t) = [f_0, \dots, f_n]$ where $v_i = [a_i, b_i]$ for $1 \le i \le t$, and $S = i^n f_n + \dots + \binom{n}{1} i f_1 + \binom{n}{0} i^0 f_0$. We need to show that $S = 0$. First notice that any entry in f is a linear combination of terms of the form $a_{i_1} a_{i_2} \cdots a_{i_{n-1-k}} b_{j_1} \cdots b_{j_k}$, where $0 \le k \le n-1$, and $\{i_1, \dots, i_{n-1-k}, j_1, \dots, j_k\} = \{1, 2, \dots, n-1\}$. Thus S is a linear combination of such terms as well. Now we compute the coefficient of each of these terms in S.

Each term $a_{i_1}a_{i_2}\cdots a_{i_{n-1-k}}b_{j_1}\cdots b_{j_k}$ appears twice in S, once in f_k and the other time in f_{k+1}. In f_k, the coefficient is $k!(n-k)!$, and in f_{k+1}, it is $i(k+1)!(n-k-1)!$. Thus, its coefficient in S is

$$\binom{n}{k+1}i^{k+1}i(k+1)!(n-k-1)!+\binom{n}{k}i^k k!(n-k)!=0.$$

The above computation works for any such term due to the symmetry of f, so all coefficients in S are 0, which means that $S=0$.

Now assume for any $t'<t$ or for the same t and any $n'<n$, the statement holds. For (n,t), where $n>t+1$, assume that $f=[f_0,\ldots,f_n]=\operatorname{Sym}_n^{n-t}(v;v_1,\ldots,v_t)$, $g=\operatorname{Sym}_{n-1}^{n-t-1}(v;v_1,\ldots,v_t)=[g_0,\ldots,g_{n-1}]$, and for any $1\le j\le t$, $h^{(j)}=\operatorname{Sym}_{n-1}^{n-t}(v;v_1,\ldots,v_{j-1},v_{j+1},\ldots,v_t)=[h_0^{(j)},\ldots,h_{n-1}^{(j)}]$. By the induction hypothesis, g satisfies the recurrence relation of order $t+1$, namely $g\in\mathscr{R}_{t+1}^+$. Also for any j, $h^{(j)}$ satisfies the recurrence relation of order t, namely $h^{(j)}\in\mathscr{R}_t^+\subseteq\mathscr{R}_{t+1}^+$.

We have the recurrence relation

$$\begin{aligned}\operatorname{Sym}_n^{n-t}(v;v_1,\ldots,v_t)=&(n-t)v\otimes\operatorname{Sym}_{n-1}^{n-t-1}(v;v_1,\ldots,v_t)\\&+\sum_{j=1}^{t}v_j\otimes\operatorname{Sym}_{n-1}^{n-t}(v;v_1,\ldots,v_{j-1},v_{j+1},\ldots,v_t).\end{aligned}\tag{4.4.10}$$

By (4.4.10), the entry of weight k in f for any $k>0$ is

$$f_k=(n-t)ig_{k-1}+\sum_{j=1}^{t}b_jh_{k-1}^{(j)}.$$

We know that $\{g_i\}$ and $\{h_i^{(j)}\}$ satisfy the recurrence relation (4.4.8) of order $t+1$. Thus, their linear combination $\{f_i\}$ also satisfies the recurrence relation (4.4.8) starting from $i=k>0$.

We also observe that by (4.4.10), the entry of weight k in f for any $k<n$ is

$$f_k=(n-t)g_k+\sum_{j=1}^{t}a_jh_k^{(j)}.$$

Since $t<n-1$, by the same argument, the recurrence relation (4.4.8) holds for f when $k=0$ as well.

Now we show $(ii) \implies (i)$. Notice that we only need to find unary signatures $\{v_i\}$ for $1 \le i \le t$ such that $\operatorname{Sym}_n^{n-t}(v; v_1, \ldots, v_t)$ matches the first $t+1$ entries of f. The theorem follows from this since we have shown that $\operatorname{Sym}_n^{n-t}(v; v_1, \ldots, v_t)$ satisfies the recurrence relation of order $t+1$ and any such signature is determined by the first $t+1$ entries.

We show that there exist $v_i = [a_i, b_i]$ $(1 \le i \le t)$ satisfying the above requirement. Since f is not identically zero, by (4.4.8), some nonzero term occurs among $\{f_0, \ldots, f_t\}$. Let $f_s \neq 0$, for $0 \le s \le t$, be the first nonzero term. By a nonzero constant multiplier, we may normalize $f_s = s!(n-s)!$, and set $v_j = [0, 1]$, for $1 \le j \le s$ (which is vacuous if $s = 0$), and set $v_{s+j} = [1, b_{s+j}]$, for $1 \le j \le t-s$ (which is vacuous if $s = t$). Let F be the function defined in (4.4.9). Then $F_k = f_k = 0$ for $0 \le k < s$ (which is vacuous if $s = 0$). By expanding the symmetrization function, for $s \le k \le t$, we get

$$F_k = k!(n-k)! \sum_{j=0}^{k-s} \binom{n-t}{k-s-j} \Delta_j i^{k-s-j},$$

where Δ_j is the elementary symmetric polynomial in $\{b_{s+1}, \ldots, b_t\}$ of degree j for $0 \le j \le t-s$. By definition, $\Delta_0 = 1$ and $F_s = f_s$. Setting $F_k = f_k$ for $s+1 \le k \le t$, this is a linear equation system on Δ_j $(1 \le j \le t-s)$, with a triangular matrix and nonzero diagonals. From this, we know that all Δ_j's are uniquely determined by $\{f_{s+1}, \ldots, f_t\}$. Moreover, $\{b_{s+1}, \ldots, b_t\}$ are the roots of the equation $\sum_{j=0}^{t-s} (-1)^j \Delta_j x^{t-j} = 0$. Thus $\{b_{s+1}, \ldots, b_t\}$ are also uniquely determined by $\{f_{s+1}, \ldots, f_t\}$ up to a permutation. □

Corollary 4.4.9. *If f is a symmetric signature and $\sigma \in \{+, -\}$, then* $\operatorname{vd}^\sigma(f) + \operatorname{rd}^\sigma(f) = \operatorname{arity}(f)$.

Thus we have an equivalent form of $\mathscr{V}^\sigma$ for $\sigma \in \{+, -\}$. Namely,

$$\mathscr{V}^\sigma = \{f \mid 2 \operatorname{rd}^\sigma(f) < \operatorname{arity}(f)\}.$$

4.4.2 Characterizing Vanishing Signature Sets

Now we show that $\mathscr{V}^+$ and $\mathscr{V}^-$ capture all symmetric vanishing signature sets. To begin, we show that a vanishing signature set cannot contain both types of nontrivial vanishing signatures.

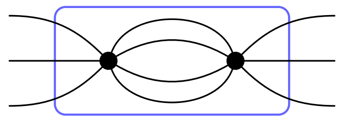


Figure 4.2: Example of a gadget used to create a degenerate vanishing signature from some general vanishing signature. This example is for a signature of arity 7 and recurrence degree 2, which is assigned to both vertices.

Lemma 4.4.10. *Let $f_+ \in \mathscr{V}^+$ and $f_- \in \mathscr{V}^-$. If neither f_+ nor f_- is the all zero signature, then the signature set $\{f_+, f_-\}$ is not vanishing.*

Proof. Let $\operatorname{arity}(f_+) = n$ and $\operatorname{rd}^+(f_+) = t$, so $2t < n$. Consider the gadget with two vertices and $2t$ edges between two copies of f_+. (See Figure 4.2 for an example of this gadget.) View f_+ in the symmetrized form. Since $\operatorname{vd}^+(f_+) = n - t$, in each term, there are $n - t$ many $[1, i]$'s and t many unary signatures not equal to (a multiple of) $[1, i]$. This is a superposition of many degenerate signatures. Then the only non-vanishing contributions come from the cases where the $n - 2t$ dangling edges on both sides are all assigned $[1, i]$, while inside, the t copies of $[1, i]$ pair up with t unary signatures not equal to $[1, i]$ from the other side perfectly. Notice that for any such contribution, the Holant value of the inside part is always the same constant and this constant is not 0 because $[1, i]$ paired up with any unary signature other than (a multiple of) $[1, i]$ is not 0. Then the superposition of all of the permutations is a degenerate signature $[1, i]^{\otimes 2(n-2t)}$ up to a nonzero constant factor.

Similarly, we can do this for f_- of arity n' and $\operatorname{rd}^-(f_-) = t'$, where $2t' < n'$, and get a degenerate signature $[1, -i]^{\otimes 2(n'-2t')}$, up to a nonzero constant factor. Then form a bipartite signature grid with $(n' - 2t')$ vertices on one side, each assigned $[1, i]^{\otimes 2(n-2t)}$, and $(n - 2t)$ vertices on the other side, each assigned $[1, -i]^{\otimes 2(n'-2t')}$. Connect edges between the two sides arbitrarily as long as it is a 1-1 correspondence. The resulting Holant is a power of 2, which is not vanishing. □

Lemma 4.4.11. *Every symmetric vanishing signature is in $\mathscr{V}^+ \cup \mathscr{V}^-$.*

Proof. Let f be a symmetric vanishing signature. We prove this by induction on n, the arity of f. For $n = 1$, by connecting $f = [f_0, f_1]$ to itself, we have $f_0^2 + f_1^2 = 0$. Then up to a constant factor, we have either $f = [1, i]$ or $f = [1, -i]$. The lemma holds.

For $n = 2$, first we do a self loop. The Holant is $f_0 + f_2$. Also, we can connect two copies of f, in which case the Holant is $f_0^2 + 2f_1^2 + f_2^2$. Since f is vanishing, both $f_0 + f_2 = 0$ and $f_0^2 + 2f_1^2 + f_2^2 = 0$. Solving them, we get $f = [1, i, -1] = [1, i]^{\otimes 2}$ or $[1, -i, -1] = [1, -i]^{\otimes 2}$ up to a constant factor.

Now assume $n > 2$ and the lemma holds for any signature of arity $k < n$. Let $f = [f_0, f_1, \ldots, f_n]$ be a vanishing signature. A self loop on f gives $f' = [f'_0, f'_1, \ldots, f'_{n-2}]$, where $f'_j = f_j + f_{j+2}$ for $0 \le j \le n-2$. Since f is vanishing, f' is vanishing as well. By the induction hypothesis, $f' \in \mathscr{V}^+ \cup \mathscr{V}^-$.

If f' is an all zero signature, then we have $f_j + f_{j+2} = 0$ for $0 \le j \le n-2$. This means that the f_j's satisfy a recurrence relation with characteristic polynomial $x^2 + 1$, so we have $f_j = ai^j + b(-i)^j$ for some a and b. Then we perform a holographic transformation with $Z = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$,

$$\begin{aligned} \operatorname{Holant}(=_2 \mid f) &\equiv_T \operatorname{Holant}([1,0,1]Z^{\otimes 2} \mid (Z^{-1})^{\otimes n} f) \\ &\equiv_T \operatorname{Holant}([0,1,0] \mid \hat{f}), \end{aligned}$$

where $\hat{f} = [a, 0, \ldots, 0, b]$. The problem $\operatorname{Holant}([0,1,0] \mid \hat{f})$ is a weighted version of testing if a graph is bipartite. Now consider a graph with only two vertices, both assigned f, and n edges between them. The Holant of this graph is $2ab$. However, we know that it must be vanishing, so $ab = 0$. If $a = 0$, then $f \in \mathscr{V}^-$. Otherwise, $b = 0$ and $f \in \mathscr{V}^+$.

Now suppose that f' is in $\mathscr{V}^+ \cup \mathscr{V}^-$ but is not an all zero signature. We consider $f' \in \mathscr{V}^+$ since the other case is similar. Then $\operatorname{rd}^+(f') = t$, so $2t < n - 2$. Consider the gadget which has only two vertices, both assigned f', and has $2t$ edges between them. (See Figure 4.2 for an example of this gadget.) It forms a signature of degree $d = 2(n - 2 - 2t)$. This gadget is valid because $n - 2 > 2t$. By the combinatorial view as in the proof of Lemma 4.4.10, this signature is $[1, i]^{\otimes d}$.

Moreover, $\operatorname{rd}^+(f') = t$ implies that the entries of f' satisfy a recurrence of order $t+1$. Replacing f'_j by $f_j + f_{j+2}$, we get a recurrence relation for the entries of f with characteristic polynomial $(x^2+1)(x-i)^{t+1} = (x+i)(x-i)^{t+2}$. Thus, $f_j = i^j p(j) + c(-i)^j$ for some polynomial $p(x)$ of degree at most $t + 1$ and some constant c. It suffices to show that $c = 0$ since $2(t + 1) < n$ as $2t < n - 2$.

Consider the signature $h = [h_0, \ldots, h_{n-1}]$ created by connecting f with a single unary signature

$[1, i]$. For any $(n-1)$-regular graph $G = (V, E)$ with h assigned to every vertex, we can define a duplicate graph of $(d+1)|V|$ vertices as follows. First for each $v \in V$, define vertices v', $v_1, \ldots, v_d$. For each i, $1 \le i \le d$, we make a copy of G on $\{v_i \mid v \in V\}$, i.e., for each edge $(u, v) \in E$, include the edge (u_i, v_i) in the new graph. Next for each $v \in V$, we introduce edges between v' and v_i for all $1 \le i \le d$. For each $v \in V$, assign the degenerate signature $[1, i]^{\otimes d}$ that we just constructed to the vertices v'; assign f to all the vertices $v_1, \ldots, v_d$. Assume the Holant of the original graph G with h assigned to every vertex is H. Then for the new graph with the given signature assignments, the Holant is H^d. By our assumption, f is vanishing, so $H^d = 0$. Thus, $H = 0$. This holds for any graph G, so h is vanishing.

Notice that $h_k = f_k + i f_{k+1}$ for any $0 \le k \le n-1$. If h is identically zero, then $f_k + i f_{k+1} = 0$ for any $0 \le k \le n-1$, which means $f = [1, i]^{\otimes n}$ up to a constant factor and we are done. Otherwise, suppose that h is not identically zero. By the inductive hypothesis, $h \in \mathscr{V}^+ \cup \mathscr{V}^-$. We claim h cannot be from $\mathscr{V}^-$. This is because, although we do not directly construct h from f, we can always realize it by the method depicted in the previous paragraph. Therefore the set $\{f', h\}$ is vanishing. As both f' and h are nonzero, and $f' \in \mathscr{V}^+$, we have $h \notin \mathscr{V}^-$, by Lemma 4.4.10.

Hence h is in $\mathscr{V}^+$. Then there exists a polynomial $q(x)$ of degree at most $t' = \lfloor \frac{n-1}{2} \rfloor$ such that $h_k = i^k q(k)$, for any $0 \le k \le n-1$. Since $2t < n-2$, we have $t \le t'$. On the other hand, $h_k = f_k + i f_{k+1}$ for any $0 \le k \le n-1$, so we have

$$\begin{aligned}
h_k &= f_k + i f_{k+1} \\
&= i^k p(k) + c(-i)^k + i\left(i^{k+1} p(k+1) + c(-i)^{k+1}\right) \\
&= i^k \left(p(k) - p(k+1)\right) + 2c(-i)^k \\
&= i^k r(k) + 2c(-i)^k \\
&= i^k q(k),
\end{aligned}$$

where $r(x) = p(x) - p(x+1)$ is another polynomial of degree at most t. Then we have

$$q(k) - r(k) = 2c(-1)^k,$$

which holds for all $0 \le k \le n-1$. Notice that the left hand side is a polynomial of degree at most t', call it $s(x)$. However, for all even $k \in \{0, \dots, n-1\}$, $s(k) = 2c$. There are exactly $\lceil \frac{n}{2} \rceil > \lfloor \frac{n-1}{2} \rfloor = t'$ many even k within the range $\{0, \dots, n-1\}$. Thus $s(x) = 2c$ for any x. Now we pick $k = 1$, so $s(1) = -2c = 2c$, which implies $c = 0$. This completes the proof. □

Combining Lemma 4.4.5, Lemma 4.4.10, and Lemma 4.4.11, we obtain the following theorem that characterizes all symmetric vanishing signature sets.

Theorem 4.4.12. *Let $\mathcal{F}$ be a set of symmetric signatures. Then $\mathcal{F}$ is vanishing if and only if $\mathcal{F} \subseteq \mathscr{V}^+$ or $\mathcal{F} \subseteq \mathscr{V}^-$.*

To finish this subsection, we prove some useful properties regarding vanishing and recurrence degrees in the construction of signatures. For two symmetric signatures f and g such that $\operatorname{arity}(f) \ge \operatorname{arity}(g)$, let $\langle f, g \rangle = \langle g, f \rangle$ denote the signature that results after connecting all edges of g to f. (If $\operatorname{arity}(f) = \operatorname{arity}(g)$, then $\langle f, g \rangle$ is a constant, which can be viewed as a signature of arity 0.)

Lemma 4.4.13. *For $\sigma \in \{+, -\}$, suppose symmetric signatures f and g satisfy $\operatorname{vd}^\sigma(g) = 0$ and $\operatorname{arity}(f) - \operatorname{arity}(g) \ge \operatorname{rd}^\sigma(f)$. Then $\operatorname{rd}^\sigma(\langle f, g \rangle) = \operatorname{rd}^\sigma(f)$.*

Proof. We consider $\sigma = +$ since the case $\sigma = -$ is similar. Let $\operatorname{arity}(f) = n$, $\operatorname{arity}(g) = m$, and $\operatorname{rd}^+(f) = t$. Denote the signature $\langle f, g \rangle$ by f'.

If $t = -1$, then f is identically zero and so is f'. Hence $\operatorname{rd}^+(f') = -1$.

Suppose $t \ge 0$. Then we have $f_k = i^k p(k)$ where $p(x)$ is a polynomial of degree exactly t. Also $\operatorname{arity}(f') = n - m \ge t$. We have

$$\begin{aligned} f'_k &= \sum_{j=0}^{m} \binom{m}{j} f_{k+j} g_j \\ &= i^k \sum_{j=0}^{m} \binom{m}{j} p(k+j) i^j g_j \\ &= i^k q(k), \end{aligned}$$

where $q(k) = \sum_{j=0}^{m} \binom{m}{j} p(k+j) i^j g_j$ is a polynomial in k. Notice that $\operatorname{vd}^+(g) = 0$. Then $\operatorname{rd}^+(g) = m$ and $g \notin \mathscr{R}^+_m$. Thus $\sum_{j=0}^{m} \binom{m}{j} i^j g_j \neq 0$. Then the leading coefficient of degree t in the polynomial

$q(k)$ is nonzero. However, $\operatorname{arity}(f') \geq t$. Thus $\operatorname{rd}^+(f') = t$ as well. $\square$

Lemma 4.4.14. *For $\sigma \in \{+, -\}$, let f be a nonzero symmetric signature and suppose that f' is obtained from f by a self loop. If $\operatorname{vd}^\sigma(f) > 0$, then $\operatorname{vd}^\sigma(f) - \operatorname{vd}^\sigma(f') = \operatorname{rd}^\sigma(f) - \operatorname{rd}^\sigma(f') = 1$.*

Proof. We may assume $\sigma = +$, $\operatorname{arity}(f) = n$, and $\operatorname{rd}^+(f) = t$. Since f is not the all zero signature, $t \geq 0$. Also since $\operatorname{vd}^+(f) > 0$, $t = n - \operatorname{vd}^+(f) < n$. By assumption, we have $f_k = i^k p(k)$, where $p(x)$ is a polynomial of degree exactly t. Then we have

$$\begin{aligned} f'_k &= f_k + f_{k+2} \\ &= i^k(p(k) - p(k+2)) \\ &= i^k q(k), \end{aligned}$$

where $q(k) = p(k) - p(k+2)$ is a polynomial in k. If $t = 0$, then $p(x)$ is a constant polynomial and $q(x)$ is identically zero. Then $\operatorname{rd}^+(f') = -1$ by definition and $\operatorname{rd}^+(f) - \operatorname{rd}^+(f') = 1$ holds. Suppose $t > 0$, then in $q(k)$, the term of degree t has a zero coefficient, but the term of degree $t-1$ is nonzero. So $q(x)$ has degree exactly $t - 1 \leq n - 2 = \operatorname{arity}(f')$. Thus $\operatorname{rd}^+(f') = t - 1$. Notice that $\operatorname{arity}(f) - \operatorname{arity}(f') = 2$, then $\operatorname{vd}^+(f) - \operatorname{vd}^+(f') = 1$ as well. $\square$

Moreover, the set of vanishing signatures is closed under orthogonal transformations. This is because under any orthogonal transformation, the unary signatures $[1, i]$ and $[1, -i]$ are either invariant or transformed into each other. Then considering the symmetrized form of any signature, we have the following lemma.

Lemma 4.4.15. *For a symmetric signature f of arity n, $\sigma \in \{+, -\}$, and an orthogonal matrix $T \in \mathbb{C}^{2\times 2}$, either $\operatorname{vd}^\sigma(f) = \operatorname{vd}^\sigma(T^{\otimes n} f)$ or $\operatorname{vd}^\sigma(f) = \operatorname{vd}^{-\sigma}(T^{\otimes n} f)$.*

4.4.3 Characterizing Vanishing Signatures via a Holographic Transformation

There is another explanation for the vanishing signatures. Given an $f \in \mathscr{V}^+$ with $\operatorname{arity}(f) = n$ and $\operatorname{rd}^+(f) = d$, we perform a holographic transformation with $Z = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$,

$$\begin{aligned} \operatorname{Holant}(=_2 \mid f) &\equiv_T \operatorname{Holant}([1,0,1]Z^{\otimes 2} \mid (Z^{-1})^{\otimes n} f) \\ &\equiv_T \operatorname{Holant}([0,1,0] \mid \hat{f}), \end{aligned}$$

where $\hat{f}$ is of the form $[\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0]$, and $\hat{f}_d \neq 0$. To see this, note that $Z^{-1} = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & -i \\ 1 & i \end{smallmatrix}\right]$ and $Z^{-1} \left[\begin{smallmatrix} 1 \\ i \end{smallmatrix}\right] = \sqrt{2} \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$. We know that f has a symmetrized form, such as $\operatorname{Sym}_n^{n-d}(\left[\begin{smallmatrix} 1 \\ i \end{smallmatrix}\right]; v_1, \ldots, v_d)$. Then up to a factor of $2^{n/2}$, we have $\hat{f} = (Z^{-1})^{\otimes n} f = \operatorname{Sym}_n^{n-d}(\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]; u_1, \ldots, u_d)$, where $u_i = Z^{-1} v_i$ for $1 \leq i \leq d$ and u_i and v_i are column vectors in $\mathbb{C}^2$. From this expression for $\hat{f}$, it is clear that all entries of Hamming weight greater than d in $\hat{f}$ are 0. Moreover, if $\hat{f}_d = 0$, then one of the u_i has to be a multiple of $[1,0]$. This contradicts the degree assumption of f, namely $\operatorname{vd}^+(f) = n - \operatorname{rd}^+(f) = n - d$ and no higher.

From this discussion, we have the following lemma.

Lemma 4.4.16. *Suppose f is a symmetric signature of arity n. Let $\hat{f} = (Z^{-1})^{\otimes n} f$. If $\operatorname{rd}^+(f) = d$, then $\hat{f} = [\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0]$ and $\hat{f}_d \neq 0$. Also $f \in \mathscr{R}_d^+$ iff all nonzero entries of $\hat{f}$ are among the first d entries in its symmetric signature notation.*

Similarly, if $\operatorname{rd}^-(f) = d$, then $\hat{f} = [0, \ldots, 0, \hat{f}_{n-d}, \ldots, \hat{f}_n]$ and $\hat{f}_{n-d} \neq 0$. Also $f \in \mathscr{R}_d^-$ iff all nonzero entries of $\hat{f}$ are among the last d entries in its symmetric signature notation.

This presentation of vanishing signatures follows the order in which we thought about them. Initially, we viewed them as summations of degenerate signatures with more than have $[1,i]$'s or $[1,-i]$'s. For this reason, we defined the notions of positive and negative vanishing degrees to indicate how many $[1,i]$'s or $[1,-i]$'s respectively can exist in such an expression. Next we realized that the entries of a symmetric vanishing signature satisfy a recurrence relation involving $\pm i$. For this reason, we defined the notions of positive and negative recurrence degrees to indicate the order of this recurrence relation. Lastly, we gained the perspective of vanishing signatures in the Z basis as described above.

In any bipartite graph for Holant($[0,1,0] \mid \hat{f}$), the binary DISEQUALITY $(\neq_2) = [0,1,0]$ on the left imposes the condition that half of the edges must take the value 0 and the other half must take the value 1. On the right side, by $f \in \mathscr{V}^+$, we have $d < n/2$, thus $\hat{f}$ requires that less than half of the edges are assigned the value 1. Therefore the Holant is always 0. A similar conclusion was reached in [50] for certain $(2,3)$-regular bipartite Holant problems with Boolean signatures. However, the importance was not realized at that time—not until the spring of 2012. At that time, we were working on [30], and Jin-Yi Cai was teaching a special topics course called Complexity of Counting Problems in which he presented [50]. When he stated that some cases were tractable because their Holant is always 0, Heng Guo immediately did the calculations to verify the vanishing and recurrence degrees of the signatures involved when expressed in the standard basis.

I now believe that this Z-basis view is the right way to think about vanishing signatures. Not only is the argument that they are vanishing trivial, but it is also a useful generalization to signatures that are not symmetric. In contrast, definition of vanishing signatures using the vanishing degrees is not that helpful, and it is not clear how to generalize recurrence degrees to asymmetric signatures. Furthermore, gadget constructions with vanishing signatures in the Z basis are more combinatorial, easier to find, and easier to explain than the same gadget constructions in the standard basis.

Open Problem 4.4.17. Are there any complex-weighted signatures over the Boolean domain that are vanishing for some reason other than the one provided by the Z-basis view?

In our characterization of symmetric vanishing signatures, one can check that all of our gadget constructions are planar. This means that there does not exist a signature that is vanishing over planar graphs but is not vanishing over general graphs. This makes me think of the following question.

Open Problem 4.4.18. Does there exists an "interesting" class $\mathcal{C}$ of graphs and a signature f such that f is vanishing over $\mathcal{C}$ but not vanishing over general graphs?

Although vanishing signatures were first introduced in [72] in the parity setting (where a signature set is vanishing if the corresponding Holant value is always 0 modulo 2), a complete characterization of those vanishing signatures was not given. Thus, the dichotomy in [72] is not currently known to be decidable. The difficulty in the parity setting is that each vanishing signature over the Gaussian integers corresponds to a vanishing signature over $\mathbb{Z}_2$. Since $1^2 \equiv -1 \pmod{2}$, the element 1 is a square root of 1 over $\mathbb{Z}_2$ just like i is a square root of 1 over $\mathbb{C}$. Therefore, a vanishing signature f over the Gaussian integers corresponds to a vanishing signature over $\mathbb{Z}_2$ by mapping any occurrence of i to 1 and then reducing modulo 2. Given our characterization for vanishing signatures over $\mathbb{C}$ and the partial characterization of vanishing signatures over $\mathbb{Z}_2$ in [72], I expect that one can now obtain a complete characterization the vanishing signatures over $\mathbb{Z}_2$, thereby making the dichotomy in [72] decidable.

Chapter 5

Dichotomy for #H-Coloring Problems over Planar 3-Regular Directed Graphs

In this chapter, we introduce an idea called anti-gadgets in complexity reductions. These combinatorial gadgets have the effect of erasing the presence of some other graph fragment, as if we had managed to include a negative copy of a graph gadget. We use this idea to prove a dichotomy for a generalization of counting graph homomorphisms that allows for complex weights. This work was published in [38, 39].

5.1 Background

Reduction, the method of transforming one problem to another, and thereby proving the hardness of a problem for an entire complexity class, is arguably the most successful tool in complexity theory to date. When expressed in terms of graph problems, a typical reduction from problem Π_1 to problem Π_2 is carried out by designing a *gadget*—a graph fragment with some desirable properties. The reduction starts from an instance graph G_1 for Π_1 and introduces one or more copies of the gadget to obtain an instance graph G_2 (or possibly multiple instance graphs) for Π_2.

The graph G_2 may contain a polynomial number of copies of the gadget. *But can it include some negative copies of a gadget?* Of course not; the notion of a *negative* graph fragment seems meaningless. However, we introduce an idea in reduction theory that has *the effect of* introducing *negative copies* of a gadget in a reduction. More precisely, we show that our new construction idea, when expressed in algebraic terms, has the same effect as *erasing* the presence of some graph fragment. It is *as if* we managed to include a negative copy of a certain gadget. We call this an *anti-gadget*. It is analogous to the pairing of a particle and its anti-particle in physics. We demonstrate the elegance and usefulness of anti-gadgets by proving a new complexity dichotomy theorem in counting complexity where anti-gadgets play a decisive role. Furthermore, we show that anti-gadgets provide a simple explanation for some miraculous cancellations that were observed in previous results [37, 36]. We also observe how anti-gadgets can guide the search for such gadget sets more by design than by chance.

The new dichotomy theorem that we prove using anti-gadgets can be stated in terms of spin systems over 3-regular graphs with vertices taking values in $\{0, 1\}$ and an arbitrary complex-valued edge function $f(\cdot, \cdot) = (w, x, y, z)$ that is not necessarily symmetric. Define the *partition function* on $G = (V, E)$ as

$$Z(G) = \sum_{\sigma : V(G) \to \{0,1\}} \prod_{(u,v) \in E(G)} f(\sigma(u), \sigma(v)).$$

Depending on the nature of the edge function f, we show that the problem $Z(\cdot)$ is either computable in polynomial time or #P-hard. The tractable cases were what we expected: (1) product-type signatures, (2) holographic reductions to affine signatures, and (3) holographic reductions to matchgate signatures with the input restricted to planar graphs. The formal statement is given in Theorem 5.4.1. This dichotomy extends several previous ones that considered a *symmetric* edge function [50, 48, 93].

This partition function $Z(G)$ computes the sum of weighted graph homomorphisms from the input graph G to the 2-vertex target graph H with weighted adjacency matrix $\left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]$. It is also known as the #H-coloring problem. Recall from Lemma 2.3.1 that we can express this problem as a counting Constraint Satisfaction Problem (#CSP). Then by Lemma 2.2.1, we can further express this problem as a Holant problem. Compared to #CSP and graph homomorphisms, the

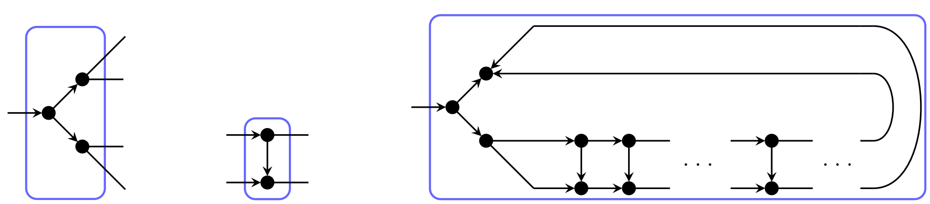
(a) As a bipartite (2,3)-regular graph (b) As a 3-regular directed graph

Figure 5.1: The two representations of an $(f \mid =_3)$-gate.

main difficulty here is bounded degree, which makes hardness proofs more challenging, and for a good reason—there are indeed more tractable cases.

We give a directed explanation of how to express $Z(\cdot)$ as a Holant problem (which is really the combined proofs of Lemma 2.3.1 and Lemma 2.2.1 that are simplified for this specific case). Given any 3-regular directed graph $G = (V, E)$, its edge-vertex incidence graph G' has vertex set $V(G') = V \cup E$ and edge set $E(G') = \{(v, e) \mid v \text{ is incident to } e \text{ in } G\}$. The graph G' is bipartite and $(2, 3)$-regular. To each $v \in V \subset V(G')$, we assign the Equality function $=_3$ of arity 3. To each $e \in E \subset V(G')$, we assign the original edge function f from G. Furthermore, we assign the first (resp. second) input of f to the edge incident to the vertex in V that is incident to the tail (resp. head) of the directed edge in G. Then the Holant value on G' is exactly the partition function $Z(G)$. Essentially $=_3$ forces all incident edges in G' at a vertex $v \in V \subset V(G')$ to take the same value, which reduces to vertex assignments on V, as in $Z(G)$.

As a Holant problem (over $(2, 3)$-regular bipartite graphs), this partition function is expressed as $\operatorname{Holant}(f \mid =_3)$, where $f = (w, x, y, z)$. Our main result is a dichotomy theorem for this problem, where $w, x, y, z \in \mathbb{C}$. To describe an $(f \mid =_3)$-gate though, it is simpler to depict it as a fragment of a 3-regular directed graph. Figure 5.1 gives an example of an $(f \mid =_3)$-gate both as a $(2, 3)$-regular



(a) (b) (c)

Figure 5.2: An arity 4-to-1 projective gadget (a), a recursive gadget (b), and a planar embedding of their interpolation construction (c). My wife thinks (c) looks like a pencil; I agree.

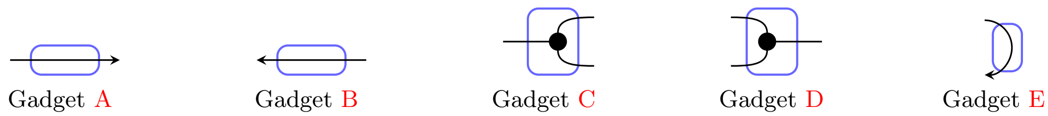


Figure 5.3: Five basic gadget components.

bipartite graph and as an equivalent 3-regular directed graph.

We create $(f \mid =_3)$-gates with two types of dangling edges. We say an edge is *leading* if it is a dangling edge and incidence to a vertex of degree 2, which must be assigned f. We say an edge is *trailing* if it is a dangling edge and incidence to a vertex of degree 3, which must be assigned $=_3$.

Suppose an $(f \mid =_3)$-gate F has n dangling edges and ℓ of them are leading edges. Then we depict F with its leading edges protruding to the left and any trailing edges protruding to the right. We do this both to easily distinguish between these two types of dangling edges, and because we are primarily interested in the corresponding signature matrix with parameter ℓ. In this chapter, we call this signature matrix with parameter ℓ the transition matrix of F. Each gadget we use is given a distinct number i, and we use M_i to denote its transition matrix unless stated otherwise.

The constructions in this chapter are primarily based upon two kinds of $(f \mid =_3)$-gates, which we call *recursive gadgets* and *projective gadgets*. An *arity-d recursive gadget* is an $(f \mid =_3)$-gate with d leading edges and d trailing edges. A *projective gadget from arity n to m* is an $(f \mid =_3)$-gate with m leading edges and n trailing edges with $m < n$. These gadget types are defined in this way to maintain the bipartite structure of the signature grid when we merge trailing edges of one gadget with leading edges of another (see Figure 5.2).

5.2 Gadgets and Anti-Gadgets

In this section, we start with a gentle primer to the association between a gadget and its transition matrix. We show that one can typically express the transition matrix starting from a few of the most basic gadget components and their matrices as atomic building blocks, after applying some well-defined operations. We then introduce anti-gadgets and explain why they are so effective.

We start with five basic gadget components as depicted in Figure 5.3. Their transition matrices are $A = \left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]$, $B = \left[\begin{smallmatrix} w & y \\ x & z \end{smallmatrix}\right]$, $C = \left[\begin{smallmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{smallmatrix}\right]$, $D = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, and $E = \left[\begin{smallmatrix} w \\ x \\ y \\ z \end{smallmatrix}\right]$.

The first operation is matrix product, which corresponds to sequentially connecting two gadgets together. For example, Gadget 1 is a simple composition of Gadget B and Gadget C, and thus its transition matrix is the matrix product $BC = \left[\begin{smallmatrix} w & 0 & 0 & y \\ x & 0 & 0 & z \end{smallmatrix}\right]$ (see Figure 5.5a). The second operation is tensor product, which corresponds to putting two gadgets in parallel (two disconnected parts). The transition matrix of Gadget 2 is $ACB^{\otimes 2}D = \left[\begin{smallmatrix} w^3+x^3 & wy^2+xz^2 \\ w^2y+x^2z & y^3+z^3 \end{smallmatrix}\right]$, where $B^{\otimes 2}$ corresponds to the parallel part of the gadget and is clearly visible in Figure 5.5b. Similarly, Gadget 3 has transition matrix $AC(A \otimes B)D$. Note that the order of the tensor product is to make the top leading edge for the row (resp. the top trailing edge for the column) the most significant bit. The transition matrices of Gadget 4 and Gadget 5 are respectively $\left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]^{\otimes 2} \operatorname{diag}(w, x, y, z)$ and $\left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]^{\otimes 2} \operatorname{diag}(w, y, x, z)$ and can be mechanically derived by our gadgetry calculus as $A^{\otimes 2}(C \otimes I_2)(I_2 \otimes A \otimes I_2)(I_2 \otimes D)$ and $A^{\otimes 2}(C \otimes I_2)(I_2 \otimes B \otimes I_2)(I_2 \otimes D)$. The composition of Gadget 4 is illustrated in Figure 5.5c. Gadget E is used to create a self-loop, as in Gadget 6, which has transition matrix $AC(A \otimes I_2)(C \otimes I_2)(BCE \otimes I_4)(A \otimes B)D$. The composition of Gadget E is illustrated in Figure 5.5c.

Now we introduce a powerful new technique called *anti-gadgets*.

Definition 5.2.1. Let G be a recursive gadget with transition matrix M. Then a recursive gadget G' is called an *anti-gadget* of G if the transition matrix of G' is λM^{-1}, for some $\lambda \in \mathbb{C} - \{0\}$.

A crucial ingredient in our proof of #P-hardness is to produce an arbitrarily large set of pairwise linearly independent signatures. These signatures are used to form a Vandermonde system of full rank. One common way to produce an arbitrarily large set of signatures is to compose copies of a recursive gadget. Let M be the transition matrix of some recursive gadget G. Composing k copies of G produces a gadget with transition matrix M^k. If M has infinite order (up to a scalar), then

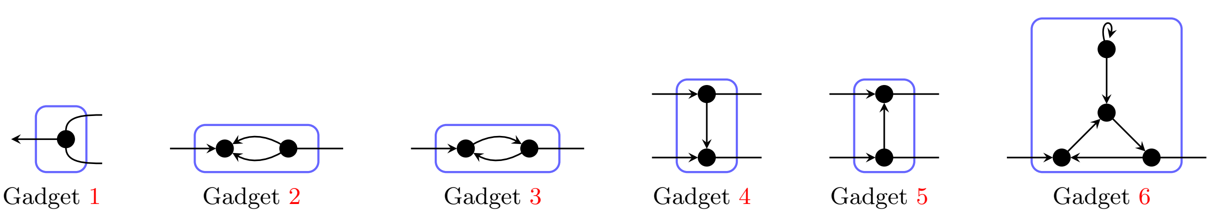
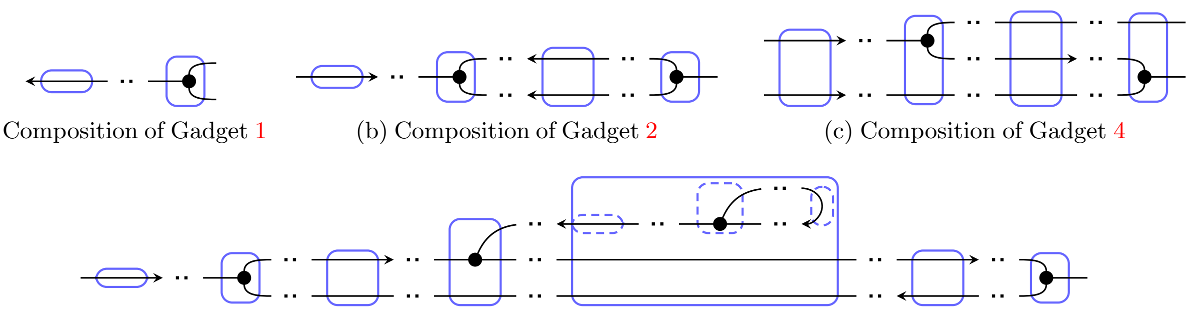


Figure 5.4: Recursive and projective gadgets.



(a) Composition of Gadget 1 (b) Composition of Gadget 2 (c) Composition of Gadget 4

(d) Composition of Gadget 6

Figure 5.5: Gadget compositions using the basic gadget components in Figure 5.3.

we have an arbitrarily large set of pairwise linearly independent signatures. Now suppose that M has finite order (up to a scalar), that is, for some positive integer k, $M^k = \lambda I$, a nonzero multiple of the identity matrix. Then composing only $k-1$ copies of G results in a gadget with a transition matrix that is the *inverse* of G's transition matrix (up to a scalar). This *is* an anti-gadget of G.

If an anti-gadget of G is composed with another gadget containing similar structure to that of G, then cancellations ensue and the composition yields a transition matrix that can be quite easy to analyze. E.g., Gadget 4 and Gadget 5 only differ by the orientation of the vertical edge. When composing an anti-gadget of Gadget 4 with Gadget 5, the contribution of the two leading edges cancel and we get $M_4^{-1} M_5 = \operatorname{diag}(w,x,y,z)^{-1} \left(\left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]^{\otimes 2} \right)^{-1} \left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]^{\otimes 2} \operatorname{diag}(w,y,x,z) = \operatorname{diag}(1, y/x, x/y, 1)$. The resulting transition matrix has infinite order unless x/y is a root of unity. This situation is analyzed formally in Lemma 5.5.1.

Another use of the anti-gadget technique can be applied with Gadget 2 and Gadget 3. Once again, the contribution of the leading edge cancels when composing an anti-gadget of Gadget 3 with Gadget 2. The resulting matrix is a bit more complicated this time. However, when this pair of gadgets is analyzed formally in Lemma 5.5.2, the assumptions are $x = 0 \wedge wyz \neq 0$. In that case, $M_3^{-1} M_2 = \left[\begin{smallmatrix} 1 & y^2/w^2 \\ 0 & 1 \end{smallmatrix}\right]$. This matrix clearly has infinite order (up to a scalar).

5.3 Interpolation Techniques

The method of polynomial interpolation has been pioneered by Valiant [126] and further developed by many others [59, 124, 14, 17, 50]. In this section, we give a new unified technique to interpolate all unary signatures. This is our main technical step to prove #P-hardness. Our method produces an infinite set of pairwise linearly independent vectors at any fixed dimension, and then projects to a lower dimension while retaining pairwise linear independence of a nontrivial fraction.

In previous work, "finisher gadgets" [93, 37, 36] were used to handle the symmetric case, mapping symmetric arity 2 signatures to arity 1 signatures. In our language, a finisher gadget is a projective gadget from a projective gadget from arity 2 to 1.

In this chapter, we introduce *projective gadget sets*. These gadget sets are completely general, in the sense that they can be used to map *any* set of pairwise linearly independent signatures (symmetric *or* asymmetric) to any lower arity, while preserving pairwise linear independence for an inverse polynomial fraction. This permits much more freedom in gadget constructions, and this power is used crucially in the proof of our dichotomy theorem. This advance is not just a simple matter of searching for the right gadgets. One must find the abstract criteria for success that can be simultaneously satisfied by gadgets that exist in practice. These developments, together with the anti-gadget concept, come together in the Group Lemma, which provides a straightforward criterion for proving #P-hardness of certain Holant problems.

Definition 5.3.1. A set of matrices $\mathcal{M} \subseteq \mathbb{C}^{2^m \times 2^n}$ forms a *projective set from arity n to m* if for any matrix $N \in \mathbb{C}^{2^n \times 2}$ with rank 2, there exists a matrix $M \in \mathcal{M}$ such that MN has rank 2.

We also call a set of gadgets *projective* from arity n to m if the set of its signature matrices is projective from arity n to m. A $(\mathcal{G} \mid \mathcal{R})$-gate set that is projective from arity 2-to-1 can be used to transform a pair of $(\mathcal{G} \mid \mathcal{R})$-gates with linearly independent binary signatures to a pair of $(\mathcal{G} \mid \mathcal{R})$-gates with linearly independent unary signatures. Projective gadgets in such a set have two trailing edges and one leading edge, but can also be viewed as operating on signatures of higher arity, with the identity transformation being performed on the other edges not connected to the projective gadget. This way of connecting the projective gadget to an existing $(\mathcal{G} \mid \mathcal{R})$-gate *automatically* gives us projective gadget sets for higher arities. But first, a quick lemma to assist with the proof.

Lemma 5.3.2. *Let v_0 and v_1 be nonzero column vectors, not necessarily the same length. Then the block matrix $\left[\begin{smallmatrix} av_0 & bv_0 \\ cv_1 & dv_1 \end{smallmatrix}\right]$ has rank 2 if and only if $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ is invertible.*

Proof. We write $\left[\begin{smallmatrix} av_0 & bv_0 \\ cv_1 & dv_1 \end{smallmatrix}\right] = \left[\begin{smallmatrix} v_0 & 0 \\ 0 & v_1 \end{smallmatrix}\right]\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$, and $\left[\begin{smallmatrix} v_0 & 0 \\ 0 & v_1 \end{smallmatrix}\right]$ has rank 2. Thus $\left[\begin{smallmatrix} av_0 & bv_0 \\ cv_1 & dv_1 \end{smallmatrix}\right]$ has rank 2 if and only if $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ is invertible. $\square$

Lemma 5.3.3. *Let $\mathcal{G}$ and $\mathcal{R}$ be signature sets. Suppose $\mathcal{P}$ is a projective $(\mathcal{G} \mid \mathcal{R})$-gate set from arity 2-to-1. Then for all integers $k \geq 2$, $\mathcal{P}$ acts as a projective $(\mathcal{G} \mid \mathcal{R})$-gate set from arity k to $k-1$.*

Proof. We are given that for any $N \in \mathbb{C}^{4\times 2}$ with rank 2, there exists an $F \in \mathcal{P}$ with $F \in \mathbb{C}^{2\times 4}$ such that FN is invertible. We want to show that for any integer $k \geq 2$ and any rank 2 matrix $B \in \mathbb{C}^{2^k\times 2}$, there exists an $F \in \mathcal{P}$ such that $(I \otimes F)B$ has rank 2, where I is the 2^{k-2}-by-2^{k-2} identity matrix.

For any $F \in \mathcal{P}$, the matrix $I \otimes F$ can be viewed as being composed of 2-by-4 blocks, with F appearing along the main diagonal and 2-by-4 zero-matrices elsewhere. We similarly view B as being composed of 4-by-2 blocks $B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_{2^{k-2}} \end{bmatrix}$. Then $(I \otimes F)B = \begin{bmatrix} FB_1 \\ FB_2 \\ \vdots \\ FB_{2^{k-2}} \end{bmatrix} \in \mathbb{C}^{2^{k-1}\times 2}$. If some B_i has rank 2, then there is an $F \in \mathcal{P}$ such that FB_i is invertible and $(I \otimes F)B$ has rank 2, as desired.

Now assume otherwise, so each B_i has rank at most 1. Since B has rank 2, there exists a 2-by-2 invertible submatrix D of B, for which the rows of D appear in B_i and B_j, for some $i < j$. It follows that B_i and B_j both have rank exactly 1. Hence for some nonzero vectors $v_0, v_1 \in \mathbb{C}^4$ and some $a, b, c, d \in \mathbb{C}$, we can write $B_i = [av_0\ bv_0]$ and $B_j = [cv_1\ dv_1]$. By Lemma 5.3.2, $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$ is invertible, as $\left[\begin{smallmatrix} B_i \\ B_j \end{smallmatrix}\right]$ has rank 2. If v_0 and v_1 are linearly independent, then choose $F \in \mathcal{P}$ such that $F[v_0\ v_1]$ is invertible; otherwise let $\tilde{v} \in \mathbb{C}^4$ be such that v_0 and $\tilde{v}$ are linearly independent, and choose $F \in \mathcal{P}$ such that $F[v_0\ \tilde{v}]$ is invertible. In either case (ignoring $\tilde{v}$ in the second case), we define $[v_0'\ v_1'] = F[v_0\ v_1]$, where v_0' and v_1' are nonzero. Then by Lemma 5.3.2, the matrix $\left[\begin{smallmatrix} av_0' & bv_0' \\ cv_1' & dv_1' \end{smallmatrix}\right] = \left[\begin{smallmatrix} FB_i \\ FB_j \end{smallmatrix}\right]$ has rank 2, and since this appears as a submatrix of $(I \otimes F)B$, we are done. $\square$

Corollary 5.3.4. *Let $\mathcal{G}$ and $\mathcal{R}$ be signature sets. Suppose $\mathcal{P}$ is a finite projective $(\mathcal{G} \mid \mathcal{R})$-gate set from arity 2-to-1. Then for any integer $k \geq 2$, $\mathcal{P}$ induces a finite projective $(\mathcal{G} \mid \mathcal{R})$-gate set from arity k to 1.*

Now we show that a finite projective $(\mathcal{G} \mid \mathcal{R})$-gadget set from arity k to 1 preserves pairwise linear independence for an inverse polynomial fraction of signatures. The essence of the next lemma is an exchange in the order of quantifiers.

Lemma 5.3.5. *Suppose $\{v_i\}_{i \geq 0}$ is a sequence of pairwise linearly independent column vectors in $\mathbb{C}^{2^k}$ and let $\mathcal{F} \subseteq \mathbb{C}^{2 \times 2^k}$ be a finite set of f matrices that is projective from arity k to 1. Then for every n, there exists some $F \in \mathcal{F}$ and some $S \subseteq \{Fv_i \mid 0 \leq i \leq n^f\}$ such that $|S| \geq n$ and the vectors in S are pairwise linearly independent.*

Proof. Let $j > i \geq 0$ be integers and let $N = [v_i \ v_j] \in \mathbb{C}^{2^k \times 2}$. Since v_i and v_j are linearly independent, $\operatorname{rank}(N) = 2$. By assumption, there exists an $F \in \mathcal{F}$ such that $FN \in \mathbb{C}^{2 \times 2}$ is invertible, so we conclude that Fv_i and Fv_j are linearly independent.

Each $F \in \mathcal{F}$ defines a coloring of the set $K = \{0, 1, \ldots, n^f\}$ as follows: color $i \in K$ with the linear subspace spanned by Fv_i. Assume for a contradiction that for each $F \in \mathcal{F}$, there is not n pairwise linearly independent vectors among $\{Fv_i \mid i \in K\}$. Then, including possibly the 0-dimensional subspace $\{0\}$, there can be at most n distinct colors assigned by each $F \in \mathcal{F}$. By the pigeonhole principle, some i and j with $0 \leq i < j \leq n^f$ must receive the same color for all $F \in \mathcal{F}$. This is a contradiction with the previous paragraph, so we are done. □

The next lemma says that under suitable conditions, we can construct all unary signatures (X, Y). The method will be interpolation at a higher dimensional iteration in a circular fashion and finishing with an appropriate projective gadget.

Lemma 5.3.6 (Group Lemma)**.** *For signature sets $\mathcal{G}$ and $\mathcal{R}$, suppose there exists a finite set of projective $(\mathcal{G} \mid \mathcal{R})$-gates from arity 2-to-1, and suppose $\mathcal{S}$ is a finite set of recursive $(\mathcal{G} \mid \mathcal{R})$-gates of arity $d \geq 1$ with nonsingular transition matrices. Let H be the group generated by the transition matrices of gadgets in $\mathcal{S}$, modulo scalar matrices λI, for $\lambda \in \mathbb{C} - \{0\}$. If H has infinite order, then*

$$\operatorname{Holant}(\mathcal{G} \cup \{(X, Y)\} \mid \mathcal{R}) \leq_T \operatorname{Holant}(\mathcal{G} \mid \mathcal{R})$$

for any $X, Y \in \mathbb{C}$.

Proof. Two matrices are unequal modulo scalar matrices λI if and only if they are linearly independent. If any member of $\mathcal{S}$, as a group element in H, has infinite order, then its powers supply an infinite set of pairwise linearly independent signatures. Otherwise they all have finite order, and the group H is identical to the monoid generated by $\mathcal{S}$, i.e., every $h \in H$ is a product over $\mathcal{S}$ with non-negative powers. Such products give a composition of gadgets in $\mathcal{S}$, which is a recursive gadget. By assumption, H has infinite order, so by composing recursive gadgets from $\mathcal{S}$, a breadth-first traversal of the Cayley graph of the monoid generated by $\mathcal{S}$ supplies an arbitrarily large set of recursive gadgets having pairwise linearly independent signatures.

Before we can use a projective gadget set to project the set of pairwise linearly independent signatures down to arity 1, we make a small modification to each corresponding gadget: connect a non-degenerate signature $g \in \mathcal{G}$ to every trailing edge. This ensures that the bipartite structure of the graph is preserved when applying projective gadgets. We claim that there is some non-degenerate signature $g \in \mathcal{G}$. If this were not the case, then any recursive gadget $s \in \mathcal{S}$ (note $\mathcal{S}$ is nonempty) could be rewritten with all leading edges internally incident to unary signatures. The recurrence matrix of such a gadget is expressible as a product of a column vector and a row vector (by partitioning s into two gadgets with no shared edges), hence the recurrence matrix of s would have rank at most 1, which is less than 2^d as promised. Let $a \geq 2$ be the arity of g. One can show by induction that any non-degenerate signature has at least one index i, such that if we express the signature as a 2-by-2^{a-1} matrix M indexed by the i-th variable for the row and the remaining $a-1$ variables for the column, then M has rank 2. We designate one such dangling edge of g as the leading edge and all other dangling edges as trailing edges. As there are d trailing edges in s, we apply d copies of g, which corresponds to multiplication by the matrix $M^{\otimes d}$. Since M has full rank, pairwise linear independence of the signatures is preserved. Now rewrite the 2^d-by-$2^{d(a-1)}$ matrix form of the signature as a column vector in $\mathbb{C}^{2^{da}}$, indexed by $c_{d(a-1)} \cdots c_1 b_1 \cdots b_d \in \{0,1\}^{da}$, where $b_1 \cdots b_d$ and $c_1 \cdots c_{d(a-1)}$ are the row and column indices. Denote these vectors as $\{v_i\}_{i \geq 0}$. Finally we can attach projective gadgets to project each v_i down to arity 1. (see Figure 5.2c).

To show $\operatorname{Holant}(\mathcal{G} \cup \{(X,Y)\} \mid \mathcal{R}) \leq_T \operatorname{Holant}(\mathcal{G} \mid \mathcal{R})$, suppose we are given as input a bipartite signature grid Ω for $\operatorname{Holant}(\mathcal{G} \cup \{(X,Y)\} \mid \mathcal{R})$, with underlying graph $G = (V, E)$. Let $Q \subseteq V$ be the set of vertices labeled with (X, Y), and let $n = |Q|$. By Corollary 5.3.4, there exists a finite projective set containing f gadgets from arity d to 1, so by Lemma 5.3.5, there is some projective gadget F in this set such that at least $n + 2$ of the first $(n + 2)^f + 1$ vectors of the form Fv_t are pairwise linearly independent. It is straightforward to efficiently find such a set; denote it by $S = \{(X_0, Y_0), (X_1, Y_1), \ldots, (X_{n+1}, Y_{n+1})\}$, and let $G_0, G_1, \ldots, G_{n+1}$ be the corresponding gadgets. At most one Y_t can be zero, so without loss of generality, assume $Y_t \neq 0$ for $0 \leq t \leq n$. If we replace every element of Q with a copy of G_t, we obtain an instance of Holant$(\mathcal{G} \mid \mathcal{R})$ (note that the correct bipartite structure is preserved), and we denote this new signature grid by Ω_t. Although Holant(Ω_t) is a sum of exponentially many terms, each nonzero term has the form $bX_t^i Y_t^{n-i}$ for some i and for some $b \in \mathbb{C}$ that does not depend on X_t or Y_t. Then for some $c_0, c_1, \ldots, c_n \in \mathbb{C}$, the sum can be rewritten as

$$\text{Holant}(\Omega_t) = \sum_{0 \leq i \leq n} c_i X_t^i Y_t^{n-i}.$$

Since each signature grid Ω_t is an instance of Holant$(\mathcal{G} \mid \mathcal{R})$, Holant$(\Omega_t)$ can be solved exactly using the oracle. Carrying out this process for every t where $0 \leq t \leq n$, we arrive at a linear system where the c_i values are the unknowns.

$$\begin{bmatrix} Y_0^{-n} \cdot \text{Holant}(\Omega_0) \\ Y_1^{-n} \cdot \text{Holant}(\Omega_1) \\ \vdots \\ Y_n^{-n} \cdot \text{Holant}(\Omega_n) \end{bmatrix} = \begin{bmatrix} X_0^0 Y_0^0 & X_0^1 Y_0^{-1} & \cdots & X_0^n Y_0^{-n} \\ X_1^0 Y_1^0 & X_1^1 Y_1^{-1} & \cdots & X_1^n Y_1^{-n} \\ \vdots & \vdots & \ddots & \vdots \\ X_n^0 Y_n^0 & X_n^1 Y_n^{-1} & \cdots & X_n^n Y_n^{-n} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{bmatrix}$$

The matrix above has entry $(X_r/Y_r)^c$ at row r and column c. Due to pairwise linear independence of (X_r, Y_r), we have that X_r/Y_r is pairwise distinct for $0 \leq r \leq n$. Hence this is a Vandermonde system of full rank, and we can solve it for the c_i values. With these values in hand, we can calculate

$$\text{Holant}(\Omega) = \sum_{0 \leq i \leq n} c_i X^i Y^{n-i},$$

completing the reduction. □

Remark. Every time we apply Lemma 5.3.6, we do so in the simpler case that H contains an element of infinite order. We had initially hoped to find a useful failure condition from group theory when H has finite order, but we did not find such a condition.

Here is how we realize a projective set of gadgets from arity 2-to-1.

Lemma 5.3.7. *Let* $\Phi_i \in \mathbb{C}^{2\times 4}$ *for* $1 \le i \le 7$ *be matrices with the following properties:* $\ker(\Phi_i) = \operatorname{span}\{u, u_i\}$ *for* $1 \le i \le 3$*,* $\ker(\Phi_{i+3}) = \operatorname{span}\{v, v_i\}$ *for* $1 \le i \le 3$*,* $\ker(\Phi_7) = \operatorname{span}\{s, t\}$*, and* $\dim\{u, u_1, u_2, u_3\} = \dim\{v, v_1, v_2, v_3\} = \dim\{u, v, s, t\} = 4$*. Then* $\{\Phi_i \mid 1 \le i \le 7\}$ *is projective from arity* 2*-to-*1*.*

Proof. Let $N \in \mathbb{C}^{4\times 2}$ be a rank 2 matrix with $\operatorname{im}(N) = \operatorname{span}\{w_1, w_2\}$. If $\operatorname{im}(N) = \operatorname{span}\{u, v\}$, then $\Phi_7 N$ has rank 2. Otherwise, either $\{w_1, w_2, u\}$ or $\{w_1, w_2, v\}$ is linearly independent. Say $\{w_1, w_2, u\}$ is linearly independent. The other case is similar. Then $\{w_1, w_2, u\}$ can be further augmented by some u_i for $1 \le i \le 3$ to form a basis, in which case $\Phi_i N$ has rank 2. □

Remark. Michael Kowalczyk and I went back-and-forth to find this theoretical condition that is satisfied by gadgets in practice. He eventually found this condition, and I found the gadgets that satisfied it. The condition is stated in terms of vector spaces, which is how we use it. However, the proof applies to the more general case of matroids. I prefer thinking of it in terms of matroids. To avoid having to introduce matroids and for consistency of presentation, we stick to the language of vector spaces.

Verifying that a specific set of gadgets forms a projective set from arity 2-to-1 only requires a straightforward linear algebra computation. Note that the exceptional cases are either symmetric signatures (for which a dichotomy exists [93]) or largely correspond to tractable cases.

Lemma 5.3.8. *Let* $w, x, y, z \in \mathbb{C}$*. Then there exists a finite projective* $((w, x, y, z) \mid =_3)$*-gate set from arity 2-to-1 unless* $x = y \lor wz = xy \lor (w, z) = (0, 0) \lor (x, y) = (0, 0) \lor (w^3 = -z^3 \land x = -y)$*.*

Proof. Let $f = (w, x, y, z)$. We are given that

$$x \ne y \land wz \ne xy \land (w, z) \ne (0, 0) \land (x, y) \ne (0, 0) \land (w^3 \ne -z^3 \lor x \ne -y).$$

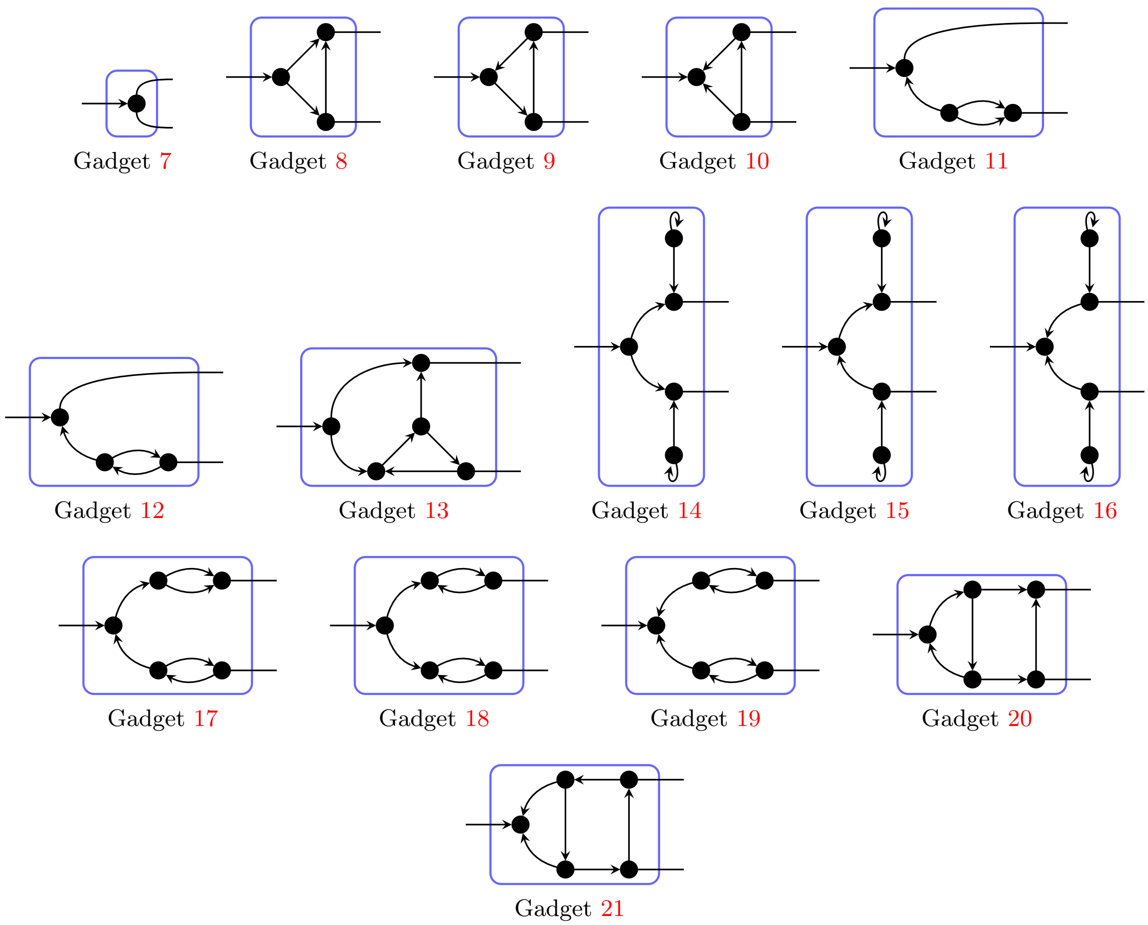


Figure 5.6: Projective gadgets from arity 2-to-1

Table 5.1: This table indicates which projective gadgets are used in each case (and the role of each gadget within each case) in the proof of Lemma 5.3.8. The seven Φ_i refer to the matrices in Lemma 5.3.7. As an example, the projective set in case 1 is $\{F_7, F_{14}, F_{16}, F_7, F_8, F_{10}, F_{12}\}$. Note that F_7 plays the role of both Φ_1 and Φ_4.

<table>
<tr><th></th><th>Φ_1</th><th>Φ_2</th><th>Φ_3</th><th>Φ_4</th><th>Φ_5</th><th>Φ_6</th><th>Φ_7</th></tr>
<tr><td>Case 1</td><td rowspan="5">F_7</td><td>F_{14}</td><td rowspan="2">F_{16}</td><td rowspan="5">F_7</td><td rowspan="5">F_8</td><td>F_{10}</td><td rowspan="2">F_{12}</td></tr>
<tr><td>Case 2</td><td rowspan="3">F_9</td><td>F_{15}</td></tr>
<tr><td>Case 3</td><td>F_{21}</td><td>F_{20}</td><td>F_{13}</td></tr>
<tr><td>Case 4</td><td rowspan="2">F_{18}</td><td>F_{17}</td><td rowspan="2">F_{11}</td></tr>
<tr><td>Case 5</td><td>F_{19}</td><td>F_{10}</td></tr>
</table>

It suffices to exhibit projective gadget sets that satisfy the hypotheses of Lemma 5.3.7, which is what we do.

Let F_i be the transition matrix of Gadget i for $7 \leq i \leq 21$. There are five cases of projective $(f \mid =_3)$-gadget sets from arity 2-to-1. We omit the verification that each set of projective gadgets forms a projective gadget set from arity 2-to-1 under its particular assumptions since this is a straightforward linear algebra computation. The five cases are

1. $wz \neq xy \wedge wxyz \neq 0 \wedge w^3x + wxyz + w^2z^2 + yz^3 \neq 0 \wedge x^2 \neq y^2$,
2. $wz \neq xy \wedge wxyz \neq 0 \wedge w^3x + wxyz + w^2z^2 + yz^3 \neq 0 \wedge x = -y \wedge w^3 \neq -z^3$,
3. $wz \neq xy \wedge wxyz \neq 0 \wedge w^3x + wxyz + w^2z^2 + yz^3 = 0 \wedge x \neq y$,
4. $wz \neq xy \wedge w = 0 \wedge z \neq 0 \wedge x \neq y$, and
5. $wz \neq xy \wedge x = 0 \wedge y \neq 0$.

Which projective gadgets are used in each case (and the role of each gadget within each case) can be found in Table 5.1. In all five cases, the vector u in the kernels of Φ_1, Φ_2, and Φ_3 is $(0, -1, 1, 0)$ and the vector v in the kernels of Φ_4, Φ_5, and Φ_6 is $(0, -x, y, 0)$.

All five cases utilize the assumption $wz \neq xy$, i.e., the edge signature is non-degenerate. Under three additional disequality assumptions, the projective gadgets in row 1 of Table 5.1 have the desired properties. The purpose of the remaining four cases is to handle the situation that these three disequalities are not all true.

Case 2 retains two of the additional disequality assumptions but assumes that $x^2 = y^2$. Since we are considering the asymmetric case, the only option is $x = -y$. By assumption, it is not the case that $x = -y \wedge w^3 = -z^3$, so we have $w^3 \neq -z^3$. Under these conditions, the projective gadgets in row 2 of Table 5.1 have the desired properties.

Like cases 1 and 2, case 3 retains the assumption that no variable is zero but now considers the case that the polynomial $w^3x + wxyz + w^2z^2 + yz^3$ is zero. Given that we are also considering the asymmetric case (i.e. $x \neq y$), the projective gadgets in row 3 of Table 5.1 have the desired properties.

Cases 4 and 5 handle the remaining case $wz \neq xy \wedge wxyz = 0$. The assumptions $wz \neq xy \wedge (w, z) \neq (0, 0) \wedge (x, y) \neq (0, 0)$ imply that at most one of w, x, y, and z is 0. By switching the

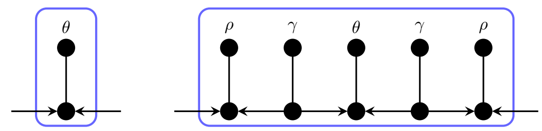


Figure 5.7: Gadgets on the left used to simulate $(0, 1, 1, 1)$.

role of 0 and 1 via the holographic transformation $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, the complexity of the case $yz = 0$ is the same as the complexity of the case $wx = 0$. Therefore, we assume that $yz \neq 0$. Case 4 considers $w = 0$, so $z \neq 0$ by assumption. Then still within the asymmetric case, the projective gadgets in row 4 of Table 5.1 have the desired properties. Case 5 considers $x = 0$, so $y \neq 0$ by assumption and the projective gadgets in row 5 of Table 5.1 have the desired properties.

These five cases cover all settings not excluded by the assumptions in the statement of the lemma, so the proof is complete. □

Once we have all unary signatures at our disposal, we can prove #P-hardness under most settings. To do so, we make use of the following lemma.

Lemma 5.3.9 (Lemma 3.3 of [93])**.** *Suppose that* $(a, b) \in \mathbb{C}^2 - \{(a, b) \mid ab = 1\} - (0, 0)$ *and let* $\mathcal{G}$ *and* $\mathcal{R}$ *be finite signature sets where* $(a, 1, 1, b) \in \mathcal{G}$ *and* $=_3 \in \mathcal{R}$*. Further assume that* $\operatorname{Holant}(\mathcal{G} \cup \{(X_i, Y_i)\} \mid 0 \leq i < m\} \mid \mathcal{R}) \leq_T \operatorname{Holant}(\mathcal{G} \mid \mathcal{R})$ *for any* $X_i, Y_i \in \mathbb{C}$ *and* $m \in \mathbb{Z}^+$*. Then* $\operatorname{Holant}(\mathcal{G} \cup \{(0, 1, 1, 1)\} \mid \mathcal{R}) \leq_T \operatorname{Holant}(\mathcal{G} \mid \mathcal{R})$ *and* $\operatorname{Holant}(\mathcal{G} \mid \mathcal{R})$ *is* #P*-hard.*

Lemma 5.3.10. *Let* $w, x, y, z \in \mathbb{C}$*, and let* $\mathcal{G}$ *and* $\mathcal{R}$ *be finite signature sets with* $(w, x, y, z) \in \mathcal{G}$ *and* $=_3 \in \mathcal{R}$*. Suppose* $\operatorname{Holant}(\mathcal{G} \cup \{(X_i, Y_i) \mid 0 \leq i < m\} \mid \mathcal{R}) \leq_T \operatorname{Holant}(\mathcal{G} \mid \mathcal{R})$ *for any* $X_i, Y_i \in \mathbb{C}$ *and* $m \in \mathbb{Z}^+$*. Then* $\operatorname{Holant}(\mathcal{G} \mid \mathcal{R})$ *is* #P*-hard unless* $wz = xy \vee (w, z) = (0, 0) \vee (x, y) = (0, 0)$.

Proof. Since $\operatorname{Holant}((0, 1, 1, 1) \mid =_3)$, #VertexCover on 3-regular graphs, is #P-hard, we only need to show how to simulate the signature $(0, 1, 1, 1)$ on the left. The assumptions $wz \neq xy \wedge (w, z) \neq (0, 0) \wedge (x, y) \neq (0, 0)$ imply that at most one of w, x, y, and z is 0. By switching the role of 0 and 1 via the holographic transformation $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, the complexity of the case $yz = 0$ is the same as the complexity of the case $wx = 0$. Therefore, we assume that $yz \neq 0$.

If $w = 0$, then Gadget 22 with $\theta = \frac{1}{x}(\frac{z}{y^2}, \frac{1}{z})$ simulates $(\frac{x}{z}, 1, 1, \frac{2z}{x})$, which can in turn simulate $(0, 1, 1, 1)$ by Lemma 5.3.9. If $x = 0$, then Gadget 22 with $\theta = \frac{1}{w}(\frac{1}{y}, \frac{y}{z^2})$ simulates $(\frac{w}{y}, 1, 1, \frac{2y}{w})$, which can in turn simulate $(0, 1, 1, 1)$ by Lemma 5.3.9. If $wx \neq 0 \wedge wz = -xy$, then Gadget 22 with $\theta = \frac{1}{xy}(\frac{2x}{w}, \frac{w}{x})$ simulates $(\frac{3w}{y}, 1, 1, \frac{3y}{w})$, which can in turn simulate $(0, 1, 1, 1)$ by Lemma 5.3.9. Finally if $wx \neq 0 \wedge wz \neq xy \wedge wz \neq -xy$, then Gadget 23 with $\theta = \frac{wz+xy}{wx(wz-xy)}(\frac{-x}{w}, \frac{w}{x})$, $\gamma = (\frac{1}{wx}, \frac{-wx}{yz(wz+xy)})$, and $\rho = \frac{1}{wz-xy}(xz, -wy)$ simulates $(0, 1, 1, 1)$. □

Remark. Let me explain how to pick θ, γ, and ρ in Gadget 23 given $wx \neq 0 \wedge wz \neq xy \wedge wz \neq -xy$. Initially, we set $\theta = (\frac{-x}{w}, \frac{w}{x})$ so that the signature of Gadget 22 has 0 for its entry of Hamming weight 0. Specifically, its signature matrix is $(wz - xy)\left[\begin{smallmatrix} 0 & 1 \\ 1 & \frac{wz+xy}{wx} \end{smallmatrix}\right]$. This signature is also symmetric, but its entries of Hamming weight 1 are different from its entry of Hamming weight 2. The purpose of γ and ρ is to make these two entries equal. A factor of $wz - xy$ has also appeared, so we cancel it by setting $\theta = \frac{1}{wz-xy}(\frac{-x}{w}, \frac{w}{x})$.

Initially, we set $\rho = (xz, -wy)$ so that, if we have $\rho = \theta$ in Gadget 22, then its signature matrix is diagonal. Specifically, this matrix is $(wz - xy)\left[\begin{smallmatrix} wx & 0 \\ 0 & -yz \end{smallmatrix}\right]$. The contribution from γ in Gadget 23 will also be a diagonal matrix, so things are very easy to analyze now. Another factor of $wz - xy$ has appeared, so we cancel it by setting $\rho = \frac{1}{wz-xy}(xz, -wy)$.

To simplify things further, we think of γ coming in two parts, γ_{out} and γ_{in}, such that $\gamma = \gamma_{\text{out}}\gamma_{\text{in}} = \gamma_{\text{in}}\gamma_{\text{out}}$. These are used to cancel terms on the "outside" (i.e. the ρ part) and the "inside" (i.e. the θ part) respectively. By setting $\gamma_{\text{out}} = (\frac{1}{wx}, \frac{-1}{yz})$, the signature matrix of the outside is I_2. Then we set $\gamma_{\text{in}} = (1, \frac{wx}{wz+xy})$ so that the signature matrix of the inside is $\frac{wx}{wz+xy}\left[\begin{smallmatrix} 0 & 1 \\ 1 & 1 \end{smallmatrix}\right]$. To cancel the remaining factor, we set $\theta = \frac{wz+xy}{wx(wz-xy)}(\frac{-x}{w}, \frac{w}{x})$.

By combining the Group Lemma (Lemma 5.3.6) with Lemma 5.3.8 and Lemma 5.3.10, we get the following theorem.

Theorem 5.3.11. *Let $w, x, y, z \in \mathbb{C}$, and let $\mathcal{G}$ and $\mathcal{R}$ be finite signature sets where $(w, x, y, z) \in \mathcal{G}$ and $(=_3) \in \mathcal{R}$. Suppose $\mathcal{S}$ is a finite set of recursive $(\mathcal{G} \mid \mathcal{R})$-gates of arity $d \geq 1$ with nonsingular transition matrices. Let H be the group generated by the transition matrices of $\mathcal{S}$, modulo scalar matrices λI, for $\lambda \in \mathbb{C} - \{0\}$. If H has infinite order, then* Holant$(\mathcal{G} \mid \mathcal{R})$ *is* #P*-hard unless*

$x = y \vee wz = xy \vee (w, z) = (0, 0) \vee (x, y) = (0, 0) \vee (w^3 = -z^3 \wedge x = -y)$.

5.4 Statement of Main Result

Theorem 5.4.1. *Let* $w, x, y, z \in \mathbb{C}$ *and* $f = (w, x, y, z)$*. Then* $\operatorname{Holant}(f \mid =_3)$ *is* #P*-hard unless one of the following conditions holds, in which case, the problem is computable in polynomial time:*

1. degenerate: $wz = xy$;
2. generalized disequality: $w = z = 0$;
3. generalized equality: $x = y = 0$;
4. holographic reduction to affine: $wz = -xy \wedge w^6 = \varepsilon z^6 \wedge x^2 = \varepsilon y^2$, where $\varepsilon = \pm 1$.

If the input is restricted to planar graphs, then another case becomes computable in polynomial time but everything else remains #P*-hard:*

5. holographic reduction to matchgates: $w^3 = \varepsilon z^3 \wedge x = \varepsilon y$, where $\varepsilon = \pm 1$.

We prove the tractability part of Theorem 5.4.1 now.

Proof of tractability. In Case 1, case 2, and case 3, both signatures are of product type, so we are done by Corollary 4.1.5.

For case 4, if $w = z = 0$, then this is already covered by case 2. Otherwise $wz \neq 0$ since $w^6 = \varepsilon z^6$, in which case we apply the holographic transformation $\left[\begin{smallmatrix} \alpha & 0 \\ 0 & \alpha^2 \end{smallmatrix}\right]$ with $\alpha = \varepsilon w^2/z^2$, which does not change the Holant value by Lemma 3.2.1. Note that $\alpha^3 = \varepsilon w^6/z^6 = 1$. The edge signature becomes $(w, x, y, z) \left[\begin{smallmatrix} \alpha & 0 \\ 0 & \alpha^2 \end{smallmatrix}\right]^{\otimes 2} = (\alpha^2 w, x, y, \alpha z)$, while $=_3$ is unchanged since $(\left[\begin{smallmatrix} \alpha & 0 \\ 0 & \alpha^2 \end{smallmatrix}\right]^{-1})^{\otimes 3} = I_8$, the 8-by-8 identity matrix. This reduces to the case $wz = -xy \wedge w^2 = \varepsilon z^2 \wedge x^2 = \varepsilon y^2$. This edge signature (as well as $=_3$) is affine, so we are done by Theorem 4.2.6.

For case 5, if the input is restricted to planar graphs, apply the same holographic transformation as in the previous case but with $\alpha = \varepsilon z/w$. Again, $\alpha^3 = \varepsilon z^3/w^3 = 1$, so the edge signature becomes $(\alpha^2 w, x, y, \alpha z)$ and $=_3$ is still unchanged. This reduces to the case $w = \varepsilon z \wedge x = \varepsilon y$. Then after a further holographic transformation by the Hadamard matrix $\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$, both signatures are matchgate signatures, so we are done by Theorem 4.3.2. $\square$

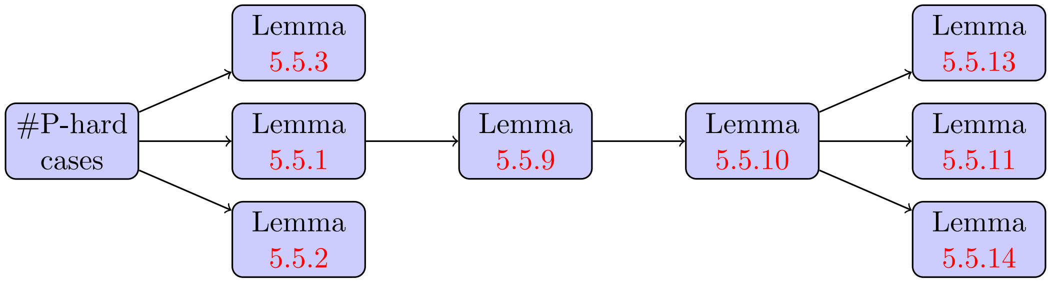


Figure 5.8: Lemmas used to prove the existence of a recursive gadget with a transition matrix that has infinite order (up to a scalar). The directed edges roughly indicate which lemma handles cases excluded by a previous lemma.

The proof of #P-hardness begins in Section 5.5. By Theorem 5.3.11, it suffices to find a recursive gadget with a transition matrix that has infinite order (up to a scalar) for each case that we did not just prove is tractable. Figure 5.8 provides a graphical view of the order in which cases are handled. Since we are proving a classification theorem involving four variables, one should expect the depth of this tree to be at least 4. A naive case analysis would produce a tree with an average branching of factor of at least 2 or 3, which would give a rather large tree width. Given our new theoretical tools (like anti-gadgets) and our thorough consideration of the possible gadgets, the tree representing our cases analysis has minimal depth and almost no branching.

5.5 Anti-Gadgets in Action

Now we use our new idea of anti-gadgets to construct explicit matrices of infinite order.

Lemma 5.5.1. *If $wz \neq xy$, $wxyz \neq 0$, and $|x| \neq |y|$, then* $\operatorname{Holant}((w,x,y,z) \mid =_3)$ *is* #P*-hard.*

Proof. The transition matrices for Gadget 4 and Gadget 5 are $M_4 = \left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]^{\otimes 2} \operatorname{diag}(w,x,y,z)$ and $M_5 = \left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]^{\otimes 2} \operatorname{diag}(w,y,x,z)$, both nonsingular. Since the matrix $M_4^{-1} M_5 = \operatorname{diag}(1, y/x, x/y, 1)$ has infinite order up to a scalar, we are done by Theorem 5.3.11. $\square$

Lemma 5.5.2. *If $x = 0$ and $wyz \neq 0$, then* $\operatorname{Holant}((w,x,y,z) \mid =_3)$ *is* #P*-hard.*

Proof. The transition matrices for Gadget 2 and Gadget 3 are $M_2 = \left[\begin{smallmatrix} w & 0 \\ y & z \end{smallmatrix}\right] \left[\begin{smallmatrix} w^2 & y^2 \\ 0 & z^2 \end{smallmatrix}\right]$ and $M_3 = \left[\begin{smallmatrix} w & 0 \\ y & z \end{smallmatrix}\right] \left[\begin{smallmatrix} w^2 & 0 \\ 0 & z^2 \end{smallmatrix}\right]$, both nonsingular. Since $M_3^{-1} M_2 = \left[\begin{smallmatrix} 1 & y^2/w^2 \\ 0 & 1 \end{smallmatrix}\right]$ has infinite order up to a scalar, we are done by Theorem 5.3.11. $\square$

Lemma 5.5.3. *If $w = 0$ and $xyz \neq 0$, then* $\operatorname{Holant}((w, x, y, z) \mid =_3)$ *is* #P-*hard.*

Proof. The transition matrices for Gadget 3 and Gadget 6 are $M_3 = \left[\begin{smallmatrix} 0 & x \\ y & z \end{smallmatrix}\right] \left[\begin{smallmatrix} 0 & xy \\ xy & z^2 \end{smallmatrix}\right]$ and $M_6 = \left[\begin{smallmatrix} 0 & x \\ y & z \end{smallmatrix}\right] \left[\begin{smallmatrix} 0 & xyz^3 \\ xyz^3 & xy^2z^2+z^5 \end{smallmatrix}\right]$, both nonsingular. Since $M_3^{-1} M_6 = z^3 \left[\begin{smallmatrix} 1 & y/z \\ 0 & 1 \end{smallmatrix}\right]$ has infinite order up to a scalar, we are done by Theorem 5.3.11. $\square$

For the remainder of the proof of #P-hardness of Theorem 5.4.1, we use our anti-gadget technique in combination with Lemma 5.5.4, Lemma 5.5.5, and Lemma 5.5.6. In the contrapositive, these lemmas provide sufficient conditions to conclude that a matrix has infinite order (up to a scalar). Their proofs follow from a few observations, some of which are known as Vieta's formulas.

For monic polynomials in $\mathbb{C}[X]$ of degree n with roots λ_i for $1 \leq i \leq n$ of the same nonnegative norm $r \in \mathbb{R}$, let $a_k \in \mathbb{C}$ be the coefficient of X^k and σ_k the elementary symmetric polynomial of degree k in λ_i / r for $1 \leq i \leq n$, the norm one (scaled) roots.[1] Thus $a_k = (-r)^{n-k} \sigma_{n-k}$. By having norm one, $\sigma_k = \overline{\sigma_{n-k}} \sigma_n$, $a_k = (-1)^n r^{n-2k} \overline{a_{n-k}} \sigma_n$, and $|a_k| = r^{n-2k} |a_{n-k}|$, for $0 \leq k < n$.

Lemma 5.5.4 (Lemma 4.4 in [93])**.** *If both roots of $X^2 + a_1 X + a_0 \in \mathbb{C}[X]$ have the same norm, then $a_1 |a_0| = \overline{a_1} a_0$. If further $a_0 a_1 \neq 0$, then $\operatorname{Arg}(a_1^2) = \operatorname{Arg}(a_0)$ thus $a_1^2 / a_0 \in \mathbb{R}^+$.*

Lemma 5.5.5. *If all roots of $X^4 + a_3 X^3 + a_2 X^2 + a_1 X + a_0 \in \mathbb{C}[X]$ have the same norm, then $a_2 |a_1|^2 = |a_3|^2 \overline{a_2} a_0$.*

Lemma 5.5.6. *If $\sum_{k=0}^{8} a_k X^k \in \mathbb{C}[X]$ is monic and all roots have the same norm, then $a_3^2 |a_1|^2 = |a_7|^2 \overline{a_5}^2 a_0^2$, $a_4 |a_2|^2 = |a_6|^2 \overline{a_4} a_0$, and $|a_3|^2 a_2 = \overline{a_6} |a_5|^2 a_0$.*

Over 3-regular directed graphs, there are some symmetries under which the Holant is invariant. The next lemma states these symmetries.

Lemma 5.5.7. *Let G be a 3-regular directed graph. Then there exists a polynomial P with integer coefficients in six variables, such that for any signature grid Ω having underlying graph G with vertex signature $=_3$ and edge signature (w, x, y, z),*

$$\operatorname{Holant}(\Omega) = P(wz, xy, w^3 + z^3, x + y, w^3 x + y z^3, w^3 y + x z^3).$$

[1] This argument assumes $r \neq 0$. However, when $r = 0$, the conclusion still holds trivially.

Proof. Consider any $0, 1$ vertex assignment σ with a nonzero valuation. If σ' is the complement assignment switching all 0's and 1's in σ, then for σ and σ', we have the sum of valuations $w^a x^b y^c z^d + w^d x^c y^b z^a$ for some a, b, c, d. Here a (resp. d) is the number of edges connecting two degree 3 vertices both assigned 0 (resp. 1) by σ. Similarly, b (resp. c) is the number of edges from one degree 3 vertex to another that are assigned 0 and 1 (resp. 1 and 0), in that order, by σ. We note that

$$w^a x^b y^c z^d + w^d x^c y^b z^a = \begin{cases} (wz)^{\min(a,d)}(xy)^{\min(b,c)} \left(w^{|a-d|} y^{|b-c|} + x^{|b-c|} z^{|a-d|}\right) & a > d \text{ XOR } b > c \\ (wz)^{\min(a,d)}(xy)^{\min(b,c)} \left(w^{|a-d|} x^{|b-c|} + y^{|b-c|} z^{|a-d|}\right) & \text{otherwise.} \end{cases}$$

We prove $a \equiv d \pmod 3$ inductively. For the all-0 assignment, this is clear since every edge contributes a factor w and the number of edges is divisible by 3 for a 3-regular graph. Now starting from any assignment, if we switch the assignment on one vertex from 0 to 1, it is easy to verify that it changes the valuation from $w^a x^b y^c z^d$ to $w^{a'} x^{b'} y^{c'} z^{d'}$, where $a - d = a' - d' + 3$. As every $\{0, 1\}$ assignment is obtainable from the all-0 assignment by a sequence of switches, the conclusion $a \equiv d \pmod 3$ follows.

Now

$$w^a x^b y^c z^d + w^d x^c y^b z^a = \begin{cases} (wz)^{\min(a,d)}(xy)^{\min(b,c)} \left(w^{3k} y^{\ell} + x^{\ell} z^{3k}\right) & a > d \text{ XOR } b > c \\ (wz)^{\min(a,d)}(xy)^{\min(b,c)} \left(w^{3k} x^{\ell} + y^{\ell} z^{3k}\right) & \text{otherwise} \end{cases}$$

for some $k, \ell \geq 0$. Consider $w^{3k} y^{\ell} + x^{\ell} z^{3k}$ (the other case is similar). Two simple inductive steps

$$\begin{aligned} w^{3k} y^{\ell+1} + x^{\ell+1} z^{3k} &= \left(w^{3k} y^{\ell} + x^{\ell} z^{3k}\right)(x + y) - xy \left(w^{3k} y^{\ell-1} + x^{\ell-1} z^{3k}\right) \\ w^{3(k+1)} y^{\ell} + x^{\ell} z^{3(k+1)} &= \left(w^{3k} y^{\ell} + x^{\ell} z^{3k}\right)(w^3 + z^3) - (wz)^3 \left(w^{3(k-1)} y^{\ell} + x^{\ell} z^{3(k-1)}\right) \end{aligned}$$

(when combined with the other case) show that the Holant is a polynomial $P(wz, xy, w^3 + z^3, x + y, w^3 x + y z^3, w^3 y + x z^3)$ with integer coefficients. □

Assuming non-degeneracy of (w, x, y, z), Lemma 5.5.1, Lemma 5.5.2, and Lemma 5.5.3 give #P-hardness unless two (or more) of w, x, y, and z are 0 or none are 0 and $|x| = |y|$. If any two (or more) of variables are 0, then the problem is tractable, as proved after Theorem 5.4.1. Therefore, the dichotomy in Theorem 5.4.1 holds unless $wxyz \neq 0$ and $|x| = |y|$. In accordance with Lemma 5.5.7, we make a change of variables to $A = wz$, $B = xy$, $C = w^3 + z^3$, $D = x + y$, $E = w^3x + yz^3$, and $F = w^3y + xz^3$. Since the complexity of a Holant remains the same under multiplication by a nonzero constant to any signature, we normalize so that $|x| = 1$ and $x = \overline{y}$ without repeatedly stating this as an assumption. Thus $B = 1$ and $D = x + y \in [-2, 2]$ with $D^2 = 4$ corresponding to the symmetric case: $x = y$. A degenerate edge signature now means $A = 1$. Additionally, notice that $E + F = CD$ and $EF = -4A^3B + BC^2 + A^3D^2$. Theorem 5.4.1 can also be stated in these symmetrized variables.

Theorem 5.5.8. *Suppose* $w, x, y, z \in \mathbb{C}$. *Then* $\operatorname{Holant}((w, x, y, z) \mid =_3)$ *is* #P*-hard except in the following cases, for which the problem is computable in polynomial time:*

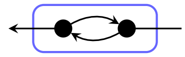
1. $wz = xy \iff A = B$;
2. $w = z = 0 \iff A = C = 0$;
3. $x = y = 0 \iff B = D = 0$;
4. $wz = -xy \wedge w^6 = z^6 \wedge x^2 = y^2 \iff A = -B \wedge 4A^3C = C^3 \wedge 4BD = D^3$;
5. $wz = -xy \wedge w^6 = -z^6 \wedge x^2 = -y^2 \iff A = -B \wedge 2A^3 = C^2 \wedge 2B = D^2$.

If the input is restricted to planar graphs, then two more cases become tractable but all other cases remain #P*-hard:*

6. $w^3 = z^3 \wedge x = y \iff 4A^3 = C^2 \wedge 4B = D^2$;
7. $w^3 = -z^3 \wedge x = -y \iff C = D = 0$.

Remark. One way to convert a polynomial in w, x, y, z to a polynomial in A, B, C, D, E, F is to use a Gröbner basis. We tried that here, but it did not give us the kinds of polynomials that we wanted. Intuitively, we want polynomials that look as though they were obtained by applying the proof of Lemma 5.5.7. Not only does the result of the Gröbner basis algorithm depend on the order in which the variables are considered, but we have an underdetermined system because we are converting from only four variables to six. In the end, I wrote my own program to do the conversion, which tried to minimize the occurrences of E and F as much as possible.

Now we continue with the proof of #P-hardness.



Gadget 24

Figure 5.9: Another unary recursive gadget

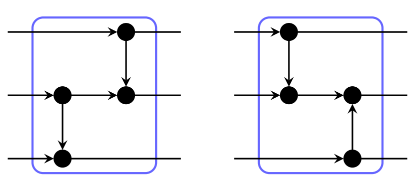
Lemma 5.5.9. *If $D^2 \neq 4$, and $A \notin \mathbb{R}$, then* $\operatorname{Holant}((w,x,y,z) \mid {=_3})$ *is* #P*-hard.*

Proof. The transition matrices for Gadget 24 and Gadget 3 are $M_{24} = \left[\begin{smallmatrix} w & y \\ x & z \end{smallmatrix}\right] \left[\begin{smallmatrix} w^2 & xy \\ xy & z^2 \end{smallmatrix}\right]$ and $M_3 = \left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right] \left[\begin{smallmatrix} w^2 & xy \\ xy & z^2 \end{smallmatrix}\right]$. Both matrices have determinant $(A-1)^2(A+1)$, which is nonzero since A is not real. Then $N = M_{24} M_3^{-1}$ has determinant 1 and trace

$$\operatorname{tr}\left(\begin{bmatrix} w & y \\ x & z \end{bmatrix} \begin{bmatrix} w & x \\ y & z \end{bmatrix}^{-1} \right) = \frac{2wz - x^2 - y^2}{wz - xy} = \frac{2A - D^2 + 2}{A - 1},$$

which is nonzero since A is not real. If the eigenvalues of N have distinct norms, then it has infinite order up to a scalar and we are done by Theorem 5.3.11, so assume that its eigenvalues are of equal norm. Then Lemma 5.5.4 says that $\frac{\operatorname{tr}(N)^2}{\det N} = \frac{(2A-D^2+2)^2}{(A-1)^2} \in \mathbb{R}^+$. Taking square roots, we have $\frac{2A-D^2+2}{A-1} \in \mathbb{R}$, which implies that $\frac{-D^2+4}{A-1} \in \mathbb{R}$. Since $D^2 \neq 4$, this gives $A \in \mathbb{R}$, a contradiction. □

Unary recursive gadgets, such as the ones used in the proof of Lemma 5.5.9, are quite useful for proving #P-hardness when variables like $A = wz$ are complex. When all variables are real, the conclusion of Lemma 5.5.4 is weak (though one can still prove #P-hardness using a related lemma with significant effort in the symmetric case [48]). For complex variables in the symmetric case, [93] showed that using higher arity (namely binary) recursive gadgets can give a much simpler proof of #P-hardness. The next lemma continues this pattern with the first ever use of ternary recursive gadgets.



Gadget 25 Gadget 26

Figure 5.10: Ternary recursive gadgets

Lemma 5.5.10. *If* $A^2 \neq 1$, $AD \neq 0$, *and* $D^2 \neq 4$, *then* $\operatorname{Holant}((w,x,y,z) \mid =_3)$ *is* #P*-hard.*

Proof. The determinants of the 8-by-8 transition matrices of Gadget 25 and Gadget 26 are both $A^2(A-1)^4 \neq 0$. If $N = M_{25}^{-1} M_{26}$ has any two eigenvalues with distinct norms, then it has infinite order up to a scalar and we are done by Theorem 5.3.11. Thus assume that all eight eigenvalues of N have the same norm. Then by Lemma 5.5.6, we know that several equations hold among the coefficients of its characteristic polynomial. After scaling by the nonzero factor $A(A-1)$, these coefficients for $A(A-1)N$ are

$$
\begin{aligned}
a_7 &= (A-1)(AD^2 + 2A + 2)\\
a_6 &= (A-1)^2(5A^2D^2 - 3A^2 + 2AD^2 + 2A + 1)\\
a_5 &= A(A-1)^3(A^2D^4 + 5A^2D^2 - 6A^2 + 7AD^2 - 6A + D^2)\\
a_4 &= A^2(A-1)^4(3A^2D^4 - 4A^2D^2 + 4A^2 + AD^4 + 4AD^2 - 4A + 2D^2 - 2)\\
a_3 &= A^3(A-1)^5(2AD^4 + 3A^2D^4 - 6A^2D^2 + 6A^2 - 4AD^2 + 6A + D^2)\\
a_2 &= A^4(A-1)^6(A^2D^4 + A^2D^2 - 3A^2 + AD^4 - 2AD^2 + 2A + 1)\\
a_1 &= A^6(A-1)^7(2AD^2 - 2A + D^2 - 2)\\
a_0 &= A^8(A-1)^8.
\end{aligned}
$$

Amazingly, C, E, and F do not appear.[2] Lemma 5.5.9 shows #P-hardness unless $A \in \mathbb{R}$, so assume that $A \in \mathbb{R}$. Because $A, D \in \mathbb{R}$, the equations in Lemma 5.5.6 are simplified by the disappearance of norms and conjugates. Using CylindricalDecomposition in Mathematica, we conclude that there are no solutions under our assumptions, which is a contradiction. ☐

Remark. Lemma 5.5.6 gives three equations that follow from the roots all having the same norm. We had found five equations, but these three were sufficient to prove Lemma 5.5.10. Even with the two additional equations, we did not see a way to prove the lemma without using Mathematica.

We explain the meaning of the assumptions in Lemma 5.5.10 after the next lemma, which considers the same assumptions except that $D = 0$ and $C \neq 0$.

[2] The runtime of CylindricalDecomposition is a double exponential in the number of variables, so it is crucial that our query include as few variables as possible.

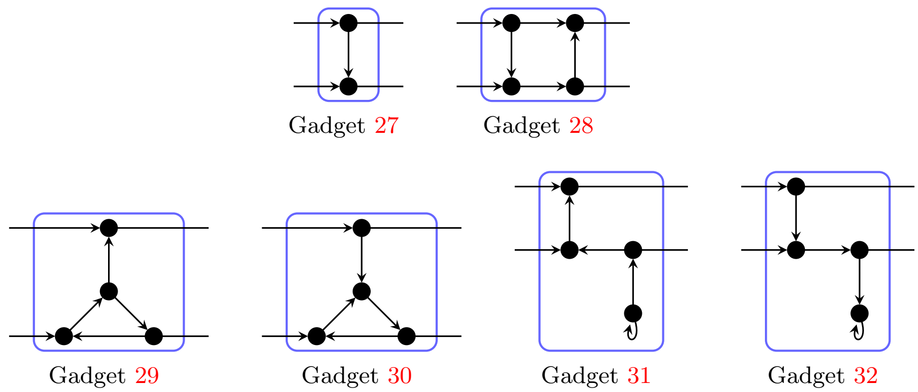


Figure 5.11: Binary recursive gadgets

Lemma 5.5.11. *If $A^2 \neq 1$, $AC \neq 0$, and $D = 0$, then* $\operatorname{Holant}((w, x, y, z) \mid =_3)$ *is* #P*-hard.*

Proof. Lemma 5.5.9 shows #P-hardness unless $A \in \mathbb{R}$, so assume that $A \in \mathbb{R}$. The transition matrix for Gadget 27 is $M_{27} = \left[\begin{smallmatrix} w & x \\ y & z \end{smallmatrix}\right]^{\otimes 2} \operatorname{diag}(w, x, y, z)$ and has determinant $A(A-1)^2 \neq 0$. If M_{27} has any two eigenvalues with distinct norms, then it has infinite order up to a scalar and we are done by Theorem 5.3.11, so assume that all eigenvalues have the same norm. However, the coefficients of the characteristic polynomial of M_{27}, which are

$$(a_3, a_2, a_1, a_0) = \left(-C, (A+1)^2(A-1), -(A-1)^2 C, A(A-1)^4\right),$$

do not satisfy the conclusion of Lemma 5.5.5 under the assumptions, a contradiction. □

The case $A = 1$ is degenerate (thus tractable), the case $A = 0$ is covered in Lemma 5.5.3, and recall that $D^2 = 4$ corresponds to the symmetric case [93], so now we assume $A \notin \{0, 1\} \wedge D^2 \neq 4$. Lemma 5.5.10 handled $A \neq -1$ and $D \neq 0$ while Lemma 5.5.11 handled $A \neq -1 \wedge D = 0 \wedge C \neq 0$. We note that $C = D = 0$ is tractable over planar graphs. Now we focus on the case $A = -1$.

The next two proofs of #P-hardness (the proofs of Lemma 5.5.13 and Lemma 5.5.14) make use of the following technical lemma.

Lemma 5.5.12. *Let $c \in \mathbb{C}$ and $\varepsilon = \pm 1$. Then the only solutions to the equation $(\overline{c + 2\varepsilon})c = \varepsilon(c + 2\varepsilon)$ are the trivial solutions $c \in \{-2\varepsilon, \varepsilon\}$.*

Proof. Assume that $c \neq -2\varepsilon$. Now we show that $c = \varepsilon$. Taking norms, we see that $|c| = 1$. Then simplifying $(\overline{c + 2\varepsilon})c = \varepsilon(c + 2\varepsilon)$ using $c\overline{c} = |c|^2 = 1$ yields $c = \varepsilon$ as claimed. □

Lemma 5.5.13. *If $A = -1$ and $E \notin \{0, \pm 2i\}$, then* $\operatorname{Holant}((w, x, y, z) \mid =_3)$ *is* #P*-hard.*

Proof. The transition matrices of Gadget 28, Gadget 29, and Gadget 31 are

$$M_{28} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}^{\otimes 2} \operatorname{diag}(w, y, x, z) \begin{bmatrix} w & y \\ x & z \end{bmatrix}^{\otimes 2} \operatorname{diag}(w, x, y, z)$$

$$M_{29} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}^{\otimes 2} \begin{bmatrix} w^4 + wxy^2 & w^2xy + xy^2z & 0 & 0 \\ w^2xy + xy^2z & wxyz + yz^3 & 0 & 0 \\ 0 & 0 & w^3x + wxyz & wx^2y + xyz^2 \\ 0 & 0 & wx^2y + xyz^2 & x^2yz + z^4 \end{bmatrix}$$

$$M_{31} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}^{\otimes 2} \operatorname{diag}(w, y, x, z) \left(I_2 \otimes \begin{bmatrix} w & x \\ y & z \end{bmatrix} \begin{bmatrix} w^2 + yz & 0 \\ 0 & wx + z^2 \end{bmatrix} \right)$$

with $\det M_{28} = 2^8$, $\det M_{31} = -2^6 E^2$, and $\det M_{29} = 2^6(E^2 + 4)$, so all are nonsingular. Let $N_1 = M_{28}^{-1} M_{31}$ and $N_2 = M_{28}^{-1} M_{29}$. The coefficients of the characteristic polynomials of $-2^4 N_1$ and $2^4 N_2$ are respectively

$$(a_3, a_2, a_1, a_0) = \left(-4, E^2 + 12, -4(E^2 + 4), 4(E^2 + 4)\right)$$
$$(a_3, a_2, a_1, a_0) = \left(4, -E^2 + 8, -4E^2, -4E^2\right).$$

If N_1 (resp. N_2) has any two eigenvalues with distinct norms, then N_1 (resp. N_2) has infinite order up to a scalar and we are done by Theorem 5.3.11, so assume that all eigenvalues of N_1 (resp. N_2) have the same norm. Then by Lemma 5.5.5, we have two equations relating these coefficients. However, after a change of variables by $c = (E^2 + 4)/4$ (for the coefficients of N_1) and $c = E^2/4$ (for the coefficients of N_2), Lemma 5.5.12 says that the only solutions to both equations require $E \in \{0, \pm 2i\}$, a contradiction. □

The next lemma is similar to Lemma 5.5.13 with E in place of F.

Lemma 5.5.14. *If $A = -1$ and $F \notin \{0, \pm 2i\}$, then* $\operatorname{Holant}((w, x, y, z) \mid =_3)$ *is* #P*-hard.*

Proof. The transition matrices of Gadget 28, Gadget 30, and Gadget 32 are

$$M_{28} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}^{\otimes 2} \operatorname{diag}(w, y, x, z) \begin{bmatrix} w & x \\ y & z \end{bmatrix}^{\otimes 2} \operatorname{diag}(w, x, y, z)$$

$$M_{30} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}^{\otimes 2} \begin{bmatrix} w^4 + wx^2y & w^2xy + x^2yz & 0 & 0 \\ w^2xy + x^2yz & wxyz + xz^3 & 0 & 0 \\ 0 & 0 & w^3y + wxyz & wxy^2 + xyz^2 \\ 0 & 0 & wxy^2 + xyz^2 & xy^2z + z^4 \end{bmatrix}$$

$$M_{32} = \begin{bmatrix} w & x \\ y & z \end{bmatrix}^{\otimes 2} \operatorname{diag}(w, x, y, z) \left(I_2 \otimes \begin{bmatrix} w & x \\ y & z \end{bmatrix} \begin{bmatrix} w^2 + xz & 0 \\ 0 & wy + z^2 \end{bmatrix} \right).$$

The rest of the proof uses the same reasoning as the proof of Lemma 5.5.13 with Gadget 29 and Gadget 31 replaced by Gadget 30 and Gadget 32 respectively. □

All remaining cases, those for which $A = -1$ and $E, F \in \{0, \pm 2i\}$, imply tractability. Since this is not immediately obvious, we prove this next. As pointed out after Lemma 5.5.7, the following equations hold and are used frequently below. They simplify to

$$E + F = CD \tag{5.5.1}$$

$$EF = -4A^3B + BC^2 + A^3D^2 = 4 + C^2 - D^2 \tag{5.5.2}$$

when $A = -1$ and $B = 1$. These next four lemmas cover all possibilities of $E, F \in \{0, \pm 2i\}$ as follows:

Lemma 5.5.15: Both zero;

Lemma 5.5.16: Both nonzero and equal;

Lemma 5.5.17: Both nonzero and not equal;

Lemma 5.5.18: Exactly one zero.

Lemma 5.5.15. *If* $A = -1 \wedge E = F = 0$*, then* $(D = 0 \wedge C^2 = -4)$ *or* $(D^2 = 4 \wedge C = 0)$*, which are both tractable.*

Proof. Since $0 = E + F = CD$, either C or D is zero. In either case, simplifying (5.5.2) gives the desired result and is covered by tractable case 4 in Theorem 5.5.8. □

Lemma 5.5.16. *If* $A = -1 \wedge E = F = \pm 2i$*, then* $C^2 = -4 \wedge D^2 = 4$*, which is tractable.*

Proof. Using $wz = A = -1$ and $xy = B = 1$, we multiply $\pm 2i = E = w^3 x + yz^3$ by $w^3 y$ to get $y^2 \pm 2iw^3 y - w^6 = 0$. Similarly, multiplying $\pm 2i = F = w^3 y + xz^3$ by $w^3 x$ gives $x^2 \pm 2iw^3 x - w^6 = 0$. This is the same quadratic polynomial with x and y as indeterminates. Its discriminant is zero, so $x = y$ which means that $D^2 = 4$. Simplifying (5.5.2) yields $C^2 = -4$ as required. This is covered by tractable case 4 in Theorem 5.5.8. □

Lemma 5.5.17. *If* $A = -1 \wedge E = -F = \pm 2i$*, then* $C = D = 0$*, which is tractable.*

Proof. Since $0 = E + F = CD$, either C or D is zero. Simplifying (5.5.2) gives $C^2 = D^2$, so both C and D are zero. This is covered by tractable case 4 in Theorem 5.5.8. □

Lemma 5.5.18. *If* $A = -1 \wedge ((E = \pm 2i \wedge F = 0) \vee (E = 0 \wedge F = \pm 2i))$*, then* $D^2 = 2 \wedge C^2 = -2$*, which is tractable.*

Proof. Since $\pm 2i = E + F = CD$, neither C or D is zero. Squaring this equation and solving for C^2 gives $C^2 = -4/D^2$. In (5.5.2), first we substitute for C^2 to conclude that $D^2 = 2$ and then substitute for D^2 to conclude that $C^2 = -2$. This is tractable case 5 in Theorem 5.5.8. □

At this point, every setting of the variables has either been proven tractable over planar graphs or #P-hard. So far, all our hardness proofs originate from #VertexCover over 3-regular graphs, which is $\operatorname{Holant}((0,1,1,1) \mid =_3)$. Recall that #VertexCover is #P-hard even for 3-regular planar graphs [144] and notice that all of our gadget constructions are planar, including our interpolation construction in the Group Lemma (see Figure 5.2c). Therefore, all of the #P-hardness results proved so far still apply when the input is restricted to planar graphs. There are, however, some cases where the problem is #P-hard in general, yet is polynomial time computable when restricted to planar graphs. We analyze this case next using a lemma from [93] that can also be found in [92].

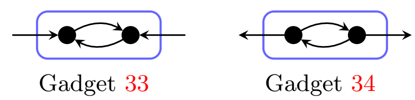


Figure 5.12: Gadgets on the left with a symmetric signature.

Lemma 5.5.19 (Lemma 33 of [92])**.** *If* $w \notin \{0, \pm 1, \pm i\}$*, then* $\operatorname{Holant}((w,1,1,w) \mid {=_3})$ *is* #P*-hard.*

Lemma 5.5.20. *If* $w \notin \{0, \pm 1, \pm i\}$*, then* $\operatorname{Holant}((w,1,-1,-w) \mid {=_3})$ *is* #P*-hard.*

Proof. Gadget 33 and Gadget 34 simulate two symmetric signatures on the left. Using Reduce in Mathematica, we conclude that at least one of the gadgets satisfies the hypothesis for #P-hardness from Lemma 5.5.19. □

Lemma 5.5.21. *Suppose* $w^3 = \varepsilon z^3 \wedge x = \varepsilon y$ *where* $\varepsilon = \pm 1$*. If* $x \neq 0 \wedge w/x \notin \{0, \pm 1, \pm i\}$*, then* $\operatorname{Holant}((w,x,y,z) \mid {=_3})$ *is* #P*-hard.*

Proof. If $wz \neq 0$, then we apply the holographic transformation $\left[\begin{smallmatrix} \alpha & 0 \\ 0 & \alpha^2 \end{smallmatrix}\right]$ with $\alpha = \varepsilon z/w$, which does not change the Holant value by Lemma 3.2.1. As in the proof of tractability, this reduces to the case $w = \varepsilon z \wedge x = \varepsilon y$. This equivalence still holds if $wz = 0$. We then normalize $x = 1$ (since it is nonzero) and replace z with εw to obtain the edge signature $(w/x, 1, \varepsilon, \varepsilon w/x)$. Depending on ε, this case is either covered in Lemma 5.5.19 (when $\varepsilon = 1$) or in Lemma 5.5.20 (when $\varepsilon = -1$), so we are done. □

5.6 Anti-Gadgets and Previous Work

To further appreciate the usefulness of anti-gadgets, we show how this technique sheds new light on previous results.

One can find *failure conditions* for a binary recursive gadget using the following lemma.

Lemma 5.6.1. *Let G be a binary recursive gadget having nonsingular transition matrix M. Then* $\{M^i\}_{i \geq 0}$ *is a sequence of pairwise linearly independent signatures unless* $a_2|a_1|^2 - |a_3|^2\overline{a_2}a_0 = 0$*, where* $x^4 + a_3x^3 + a_2x^2 + a_1x + a_0$ *is the characteristic polynomial of M.*

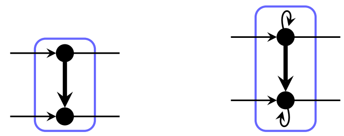


Figure 5.13: Recursive gadgets from [37] on k-regular graphs. Bold edges represent parallel edges. In Gadget 35 (resp. Gadget 36), the multiplicity is $k-2$ (resp. $k-4$) so that the vertices have degree k.

Analyzing a failure condition such as $a_2|a_1|^2 - |a_3|^2\overline{a_2}a_0 = 0$ simultaneously for several gadgets is quite difficult, even with the aid of symbolic computation. Previous work [37, 36] relied heavily on miraculous cancellations in the failure conditions to contend with this. For example, consider the two gadgets in Figure 5.13. They are from [37], where symmetric (i.e. $x = y$) signatures (w, x, y, z) were considered on k-regular graphs.

After a change of variables $X = wzx^{-2}$ and $Y = (w/x)^3 + (z/x)^3$ and making a few assumptions to guarantee that M_{35} and M_{36} are nonsingular (which we omit in this discussion), the failure conditions of Gadget 35 and Gadget 36 (when restricted to the real numbers) simplify to

$$(X-1)^3(X^{k-2}-1)(X^{k-2}(X+1)^2(X^{k-1}+X^{k-2}+X+3Y+1)-Y^3) = 0,$$
$$X^3(X-1)^3(X^{k-4}-1)(X^{k-2}(X+1)^2(X^2+X^{k-2}+X+3Y+X^{k-3})-Y^3) = 0.$$

Assuming that both gadgets fail and $X \notin \{0, \pm 1\}$, this gives two polynomial expressions for Y^3. Setting these equal to each other and refactoring results in the contradiction $X^{k-2}(X+1)^3(X-1)(X^{k-3}-1) = 0$, implying that either one or the other gadget works. At the time of this discovery, it was a mystery whether there was any underlying explanation for such miraculous cancellations. Now we see how anti-gadgets reveal a better understanding of this same gadget pair.

By assuming that M_{35} *fails* to produce an infinite set of pairwise linearly independent signatures, we have an explicit recursive gadget for M_{35}^{-1}. Then $M_{35}^{-1}M_{36} = \operatorname{diag}(1, X, X, 1)$ clearly produces an infinite set of pairwise linearly independent signatures unless X is zero or a root of unity. Note that in the "gadget language" of $M_{35}^{-1}M_{36}$, the two leading directed edges of Gadget 35 and Gadget 36 simply annihilate each other, as do $k-4$ copies of the vertical edge. The signatures $=_3$ at the

Figure 5.14: Recursive gadgets from [36] on k-regular graphs for k even. The gadgets are pictured for $k = 4$ but generalize to all even $k \geq 4$ by adding self loops to the vertices.

degree 3 vertices force the matrix $M_{35}^{-1} M_{36}$ to be diagonal. Thus, with almost no effort we have a strictly stronger result (i.e. over the complex numbers) through the use of an anti-gadget. This also shows that the anti-gadget concept is useful in the symmetric setting as well as the asymmetric setting.

In [36], a similarly fantastic cancellation occurred involving Gadget 37 and Gadget 38 (see Figure 5.14). They form a suitable gadget and anti-gadget pair, as $M_{37}^{-1} M_{38}$ is a diagonal matrix. While this diagonal matrix is not as easy to analyze as the previous example, anti-gadgets would *inform* the search for such useful gadgets, even if the analysis is carried out with different techniques.

5.7 Closing Thoughts

General purpose binary starter gadgets In Section 5.3, we said

> [projective gadget sets permit] much more freedom in gadget constructions, and this power is used crucially in the proof of our dichotomy theorem.

We definitely believed this at the time, but now I think there is a sense in which this is a bit over stated. Let me explain.

When I begin this work of generalizing [94, 93], I decided not to use the linear interpolation that they used. They were unable to find general purpose binary starter gadgets, which made their proof more difficult. In his thesis [92, p. 69], Michael Kowalczyk suggested this difficulty could be avoided by using the circular interpolation construction from their later work [37, 36]. They had used it to overcome a parity restriction when the arity of Equality signature is even (in which case, one can only construct $((w, x, y, z) \mid =_k)$-gates with an even number of dangling edges). I took this suggestion.

Using the circular construction when doing k-dimensional interpolation brings its own challenge. Namely, how to construct projective gadgets from arity k-to-1? We decisively solved this problem by Lemma 5.3.3, which says that it suffices to find projective gadgets from arity 2-to-1. This is what the previous [94, 93] work had called finisher gadgets, so given this lemma, we now had strictly fewer conditions to satisfy to prove our dichotomy than [94, 93]. However, it is better than this.

After publishing [39], we realized that we could have simplified our proof. Lemma 5.3.3 can be strengthened to say that it suffices to find projective gadgets from arity *1-to-0*. (These projective gadgets only have one dangling edge, so this is only possible when the arity of the Equality signature is odd.) Then we can get drastically simpler versions of Lemma 5.3.7 and Lemma 5.3.8. Instead of needing all the projective gadgets from arity 2-to-1 in Figure 5.6 and the projective gadget sets they form in Table 5.1, we should be able to find projective gadget set from arity *1-to-0*.

Projective gadgets from arity 1-to-0 are the same as unary starter gadgets in [94, 93], and they had no problem finding general purpose unary starter gadgets, so I expect that we could also find general purpose unary starter gadgets in our setting. Now look more closely at the proof of Lemma 5.3.3. The projective gadget F from arity 2-to-1 was used to create the projective gadget $F' = I_{2^{k-2}} \otimes F$ from arity k-to-1. This was of creating F' makes it easy to reason about, but any way of tensoring F and identity matrices (and ending up with a 2^{k-1}-by-2^k) should work. In particular, we could use $F' = F \otimes I_{2^{k-1}}$ in the strengthening of Lemma 5.3.3 when F is a projective gadget from arity 1-to-0.

This construction actually yields general purpose binary starter gadgets from general purpose unary starter gadgets, which solve the open question Kowalczyk has posed in his thesis. The final construction is not really the circular interpolation construction nor is it really the linear interpolation construction. Instead, it is some hybrid of the two and superior to both.

Conjecture for generalization Based on the result in this chapter and the results in [37, 36], I make the following conjecture.

Conjecture 5.7.1. *Let $w, x, y, z \in \mathbb{C}$ and $f = (w, x, y, z)$. Suppose $k \geq 1$ is an integer. Then* $\operatorname{Holant}(f \mid =_k)$ *is #*P*-hard unless one of the following conditions holds, in which case, the problem is computable in polynomial time:*

1. $k \leq 2$;
2. degenerate: $wz = xy$;
3. generalized disequality: $w = z = 0$;
4. generalized equality: $x = y = 0$;
5. holographic reduction to affine: $wz = -xy \wedge w^{2k} = \varepsilon^k z^{2k} \wedge x^2 = \varepsilon y^2$, where $\varepsilon = \pm 1$.

*If the input is restricted to planar graphs, then another case becomes computable in polynomial time but everything else remains #*P*-hard:*

6. holographic reduction to matchgates: $w^k = \varepsilon^k z^k \wedge x = \varepsilon y$, where $\varepsilon = \pm 1$.

I can show that the conjectured tractable cases are indeed so.

Proof of tractability. For case 1, the Holant is a product over disconnected components, and on each connected component, the Holant can be computed with matrix product and trace. In Case 1, case 2, and case 3, both signatures are of product type, so we are done by Corollary 4.1.5.

For case 5, we do a holographic transformation by $\left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right]$ with $\alpha = (\varepsilon^k w^2 / z^2)^{1/4}$. Reusing variable names, the transformed edge signature $f = (w, x, y, z)$ satisfies $wz = -xy \wedge w^2 = \varepsilon^k z^2 \wedge x^2 = \varepsilon^k y^2$, which is affine. The EQUALITY signature $=_k$ is transformed to $[1, 0, \ldots, 0, \alpha^k]$, which is affine since $\alpha^{4k} = 1$. Thus we are done by Theorem 4.2.6.

For case 6, we do a holographic transformation by $\left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right]$ with $\alpha = (\varepsilon^k w / z)^{1/2}$. Reusing variable names, the transformed edge signature $f = (w, x, y, z)$ satisfies $w = \varepsilon^k z \wedge x = \varepsilon^k y$. The EQUALITY signature $=_k$ is transformed to $[1, 0, \ldots, 0, \alpha^k]$. Since $\alpha^{2k} = 1$, both of these signatures become matchgate signatures after a further holographic transformation by the Hadamard matrix $\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$, so we are done by Theorem 4.3.2. $\square$

A proof of hardness for the remaining cases should follow without too much difficulty by using the Pl-#CSP$^2(\mathcal{F})$ dichotomy recently proved in [27].

Chapter 6

Dichotomy for Holant Problems over Planar 4-Regular Graphs

For Holant problems, it is important to understand the complexity of the small arity cases first. In this chapter, we prove a dichotomy theorem for a symmetric signature of arity 4 with complex weights in the planar Holant framework. This result is used to obtain both the Holant dichotomy in Chapter 7 and the planar #CSP dichotomy in Chapter 8. To prove this result, we improve known interpolation techniques in both the one- and two-dimensional settings. In one dimension, we provide a tight characterization of when it succeeds. We apply this result alongside the anti-gadget technique from Chapter 5 and a new technique called planar pairing. The case for two dimensions involves a demanding interpolation step using asymmetric signatures. We found that in order to prove a dichotomy for a symmetric signature, we must consider asymmetric ones. In particular, we prove that counting Eulerian orientations over planar 4-regular graphs is #P-hard. The reduction is from the problem of evaluating the Tutte polynomial of a planar graph at the point $(3, 3)$. Part of this work was published in [29, 30] and part of it was published in [73, 74].

6.1 Background

For Holant problems, it is important to understand the complexity of the small arity cases first. We use the following theorem about edge-weighted signatures on degree prescribed graphs. Specifically, we apply it with $\mathcal{G} = \{=_4\}$. See also Theorem 22 in [92], which contains a proof.

Theorem 6.1.1 (Theorem 3 in [36]). *Let $S \subseteq \mathbb{Z}^+$ contain $k \geq 3$, let $\mathcal{G} = \{=_k \mid k \in S\}$, and let $d = \gcd(S)$. Suppose $f_0, f_1, f_2 \in \mathbb{C}$. Then $\operatorname{Holant}([f_0, f_1, f_2] \mid \mathcal{G})$ is #P-hard unless one of the following conditions hold, in which case the problem is computable in polynomial time:*

1. $f_0 f_2 = f_1^2$;
2. $f_0 = f_2 = 0$;
3. $f_1 = 0$;
4. $f_0 f_2 = -f_1^2 \wedge f_0^{2d} = f_2^{2d}$.

If the input is restricted to planar graphs, then the problem remains #P-hard unless

5. $f_0^d = f_2^d$,

in which case the problem is computable in polynomial time.

Theorem 6.1.1 is very explicit, but its restatement as Theorem 6.1.2 is more conceptual. Let $\mathcal{T}_k = \left\{\left[\begin{smallmatrix} 1 & 0 \\ 0 & \omega \end{smallmatrix}\right] \in \mathbb{C}^{2\times 2} \mid \omega^k = 1\right\}$.

Theorem 6.1.2 (Theorem 3 in [36]). *Let $S \subseteq \mathbb{Z}^+$ contain $k \geq 3$, let $\mathcal{G} = \{=_k \mid k \in S\}$, and let $d = \gcd(S)$. Suppose f is a non-degenerate, symmetric, complex-valued binary signature in Boolean variables. Then $\operatorname{Holant}(f \mid \mathcal{G})$ is #P-hard unless one of the following conditions hold, in which case the problem is computable in polynomial time:*

1. $f \in \mathscr{P}$;
2. *there exists* $T \in \mathcal{T}_{4d}$ *such that* $T^{\otimes 2} f \in \mathscr{A}$;

If the input is restricted to planar graphs, then the problem remains #P-hard unless

3. *there exists* $T \in \mathcal{T}_{2d}$ *such that* $T^{\otimes 2} f \in \left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right] \mathscr{M}$,

in which case the problem is computable in polynomial time.

When f is a symmetric signature of arity 3 with complex weights, Cai, Huang, and Lu [35, Theorem 3] gave a dichotomy for $\operatorname{Holant}(f)$ while Cai, Lu, and Xia [46, special case of Theorem V.1] gave a dichotomy for Pl-Holant(f). We state their dichotomies as a single result.

Theorem 6.1.3. *If f is a non-degenerate, symmetric, complex-valued signature of arity 3 in Boolean variables, then* Holant(f) *is* #P*-hard unless f is $\mathscr{A}$-transformable or $\mathscr{P}$-transformable or vanishing, in which case, the problem is computable in polynomial time. If the input is restricted to planar graphs, then the problem remains* #P*-hard unless $\mathscr{M}$-transformable, in which case, the problem is computable in polynomial time.*

In this chapter, we use the dichotomy for Pl-Holant(f) prove the planar dichotomy for any signature of arity 4 with complex weights. This result is used to obtain both the Holant dichotomy in Chapter 7 and the planar #CSP dichotomy in Chapter 8. To prove this result, we improve known interpolation techniques in both the one- and two-dimensional settings. Some previous dichotomy theorems were for Holant problems with real weights [48, 37, 46, 80]. A dichotomy for complex weights is more technically challenging. The proof technique of polynomial interpolation often has infinitely many failure cases in $\mathbb{C}$ corresponding to the infinitely many roots of unity, which prevents a brute force analysis of failure cases. This increased difficulty requires us to develop new ideas to bypass previous interpolation proofs.

For interpolation in one dimension, we provide a tight characterization of when it succeeds. When it fails is the perfect time to use the anti-gadget technique from Chapter 5. We also introduce the notion of planar pairings to build reductions. We show that every planar 3-regular graph has a planar pairing and that one can be efficiently computed. By combining these three techniques, we show that counting complex-weighted matchings over planar 4-regular graphs is #P-hard.

Our interpolation result in two dimensions involves a demanding interpolation step using asymmetric signatures. We found that in order to prove a dichotomy for a symmetric signature, we must consider asymmetric ones. Huang and Lu [80] proved that counting Eulerian orientations is #P-hard over 4-regular graphs but left open its complexity when the input is also planar. We perform a planar interpolation with a rotationally invariant signature to show that this problems remains #P-hard when further restricted to planar graphs. The reduction is from the problem of evaluating the Tutte polynomial of a planar graph at the point $(3, 3)$, which has a natural expression in the Holant framework.

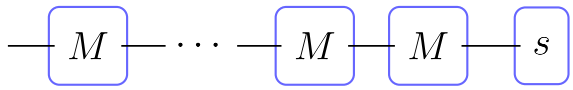


Figure 6.1: Unary recursive construction with starter gadget.

Theorem 6.1.4 (Theorem 5.1 in [138]). *For $x, y \in \mathbb{C}$, evaluating the Tutte polynomial at (x, y) is* #P*-hard over planar graphs unless*

$$(x-1)(y-1) \in \{1, 2\} \qquad \textit{or} \qquad (x, y) \in \{(1, 1), (-1, -1), (\omega, \omega^2), (\omega^2, \omega)\},$$

where $\omega = e^{2\pi i/3}$. In each exceptional case, the computation can be done in polynomial time.

6.2 Improving Unary Interpolation

In this section, we discuss a common interpolation method called the *unary recursive construction* and characterize of when it succeeds. The goal of this construction is to interpolate a unary signature and is based on work by Vadhan [124] and further developed by others [50, 47, 36].

There are two gadgets in the unary recursive construction: a *starter* gadget of arity 1 and a *recursive* gadget of arity 2. The signature of the starter gadget is represented by its signature matrix with parameter 1, which is a two-dimensional column vectors s. The signature of the recursive gadget is also represented by its signature matrix with parameter 1, which is a 2-by-2 matrix M. The construction begins with the starter gadget and proceeds by connecting $k \geq 0$ recursive gadgets, one at a time, to the only available edge (see Figure 6.1). The signature matrix with parameter 1 of the resulting gadget $M^k s$. This construction is denoted by (M, s).

The essential difficulty in using polynomial interpolation is constructing an infinite set of signatures that are pairwise linearly independent [36]. The pairwise linear independence of signatures translates into distinct evaluation points for the polynomial being interpolated. Thus, the essence of this interpolation technique can be stated as follows.

Lemma 6.2.1 (Lemma 5.2 in [47]). *Suppose $M \in \mathbb{C}^{2\times 2}$ and $s \in \mathbb{C}^{2\times 1}$. If the following three conditions are satisfied,*

1. $\det(M) \neq 0$;

2. s is not a column eigenvector of M (nor the zero vector);

3. the ratio of the eigenvalues of M is not a root of unity;

then the vectors in the set $V = \{M^k s\}_{k \geq 0}$ *are pairwise linearly independent.*

Clearly the first condition is necessary. The second condition is equivalent to $\det([s\ Ms]) \neq 0$, which is necessary since it checks the linear dependence of the first two vectors in V.

The unary recursive construction can be generalized to larger dimensions, where the starter gadget has arity d and the recursive gadget has arity $2d$ [93]. Then the starter gadget is represented by its signature matrix with parameter d, which is a column vector in $\mathbb{C}^{2^d}$ The recursive gadget is also represented by its signature matrix with parameter d, which is a matrix in $\mathbb{C}^{2^d \times 2^d}$.

For dimensions larger than one, the second condition in Lemma 6.2.1 must be replaced by a stronger assumption, such as "*s is not orthogonal to any row eigenvector of M*" [50]. Previous work (Lemma 4.10 in [94]) satisfied this stronger condition by showing that it follows from $\det([s\ Ms\ \ldots\ M^{n-1}s]) \neq 0$. For completeness, we show that these two conditions are equivalent. The use of n instead of 2^d in the next two lemmas is not overly general. Sometimes degeneracies or redundancies in the starter and recursive gadgets warrant the consideration of such cases.

Lemma 6.2.2. *Suppose* $M \in \mathbb{C}^{n \times n}$ *and* $s \in \mathbb{C}^{n \times 1}$*. Then s is not orthogonal to any row eigenvector of M iff* $\det([s\ Ms\ \ldots\ M^{n-1}s]) \neq 0$*.*

Proof. Suppose that s is not orthogonal to any row eigenvector of M and assume for a contradiction that $\det([s\ Ms\ \ldots\ M^{n-1}s]) = 0$. Then there is a nonzero row vector v such that $v[s\ Ms\ \ldots\ M^{n-1}s] = \mathbf{0}$ is the zero vector. Consider the linear span S by row vectors in the set $\{v, vM, \ldots, vM^{n-1}\}$. We claim that S is an invariant subspace of row vectors under the action of multiplication by M from the right.

By the Cayley-Hamilton theorem, M satisfies its own characteristic polynomial, which is a monic polynomial of degree n. Thus, M^n is a linear combination of $I_n, M, \ldots, M^{n-1}$. This shows that for any $u \in S$, uM still belongs to S. Therefore, there exists a $u \in S$ such that u is a row eigenvector of M. By the definition of S, this u is orthogonal to s, which is a contradiction.

In the other direction, suppose $\det([s\ Ms\ \ldots\ M^{n-1}s]) \neq 0$ and assume for a contradiction that

s is orthogonal to some row eigenvector v of M with eigenvalue λ. Then $v[s\ Ms\ \ldots\ M^{n-1}s] = \mathbf{0}$ is the zero vector because $vM^i s = \lambda^i vs = 0$. Since $v \neq \mathbf{0}$, this a contradiction. □

Another necessary condition, even for the d-dimensional case, is that M has infinite order modulo a scalar. Otherwise, $M^k = \beta I_n$ for some k and any vector of the form $M^\ell s$ for $\ell \geq k$ is some multiple of a vector in the set $\{M^i s\}_{0 \leq i < k}$. We improve the d-dimensional version of Lemma 6.2.1 by replacing the third condition with this necessary condition.

Lemma 6.2.3. *Suppose $M \in \mathbb{C}^{n\times n}$ and $s \in \mathbb{C}^{n\times 1}$. If the following three conditions are satisfied,*

1. $\det(M) \neq 0$*;*
2. *s is not orthogonal to any row eigenvector of M;*
3. *M has infinite order modulo a scalar;*

then the vectors in the set $V = \{M^k s\}_{k\geq 0}$ are pairwise linearly independent.

Proof. Since $\det(M) \neq 0$, M is nonsingular and the eigenvalues λ_i of M, for $1 \leq i \leq n$, are nonzero. Let $M = P^{-1}JP$ be the Jordan decomposition of M and let $p = Ps \in \mathbb{C}^{n\times 1}$. Suppose for a contradiction that the vectors in V are not pairwise linearly independent. This means that there exists integers $k > \ell \geq 0$ such that $M^k s = \beta M^\ell s$ for some nonzero complex value β. Let $t = k - \ell > 0$. Then we have $P^{-1}J^t P s = M^t s = \beta s$ and $J^t p = \beta p$.

Suppose that J contains some nontrivial Jordan block and consider the 2-by-2 submatrix in the bottom right corner of this block. From this portion of J, the two equations given by $J^t p = \beta p$ are $\lambda_i^t p_{i-1} + t\lambda_i^{t-1} p_i = \beta p_{i-1}$ and $\lambda_i^t p_i = \beta p_i$. Since s is not orthogonal to any row eigenvector of M, $p_i \neq 0$. But then these equations imply that $t\lambda_i^{t-1} p_i = 0$, a contradiction.

Otherwise, J contains only trivial Jordan blocks. From $J^t p = \beta p$, we get the equations $\lambda_i p_i = \beta p_i$ for $1 \leq i \leq n$. Since s is not orthogonal to any row eigenvector of M, $p_i \neq 0$ for $1 \leq i \leq n$. But then $M^t = \beta I_n$, which contradicts that fact that M has infinite order modulo a scalar. □

With this lemma, we obtain a tight characterization for the success of interpolation by a unary recursive construction. For example, the construction using a recursive gadget with signature matrix $M = \left[\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right]$ and a starter gadget with signature $s = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]$ is successful because M and s satisfy our conditions but do not satisfy previous sufficient conditions.

Lemma 6.2.4. *Let $\mathcal{F}$ be a set of signatures. If there exists a planar $\mathcal{F}$-gate with signature matrix $M \in \mathbb{C}^{2\times 2}$ and a planar $\mathcal{F}$-gate with signature $s \in \mathbb{C}^{2\times 1}$ satisfying the following conditions,*

1. $\det(M) \neq 0$;
2. $\det([s\ Ms]) \neq 0$;
3. *M has infinite order modulo a scalar;*

then $\text{Pl-Holant}(\mathcal{F} \cup \{[a, b]\}) \leq_T \text{Pl-Holant}(\mathcal{F})$ *for any* $a, b \in \mathbb{C}$.

Proof. Consider an instance $\Omega = (G, \pi)$ of $\text{Pl-Holant}(\mathcal{F} \cup \{[a, b]\})$. Let V' be the subset of vertices assigned $[a, b]$ by π and suppose that $|V'| = n$. We construct from Ω a sequence of instances Ω_k of $\text{Pl-Holant}(\mathcal{F})$ indexed by $k \geq 1$. We obtain Ω_k from Ω by replacing each occurrence of $[a, b]$ with the unary recursive construction (M, s) in Figure 6.1 containing k copies of the recursive gadget. This unary recursive construction has the signature $[x_k, y_k] = M^k s$.

By applying our assumptions to Lemmas 6.2.2 and 6.2.3, we know that the signatures in the set $V = \{[x_k, y_k] \mid 0 \leq k \leq n + 1\}$ are pairwise linearly independent. In particular, at most one y_k can be 0, so we may assume that $y_k \neq 0$ for $0 \leq k \leq n$, renaming variables if necessary.

We stratify the assignments in Ω based on the assignment to $[a, b]$. Let c_ℓ be the sum over all assignments of products of evaluations at all $v \in V(G) - V'$ such that exactly ℓ occurrences of $[a, b]$ have their incident edge assigned 0 (and $n - \ell$ have their incident edge assigned 1). Then

$$\text{Holant}(\Omega) = \sum_{0\leq \ell \leq n} a^\ell b^{n-\ell} c_\ell$$

and the value of the Holant on Ω_k, for $k \geq 1$, is

$$\begin{aligned}\text{Holant}(\Omega_k) &= \sum_{0\leq \ell \leq n} x_k^\ell y_k^{n-\ell} c_\ell \\ &= y_k^n \sum_{0\leq \ell \leq n} \left(\frac{x_k}{y_k}\right)^\ell c_\ell.\end{aligned}$$

The coefficient matrix of this linear system is Vandermonde. Since the signatures in V are pairwise linearly independent, the ratios x_k/y_k are distinct (and well-defined since $y_k \neq 0$), which means that the Vandermonde matrix has full rank. Therefore, we can solve the linear system for the

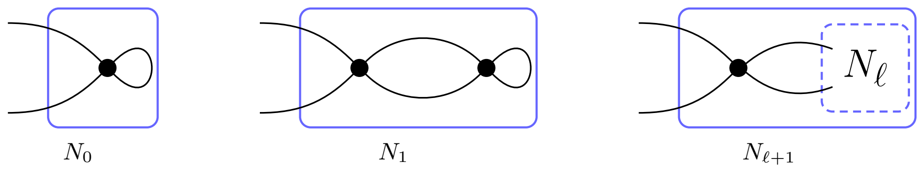


Figure 6.2: Binary recursive construction with starter gadget used to interpolate $[1, 0, 0]$. The vertices are assigned $[v, 1, 0, 0, 0]$.

unknown c_ℓ's and obtain the value of Holant(Ω). □

The first two conditions of Lemma 6.2.4 are easy to check. The third condition holds in one of these two cases: either the eigenvalues are the same but M is not a multiple of the identity matrix, or the eigenvalues are different but their ratio is not a root of unity.

Our refined conditions work well with the anti-gadget technique. The power of this lemma is that when the third condition fails to hold, there exists an integer k such that $M^k = I_2$, where I_2 is the 2-by-2 identity matrix. Therefore we can construct $M^{k-1} = M^{-1}$ and use this in other gadget constructions.

6.3 Planar Pairing

In this section, we consider the problem Pl-Holant($[v, 1, 0, 0, 0]$) when v is different from 0. Over the next two lemmas, we prove that this problem is #P-hard by reducing from Pl-Holant($[v, 1, 0, 0]$). These problems are weighted versions of counting matchings over planar k-regular graphs for $k = 4$ and $k = 3$ respectively. The key idea, defined in Definition 6.3.3, is that of a planar pairing.

In the first lemma, we show how to use either the anti-gadget technique or interpolation by our tight characterization of the unary recursive construction from Section 6.2 to effectively obtain $[1, 0, 0]$. The construction in this proof is actually not a unary recursive construction, but a binary recursive construction. However, degeneracies in the starter and recursive gadgets permit analysis equivalent to that of the unary recursive construction.

Lemma 6.3.1. *For any $v \in \mathbb{C}$ and signature set $\mathcal{F}$ containing $[v, 1, 0, 0, 0]$,*

$$\text{Pl-Holant}(\mathcal{F} \cup \{[1, 0, 0]\}) \leq_T \text{Pl-Holant}(\mathcal{F}).$$

Proof. Consider the gadget construction in Figure 6.2. For $k \geq 0$, the signature of N_k is of the form $[a_k, b_k, 0]$, and $N_0 = [v, 1, 0]$. Since N_k is symmetric and always ends with 0, we can analyze this construction as though it were a unary recursive construction. Let $s_k = \left[\begin{smallmatrix} a_k \\ b_k \end{smallmatrix}\right]$, so $s_0 = \left[\begin{smallmatrix} v \\ 1 \end{smallmatrix}\right]$. It is clear that $s_k = M^k s_0$, where $M = \left[\begin{smallmatrix} v & 2 \\ 1 & 0 \end{smallmatrix}\right]$.

Now M is nonsingular since since $\det(M) = -2$. If M has finite order modulo a scalar, then $M^\ell = \beta I_2$ for some integer $\ell \geq 1$ and some nonzero $\beta \in \mathbb{C}$. Then $N_{\ell-1}$ is an anti-gadget of M with signature $M^{\ell-1} s_0 = \beta M^{-1} s_0 = \beta \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$, and we have $[1, 0, 0]$ after normalizing.

Now assume that M has infinite order modulo a scalar. Since $\det([s_0 \;\, M s_0]) = -2$, we can interpolate any signature of the form $[x, y, 0]$ by Lemma 6.2.4, including $[1, 0, 0]$. $\square$

For the next lemma, we use a well-known and easy generalization of a classic result of Petersen [110]. Petersen's theorem considers 3-regular, bridgeless, simple graphs and concludes that there exists a perfect matching. The same conclusion holds even if the graphs are not simple. We provide a proof for completeness.

Theorem 6.3.2. *Any 3-regular bridgeless graph G has a perfect matching.*

Proof. We may assume that G is connected. If G has a vertex v with a self-loop, then the other edge of v is a bridge since G is 3-regular, which is a contradiction. If there exists some pair of vertices of G joined by exactly three parallel edges, then G has only these two vertices since it is connected and the theorem holds.

In the remaining case, there exists some pair of vertices joined by exactly two parallel edges. We build a new graph G' without any parallel edges. For vertices u and v joined by exactly two parallel edges, we remove these two parallel edges and introduce two new vertices w_1 and w_2. We also introduce the new edges (u, w_1), (u, w_2), (v, w_1), (v, w_2), and (w_1, w_2). Then G' is a 3-regular, bridgeless, simple graph.

By Petersen's theorem, G' has a perfect matching P'. Now we construct a perfect matching P in G using P'. We put any edge in both G and P' into P. If u is matched by a new edge in G', then v must be matched by a new edge in G' as well and we put the edge (u, v) into P. If u and v are not matched by a new edge, then we do not add anything to P. It is easy to see that P is a

(a) The neighborhood around u and v in G. (b) The same neighborhood in H.

Figure 6.3: The neighborhood around u and v both before and after they are removed.

perfect matching in G. □

We use this result to show the existence of what we call a *planar pairing* for any planar 3-regular graph, which we use in our proof of #P-hardness.

Definition 6.3.3 (Planar pairing)**.** A *planar pairing* in a graph $G = (V, E)$ is a set of edges $P \subset V \times V$ such that P is a perfect matching in the graph $(V, V \times V)$, and the graph $(V, E \cup P)$ is planar.

Obviously, a perfect matching in the original graph is a planar pairing.

Lemma 6.3.4. *For any planar 3-regular graph G, there exists a planar pairing that can be computed in polynomial time.*

Proof. We efficiently find a planar pairing in G by induction on the number of vertices in G. Since G is a 3-regular graph, it must have an even number of vertices. If there are no vertices in G, then there is nothing to do. Suppose that G has $n = 2k$ vertices and that we can efficiently find a planar pairing in graphs containing fewer vertices. If G is not connected, then we can already apply our inductive hypothesis on each connected component of G. The union of planar pairings in each connected component of G is a planar pairing in G, so we are done. Otherwise assume that G is connected.

Suppose that G contains a bridge (u, v). Let the three (though not necessarily distinct) neighbors of u be v, u_1, and u_2, and let the three (though not necessarily distinct) neighbors of v be u, v_1, and v_2 (see Figure 6.3a). Furthermore, let H_u be the connected component in $G - \{(u, v)\}$ containing u and let H_v be the connected component in $G - \{(u, v)\}$ containing v. Consider the induced subgraph H'_u of H_u after adding the edge $e_u = (u_1, u_2)$ (which might be a self-loop on $u = u_1 = u_2$) and then removing u. Similarly, consider the induced subgraph H'_v of H_v after adding the edge $e_v = (v_1, v_2)$ (which might be a self-loop on $v = v_1 = v_2$) and then removing v. Both H'_u and H'_v are 3-regular graphs and their disjoint union gives a graph H' with $n-2 = 2(k-1)$ vertices (see Figure 6.3b).

By induction on both H'_u and H'_v, we have planar pairings P_u and P_v in H'_u and H'_v respectively. Let H'' be the graph H' including the edges $P_u \cup P_v$. If H'' contains both e_u and e_v, then embed H'' in the plane so that both e_u and e_v are adjacent to the outer face. Otherwise, any planar embedding will do. Then the graph G including the edges $P_u \cup P_v$ is also planar, so $P_u \cup P_v \cup \{(u, v)\}$ is a planar pairing in G.

Otherwise, G is bridgeless. Then by Theorem 6.3.2, G has a perfect matching, which is also a planar pairing in G. Since a perfect matching can be found in polynomial time by Edmond's blossom algorithm [61], the whole procedure is in polynomial time. □

After publishing a preliminary version of [74], we realized that a previous construction by Cai and Kowalczyk uses a planar pairing to show that counting vertex covers over k-regular graphs is #P-hard for even $k \geq 4$ (see the proof of Lemma 15 in [37]). Their algorithm to find a planar pairing starts by taking a spanning tree and then pairing up the vertices on this tree, which is simpler than our approach. We believe that it is worth emphasizing the importance of a planar pairing. Most gadget constructions in hardness proofs for Holant problems are local but the planar pairing technique is a global argument, which permits reductions that are not otherwise possible.

Now we use the planar pairing technique to show the following.

Lemma 6.3.5. *Let* $v \in \mathbb{C}$. *Then* Pl-Holant$([v, 1, 0, 0]) \leq_T$ Pl-Holant$([v, 1, 0, 0, 0])$.

Proof. An instance of Pl-Holant$([v, 1, 0, 0])$ is a signature grid Ω with underlying graph $G = (V, E)$ that is planar and 3-regular. By Lemma 6.3.4, there exists a planar pairing P in G and it can be found in polynomial time. Then the graph $G' = (V, E \cup P)$ is planar and 4-regular. We assign $[v, 1, 0, 0, 0]$ to every vertex in G'. By Lemma 6.3.1, we can assume that we have $[1, 0, 0]$. We replace each edge in P with a path of length 2 to form a graph G'' and assign $[1, 0, 0] = [1, 0]^{\otimes 2}$ to each of the new vertices. Then the signature grid Ω'' with underlying graph G'' has the same Holant value as the original signature grid Ω. □

Corollary 6.3.6. *Let* $v \in \mathbb{C}$. *If* $v \neq 0$, *then* Pl-Holant$([v, 1, 0, 0, 0])$ *is* #P*-hard.*

Proof. By Lemma 6.3.5, it suffices to show that Pl-Holant$([v, 1, 0, 0])$ is #P-hard. Then we are done by applying Theorem 6.1.3, the planar Holant dichotomy for a single ternary signature. □

6.4 Counting Eulerian Orientations

Recall the definition of an Eulerian orientation.

Definition 6.4.1. Given a graph G, an orientation of its edges is an *Eulerian orientation* if for each vertex v of G, the number of incoming edges of v equals the number of outgoing edges of v.

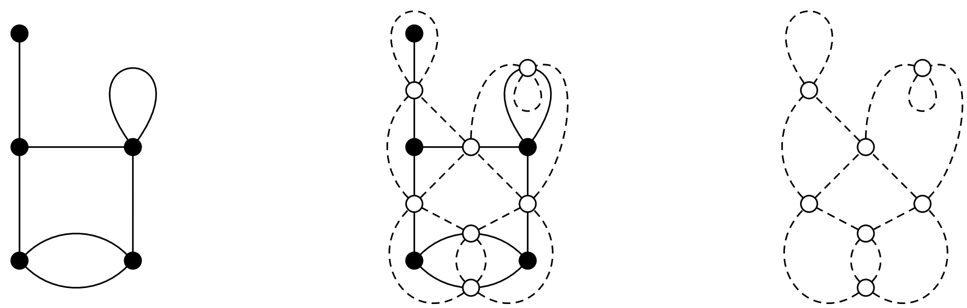
Counting (unweighted) Eulerian orientations over 4-regular graphs was shown to be #P-hard in Theorem V.10 of [80]. We improve this result by showing that this problem remains #P-hard when the graph is also planar. Our reduction is from evaluating the Tutte polynomial of a planar graph at a point on its diagonal, which have an alternative expression involving medial graphs.

Definition 6.4.2 (cf. [12])**.** For a connected plane graph G (i.e. a planar embedding of a connected planar graph), its *medial graph* H has a vertex for each edge of G and two vertices in H are joined by an edge for each face of G in which their corresponding edges occur consecutively.

An example of a plane graph and its medial graph are given in Figure 6.4. Any medial graph is planar and 4-regular. Las Vergnas [136] connected the evaluation of the Tutte polynomial of a plane graph at the point (3,3) with a sum of weighted Eulerian orientations on its medial graph.

Theorem 6.4.3 (Theorem 2.1 in [136])**.** *Let* G *be a connected plane graph and let* $\mathscr{O}(G_m)$ *be the set of all Eulerian orientations in the medial graph* G_m *of* G. *Then*

$$2 \cdot \text{Tutte}(G; 3, 3) = \sum_{O \in \mathscr{O}(G_m)} 2^{\beta(O)}, \tag{6.4.1}$$



(a) (b) (c)

Figure 6.4: A plane graph (a), its medial graph (c), and both graphs superimposed (b).

where $\beta(O)$ is the number of saddle vertices in the orientation O, i.e. the number of vertices in which the edges are oriented "in, out, in, out" in cyclic order.

Although the medial graph depends on a particular embedding of the planar graph G, the right side of (6.4.1) is invariant under different embeddings of G. This follows from (6.4.1) and the fact that the Tutte polynomial does not depend on the embedding of G.

Now we can prove our hardness result.

Theorem 6.4.4. #Eulerian-Orientations *is* #P*-hard over planar 4-regular graphs.*

Proof. We reduce from the right side of (6.4.1) to Pl-Holant($\neq_2$ | $[0, 0, 1, 0, 0]$), which is the problem of counting Eulerian orientations over planar 4-regular graphs. Then by Theorem 6.1.4 and Theorem 6.4.3, we conclude that Pl-Holant($\neq_2$ | $[0, 0, 1, 0, 0]$) is #P-hard.

The right side of (6.4.1) is the bipartite Holant problem Pl-Holant($\neq_2$ | f), where

$$M_f = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

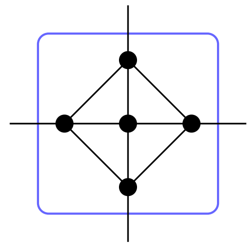


Figure 6.5: The planar tetrahedron gadget. Each vertex is assigned $[3,0,1,0,3]$.

We perform a holographic transformation by $Z = \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$ to get

$$\begin{aligned}
\text{Pl-Holant}(\neq_2 \mid f) &\equiv_T \text{Pl-Holant}([0,1,0](Z^{-1})^{\otimes 2} \mid Z^{\otimes 4} f) \\
&\equiv_T \text{Pl-Holant}([1,0,1]/2 \mid 4\hat{f}) \\
&\equiv_T \text{Pl-Holant}(\hat{f}),
\end{aligned}$$

where the signature matrix of $\hat{f}$ is

$$M_{\hat{f}} = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}.$$

We also perform the same holographic transformation by Z on our target counting problem Pl-Holant$(\neq_2 \mid [0,0,1,0,0])$ to get

$$\begin{aligned}
\text{Pl-Holant}(\neq_2 \mid [0,0,1,0,0]) &\equiv_T \text{Pl-Holant}([0,1,0](Z^{-1})^{\otimes 2} \mid Z^{\otimes 4}[0,0,1,0,0]) \\
&\equiv_T \text{Pl-Holant}([1,0,1]/2 \mid 2[3,0,1,0,3]) \\
&\equiv_T \text{Pl-Holant}([3,0,1,0,3]).
\end{aligned}$$

Using the planar tetrahedron gadget in Figure 6.5, we assign $[3,0,1,0,3]$ to every vertex and obtain

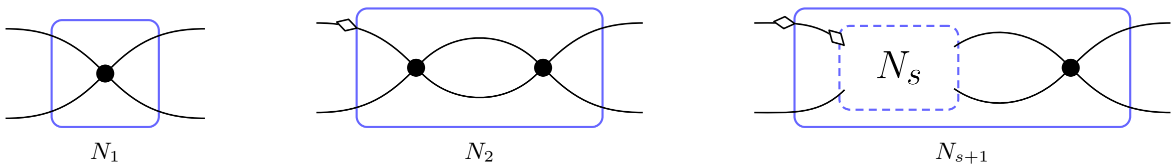


Figure 6.6: Simple binary recursive construction.

a gadget with signature $32\hat{g}$, where the signature matrix of $\hat{g}$ is

$$M_{\hat{g}} = \frac{1}{2}\begin{bmatrix} 19 & 0 & 0 & 7 \\ 0 & 7 & 5 & 0 \\ 0 & 5 & 7 & 0 \\ 7 & 0 & 0 & 19 \end{bmatrix}.$$

Now we show how to reduce Pl-Holant$(\hat{f})$ to Pl-Holant$(\hat{g})$ by interpolation. Consider an instance Ω of Pl-Holant$(\hat{f})$. Suppose that $\hat{f}$ appears n times in Ω. We construct from Ω a sequence of instances Ω_s of Holant$(\hat{g})$ indexed by $s \geq 1$. We obtain Ω_s from Ω by replacing each occurrence of $\hat{f}$ with the gadget N_s in Figure 6.6 with $\hat{g}$ assigned to all vertices. Since $\hat{f}$ and $\hat{g}$ are rotationally symmetric, it is unnecessary to specify which edge corresponds to which input.

To obtain Ω_s from Ω, we effectively replace $M_{\hat{f}}$ with $M_{N_s} = (M_{\hat{g}})^s$, the sth power of the signature matrix $M_{\hat{g}}$. Let

$$T = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Then

$$M_{\hat{f}} = T\Lambda_{\hat{f}}T^{-1} = T\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}T^{-1} \qquad \text{and} \qquad M_{\hat{g}} = T\Lambda_{\hat{g}}T^{-1} = T\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 13 \end{bmatrix}T^{-1}.$$

We can view our construction of Ω_s as first replacing each $M_{\hat{f}}$ by $T\Lambda_{\hat{f}}T^{-1}$ to obtain a signature grid Ω', which does not change the Holant value, and then replacing each $\Lambda_{\hat{f}}$ with $\Lambda_{\hat{g}}^s$. We stratify the assignments in Ω' based on the assignment to $\Lambda_{\hat{f}}$. We only need to consider the assignments to $\Lambda_{\hat{f}}$ that assign

- 0000 j many times,
- 0110 or 1001 k many times, and
- 1111 ℓ many times.

Let $c_{jk\ell}$ be the sum over all such assignments of the products of evaluations from T and T^{-1} but excluding $\Lambda_{\hat{f}}$ on Ω'. Then

$$\text{Holant}(\Omega) = \sum_{j+k+\ell=n} 3^\ell c_{jk\ell}$$

and the value of the Holant on Ω_s, for $s \geq 1$, is

$$\text{Holant}(\Omega_s) = \sum_{j+k+\ell=n} (6^k 13^\ell)^s c_{jk\ell}. \tag{6.4.2}$$

This coefficient matrix in the linear system of (6.4.2) is Vandermonde and of full rank since for any $0 \leq k + \ell \leq n$ and $0 \leq k' + \ell' \leq n$ such that $(k, \ell) \neq (k', \ell')$, $6^k 13^\ell \neq 6^{k'} 13^{\ell'}$. Therefore, we can solve the linear system for the unknown $c_{jk\ell}$'s and obtain the value of Holant(Ω). □

The previous proof can be easily modified to reduce from #EO over 4-regular graphs by interpolating the so-called crossover signature, which has signature matrix

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Conceptually, the current proof is simpler because the #P-hardness proof for #EO over 4-regular graphs in [80] reduces from the same starting point as our current proof.

6.5 A Demanding Binary Interpolation

Our main result in this chapter is a planar Holant dichotomy for a symmetric signature f of arity 4, and our main technique to prove the hardness is polynomial interpolation. The binary recursive construction is a particularly simple construction for which polynomial interpolation is likely to succeed. To verify the success of this interpolation, we analyze the signature matrix of f. Since f is symmetric, the signature matrix of f has some redundancies: its second and third rows are the same and its second and third columns are the same. We "remove" this extra information using the following definition.

Definition 6.5.1. A 4-by-4 matrix is *redundant* if its middle two rows and middle two columns are the same. Denote the set of all redundant 4-by-4 matrices over $\mathbb{C}$ by $\mathrm{RM}_4(\mathbb{C})$.

Consider the function $\varphi : \mathbb{C}^{4\times 4} \to \mathbb{C}^{3\times 3}$ defined by

$$\varphi(M) = AMB,$$

where

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Intuitively, the operation φ replaces the middle two columns of M with their sum and then the middle two rows of M with their average. (These two steps commute.) Conversely, we have the following function $\psi : \mathbb{C}^{3\times 3} \to \mathrm{RM}_4(\mathbb{C})$ defined by

$$\psi(N) = BNA.$$

Intuitively, the operation ψ duplicates the middle row of N and then splits the middle column evenly into two columns. Notice that $\varphi(\psi(N)) = N$. When restricted to $\mathrm{RM}_4(\mathbb{C})$, φ is an isomorphism between the semi-group of 4-by-4 redundant matrices and the semi-group of 3-by-3 matrices, under matrix multiplication, and ψ is its inverse. To see this, just notice that

$$AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad BA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

are the identity elements of their respective semi-groups. If the signature matrix M_g of an arity 4 signature g is redundant, we also define the *compressed signature matrix* of g as $\widetilde{M_g} = \varphi(M_g)$.

If all signatures in an $\mathcal{F}$-gate have even arity, then the signature of any $\mathcal{F}$-gate also has even arity. Knowing that binary signatures alone do not produce #P-hardness, and with the above constraint in mind, we would like to interpolate another signature of arity 4 using a given signature of arity 4. We are particularly interested in the signature g with signature matrix

$$M_g = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6.5.3}$$

the identity element in the semi-group of redundant matrices. Thus $\widetilde{M_g} = I_3$.

We prove that this signature g is hard. For an asymmetric signature, we often want to reorder the input bits under a circular permutation. For a single counterclockwise rotation by $90°$, the effect on the entries of the signature matrix of an arity 4 signature is given in Figure 3.2.

Lemma 6.5.2. *Let g be the arity 4 signature with M_g given in (6.5.3) so that $\widetilde{M_g} = I_3$. Then* Pl-Holant(g) *is* #P*-hard.*

Proof. We reduce from Pl-Holant$(\neq_2 \mid [0,0,1,0,0])$, the problem of counting Eulerian orientation over planar 4-regular graphs, which is #P-hard by Theorem 6.4.4. Recall from the proof of that

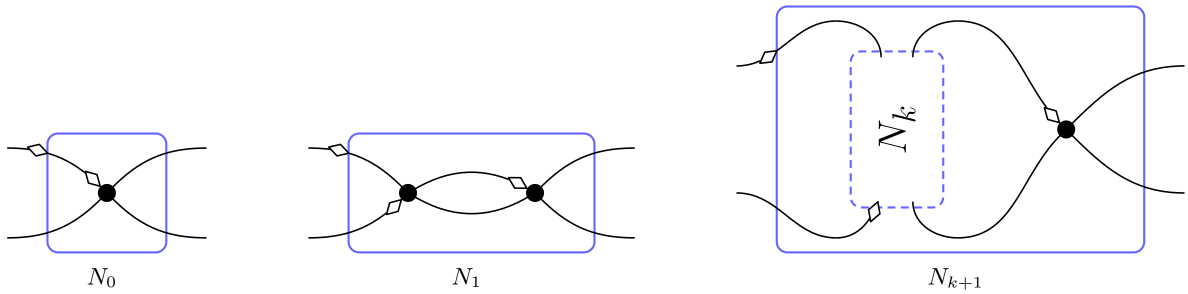


Figure 6.7: Recursive construction to approximate $[1, 0, \frac{1}{3}, 0, 1]$. Vertices are assigned g.

theorem that

$$\begin{aligned}\text{Pl-Holant}(\neq_2 \mid [0,0,1,0,0]) &\equiv_T \text{Pl-Holant}([0,1,0](Z^{-1})^{\otimes 2} \mid Z^{\otimes 4}[0,0,1,0,0]) \\ &\equiv_T \text{Pl-Holant}(\tfrac{1}{2}[1,0,1] \mid \tfrac{2}{3}[1,0,\tfrac{1}{3},0,1]) \\ &\equiv_T \text{Pl-Holant}([1,0,\tfrac{1}{3},0,1]),\end{aligned}$$

where $Z = \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$. Let $\mathscr{O} = [1, 0, \frac{1}{3}, 0, 1]$, We reduce from Holant($\mathscr{O}$) via an arbitrarily close approximation using the recursive construction in Figure 6.7 with g assigned to every vertex.

We claim that the signature matrix M_{N_k} of Gadget N_k is

$$M_{N_k} = \begin{bmatrix} 1 & 0 & 0 & a_k \\ 0 & a_{k+1} & a_{k+1} & 0 \\ 0 & a_{k+1} & a_{k+1} & 0 \\ a_k & 0 & 0 & 1 \end{bmatrix},$$

where $a_k = \frac{1}{3} - \frac{1}{3}\left(-\frac{1}{2}\right)^k$. This is true for N_0. Inductively assume M_{N_k} has this form. Then the rotated form of the signature matrix for N_k, as described in Figure 3.2, is

$$\begin{bmatrix} 1 & 0 & 0 & a_{k+1} \\ 0 & a_k & a_{k+1} & 0 \\ 0 & a_{k+1} & a_k & 0 \\ a_{k+1} & 0 & 0 & 1 \end{bmatrix}. \tag{6.5.4}$$

The action of g on the far right side of N_{k+1} is to replace each of the middle two entries in the middle two rows of the matrix in (6.5.4) with their average, $(a_k + a_{k+1})/2 = a_{k+2}$. This gives $M_{N_{k+1}}$.

Let G be a graph with n vertices and $H_{\mathscr{O}}$ (resp. H_{N_k}) be the Holant value on G with all vertices assigned $\mathscr{O}$ (resp. N_k). Since each signature entry in $\mathscr{O}$ can be expressed as a rational number with denominator 3, each term in the sum of $H_{\mathscr{O}}$ can be expressed as a rational number with denominator 3^n, and $H_{\mathscr{O}}$ itself is a sum of 2^{2n} such terms, where $2n$ is the number of edges in G. If the error $|H_{N_k} - H_{\mathscr{O}}|$ is at most $1/3^{n+1}$, then we can recover $H_{\mathscr{O}}$ from H_{N_k} by selecting the nearest rational number to H_{N_k} with denominator 3^n.

For each signature entry x in $M_{\mathscr{O}}$, its corresponding entry $\tilde{x}$ in M_{N_k} satisfies $|\tilde{x} - x| \leq x/2^k$. Then for each term t in the Holant sum $H_{\mathscr{O}}$, its corresponding term $\tilde{t}$ in the sum H_{N_k} satisfies $t(1 - 1/2^k)^n \leq \tilde{t} \leq t(1 + 1/2^k)^n$, thus $-t(1 - (1 - 1/2^k)^n) \leq \tilde{t} - t \leq t((1 + 1/2^k)^n - 1)$. Since $1 - (1 - 1/2^k)^n \leq (1 + 1/2^k)^n - 1$, we get $|\tilde{t} - t| \leq t((1 + 1/2^k)^n - 1)$. Also each term $t \leq 1$. Hence

$$|H_{N_k} - H_{\mathscr{O}}| \leq 2^{2n} \left(\left(1 + 1/2^k\right)^n - 1 \right) < 1/3^{n+1},$$

if we take $k = 4n$. □

When I explained this proof to Xi Chen, he thought that it should be possible to prove hardness using polynomial interpolation instead of this approximation argument. He was right, as I now demonstrate.

Alternate proof of Lemma 6.5.2. As in the proof of Theorem 6.4.4, we reduce from Pl-Holant(f), where

$$M_f = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}.$$

Consider an instance Ω of this problem. Suppose that f appears n times in Ω. We construct from Ω a sequence of instances Ω_s of Holant(g) indexed by $s \geq 0$. We obtain Ω_s from Ω by replacing each occurrence of f with the gadget N_s in Figure 6.8 with g' assigned to all vertices, where g' is obtained from $2g$ after a counterclockwise rotation by a quarter turn as in Figure 3.2. The signature matrix of g' is

$$M_{g'} = \begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}.$$

To obtain Ω_s from Ω, we effectively replace M_f with $M_{N_s} = (M_{g'})^s$, the sth power of the signature matrix $M_{g'}$. Let

$$T = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Then

$$M_f = T \Lambda_f T^{-1} = T \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} T^{-1} \qquad \text{and} \qquad M_{g'} = T \Lambda_{g'} T^{-1} = T \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} T^{-1}.$$

We stratify the assignments in Ω based on the assignment to Λ_f. We only need to consider the assignments to f that assign

- 0000 or 0110 or 1001 $n-i$ many times and
- 1111 i many times

since any other assignment contributes a factor of 0. Let c_i be the sum over all such assignments of the products of evaluations from T and T^{-1} but excluding Λ_f on Ω. Then

$$\operatorname{Holant}(\Omega) = \sum_{0 \le i \le n} 3^i c_i \qquad \text{and} \qquad \operatorname{Holant}(\Omega_t) = \sum_{0 \le i \le n} 9^{ti} c_i$$

for $t = 2s \ge 0$.

Let $p(x) = \sum_{0 \le i \le n} x^i c_i$, which has degree at most n with integer coefficients. Using our oracle for Pl-Holant(g'), we can evaluate this polynomial at $n+1$ distinct points $x = 9^t$ for $0 \le t \le n$. Then via polynomial interpolation, we can recover the coefficients of this polynomial efficiently. Evaluating this polynomial at $x = 3$ gives the value of Holant(Ω), as desired. □

Remark. One advantage of the alternative proof is that it directly reduces from Pl-Holant(f) instead of going through Pl-Holant($[3,0,1,0,3]$). Furthermore, there is no longer a need to give special attention to Pl-Holant($[3,0,1,0,3]$), the problem of counting Eulerian orientations over planar 4-regular graphs. With this alternative proof, the signature $[3,0,1,0,3]$ is now swept up by Lemma 6.5.4 with all the other signatures with redundant and nonsingular compressed signature matrices.

Most symmetric signatures of arity 4 are hard because they can be used to interpolate the signature g whose signature matrix is given in (6.5.3). We prove this in Lemma 6.5.4. There are three cases in Lemma 6.5.4 and one of them requires the following technical lemma.

Lemma 6.5.3. *Let $M = [B_0\ B_1\ \cdots\ B_t]$ be an n-by-n block matrix such that there exists a $\lambda \in \mathbb{C}$, for all integers $0 \le k \le t$, block B_k is an n-by-c_k matrix for some integer $c_k \ge 0$, and the entry of B_k at row r and column c is $(B_k)_{rc} = r^{c-1} \lambda^{kr}$, where $r, c \ge 1$. If λ is nonzero and is not a root of unity, then M is nonsingular.*

Proof. We prove by induction on n. If $n = 1$, then the sole entry is λ^k for some nonnegative integer k. This is nonzero since $\lambda \neq 0$. Assume $n > 1$ and let the left-most nonempty block be B_j. We divide row r by λ^{jr}, which is allowed since $\lambda \neq 0$. This effectively changes block B_ℓ into a block of the form $B_{\ell-j}$. Thus, we have another matrix of the same form as M but with a nonempty block B_0. To simplify notation, we also denote this matrix again by M. The first column of B_0 is all 1's. We subtract row $r-1$ from row r, for r from n down to 2. This gives us a new matrix M', and $\det M = \det M'$. Then $\det M'$ is the determinant of the $(n-1)$-by-$(n-1)$ submatrix M'' obtained from M' by removing the first row and column. Now we do column operations (on M'') to return the blocks to the proper form so that we can invoke the induction hypothesis.

For any block B'_k different from B'_0, we prove by induction on the number of columns in B'_k that B'_k can be repaired. In the base case, the rth element of the first column is $(B'_k)_{r1} = \lambda^{kr} - \lambda^{k(r-1)} = \lambda^{k(r-1)}(\lambda^k - 1)$ for $r \geq 2$. We divide this column by $\lambda^k - 1$ to obtain $\lambda^{k(r-1)}$, which is allowed since λ is not a root of unity and $k \neq 0$. This is now the correct form for the rth element of the first column of a block in M''.

Now for the inductive step, assume that the first $d-1$ columns of block B'_k are in the correct form to be a block in M''. That is, for row index $r \geq 2$, which denotes the $(r-1)$-th row of M'', the rth element in the first $d-1$ columns of B'_k have the form $(B'_k)_{rc} = (r-1)^{c-1}\lambda^{k(r-1)}$. The rth element in column d of B'_k currently has the form $(B'_k)_{rd} = r^{d-1}\lambda^{kr} - (r-1)^{d-1}\lambda^{k(r-1)}$. Then we do column operations

$$\begin{aligned}(B'_k)_{rd} - \sum_{c=1}^{d-1}\binom{d-1}{c-1}(B'_k)_{rc} &= r^{d-1}\lambda^{kr} - (r-1)^{d-1}\lambda^{k(r-1)} - \sum_{c=1}^{d-1}\binom{d-1}{c-1}(r-1)^{c-1}\lambda^{k(r-1)} \\ &= r^{d-1}\lambda^{kr} - r^{d-1}\lambda^{k(r-1)} \\ &= r^{d-1}\lambda^{k(r-1)}(\lambda^k - 1)\end{aligned}$$

and divide by $(\lambda^k - 1)$ to get $r^{d-1}\lambda^{k(r-1)}$. Once again, this is allowed since λ is not a root of unity and $k \neq 0$. Then more (of the same) column operations yield

$$r^{d-1}\lambda^{k(r-1)} - \sum_{c=1}^{d-1}\binom{d-1}{c-1}(r-1)^{c-1}\lambda^{k(r-1)} = \lambda^{k(r-1)}\left(r^{d-1} + (r-1)^{d-1} - \sum_{c=1}^{d}\binom{d-1}{c-1}(r-1)^{c-1}\right)$$

and the term in parentheses is precisely $(r-1)^{d-1}$. This gives the correct form for the rth element

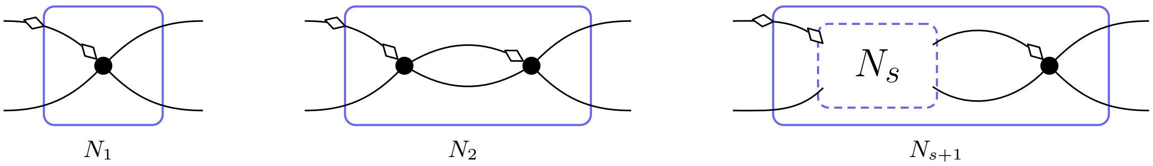


Figure 6.8: Recursive construction to interpolate g. The vertices are assigned f.

in column d of B'_k in M''.

Now we repair the columns in B'_0, also by induction on the number of columns. In the base case, if B'_0 only has one column, then there is nothing to prove, since this block has disappeared in M''. Otherwise, $(B'_0)_{r2} = r - (r-1) = 1$, so the second column is already in the correct form to be the first column in M'', and there is still nothing to prove. For the inductive step, assume that columns 2 to $d-1$ are in the correct form to be the first block in M'' for $d \geq 3$. That is, the entry at row $r \geq 2$ and column c from 2 through $d-1$ has the form $(B'_0)_{rc} = (r-1)^{c-2}$. The rth element in column d currently has the form $(B'_0)_{rd} = r^{d-1} - (r-1)^{d-1}$. Then we do the column operations

$$\begin{aligned}(B'_0)_{rd} - \sum_{c=2}^{d-1} \binom{d-1}{c-2} (B'_0)_{rc} &= r^{d-1} - (r-1)^{d-1} - \sum_{c=2}^{d-1} \binom{d-1}{c-2} (r-1)^{c-2} \\ &= (d-1)(r-1)^{d-2}\end{aligned}$$

and divide by $d-1$, which is nonzero, to get $(r-1)^{d-2}$. This is the correct form for the rth element in column d of B'_0 in M''. Therefore, we invoke our original induction hypothesis that the $(n-1)$-by-$(n-1)$ matrix M'' has a nonzero determinant, which completes the proof. □

Remark. The previous proof is similar to (but more complicated than) the derivation for the determinant of a Vandermonde matrix.

Lemma 6.5.4. *Let f be a signature of arity 4 with complex weights. If M_f is redundant and $\widetilde{M_f}$ is nonsingular, then* Pl-Holant(f) *is* #P*-hard.*

Proof. Let g be the arity 4 signature with M_g given in (6.5.3). We reduce from Pl-Holant$(\mathcal{F} \cup \{g\})$, which is #P-hard by Lemma 6.5.2.

Consider an instance Ω of Pl-Holant$(\mathcal{F} \cup \{g\})$. Suppose that g appears n times in Ω. We construct from Ω a sequence of instances Ω_s of Pl-Holant$(\mathcal{F})$ indexed by $s \geq 1$. We obtain Ω_s from Ω by replacing each occurrence of g with the gadget N_s in Figure 6.8 with f assigned to all vertices. In Ω_s, the edge corresponding to the ith significant index bit of N_s connects to the same location as the edge corresponding to the ith significant index bit of g in Ω.

Now to determine the relationship between Holant(Ω) and Holant(Ω_s), we use the isomorphism between redundant 4-by-4 matrices and 3-by-3 matrices. To obtain Ω_s from Ω, we effectively replace M_g with $M_{N_s} = (M_f)^s$, the sth power of the signature matrix M_f. By the Jordan normal form of $\widetilde{M_f}$, there exist $T, \Lambda \in \mathbb{C}^{3\times 3}$ such that

$$\widetilde{M_f} = T \Lambda T^{-1} = T \begin{bmatrix} \lambda_1 & b_1 & 0 \\ 0 & \lambda_2 & b_2 \\ 0 & 0 & \lambda_3 \end{bmatrix} T^{-1},$$

where $b_1, b_2 \in \{0, 1\}$. Note that $\lambda_1 \lambda_2 \lambda_3 = \det(\widetilde{M_f}) \neq 0$. Also since $\widetilde{M_g} = \varphi(M_g) = I_3$, and $T I_3 T^{-1} = I_3$, we have $\psi(T) M_g \psi(T^{-1}) = M_g$. We can view our construction of Ω_s as first replacing each M_g by $\psi(T) M_g \psi(T^{-1})$, which does not change the Holant value, and then replacing each new M_g with $\psi(\Lambda^s) = \psi(\Lambda)^s$ to obtain Ω_s. Observe that

$$\varphi(\psi(T)\psi(\Lambda^s)\psi(T^{-1})) = T\Lambda^s T^{-1} = (\widetilde{M_f})^s = (\varphi(M_f))^s = \varphi((M_f)^s),$$

hence, $\psi(T)\psi(\Lambda^s)\psi(T^{-1}) = M_{N_s}$. (Since $M_g = \psi(T) M_g \psi(T^{-1})$ and $M_{N_s} = \psi(T)\psi(\Lambda^s)\psi(T^{-1})$, replacing each M_g, sandwiched between $\psi(T)$ and $\psi(T^{-1})$, by $\psi(\Lambda^s)$ indeed transforms Ω to Ω_s. We also note that, by the isomorphism, $\psi(T^{-1})$ *is* the multiplicative inverse of $\psi(T)$ within the semi-group of redundant 4-by-4 matrices; but we prefer not to write it as $\psi(T)^{-1}$ since it is not the usual matrix inverse as a 4-by-4 matrix. Indeed, $\psi(T)$ is not invertible as a 4-by-4 matrix.)

In the case analysis below, we stratify the assignments in Ω_s based on the assignment to $\psi(\Lambda^s)$. The inputs to $\psi(\Lambda^s)$ are from $\{0, 1\}^2 \times \{0, 1\}^2$. However, we can combine the input 01 and 10, since $\psi(\Lambda^s)$ is redundant. Thus we actually stratify the assignments in Ω_s based on the assignment to Λ^s, which takes inputs from $\{0,1,2\} \times \{0,1,2\}$. In this compressed form, the row and column assignments to Λ^s are the Hamming weight of the two actual binary valued inputs to the uncompressed form $\psi(\Lambda^s)$.

Now we begin the case analysis on the values of b_1 and b_2.

1. Assume $b_1 = b_2 = 0$. We only need to consider the assignments to Λ^s that assign
   - $(0,0)$ i many times,
   - $(1,1)$ j many times, and
   - $(2,2)$ k many times

   since any other assignment contributes a factor of 0. Let c_{ijk} be the sum over all such assignments of the products of evaluations of all signatures in Ω_s except for Λ^s (including the contributions from T and T^{-1}). Note that this quantity is the same in Ω as in Ω_s. In particular it does not depend on s. Then

   $$\operatorname{Holant}(\Omega) = \sum_{i+j+k=n} \frac{c_{ijk}}{2^j}$$

   and the value of the Holant on Ω_s, for $s \geq 1$, is

   $$\operatorname{Holant}(\Omega_s) = \sum_{i+j+k=n} \left(\lambda_1^i \lambda_2^j \lambda_3^k\right)^s \left(\frac{c_{ijk}}{2^j}\right).$$

   The coefficient matrix is Vandermonde, but it may not have full rank because it might be that $\lambda_1^i \lambda_2^j \lambda_3^k = \lambda_1^{i'} \lambda_2^{j'} \lambda_3^{k'}$ for some $(i,j,k) \neq (i',j',k')$, where $i+j+k = i'+j'+k' = n$. However, this is not a problem since we are only interested in the sum $\sum \frac{c_{ijk}}{2^j}$. If two coefficients are the same, we replace their corresponding unknowns $c_{ijk}/2^j$ and $c_{i'j'k'}/2^{j'}$ with their sum as a new variable. After all such combinations, we have a Vandermonde system of full rank. In particular, none of the entries are 0 since $\lambda_1 \lambda_2 \lambda_3 = \det(\widetilde{M_f}) \neq 0$. Therefore, we can solve the linear system and obtain the value of $\operatorname{Holant}(\Omega)$.

2. Assume $b_1 \neq b_2$. We can permute the Jordan blocks in Λ so that $b_1 = 1$ and $b_2 = 0$, then $\lambda_1 = \lambda_2$, denoted by λ. We only need to consider the assignments to Λ^s that assign
   - $(0,0)$ i many times,

- $(1,1)$ j many times,
- $(2,2)$ k many times, and
- $(0,1)$ ℓ many times

since any other assignment contributes a factor of 0. Let $c_{ijk\ell}$ be the sum over all such assignments of the products of evaluations of all signatures in Ω_s except for Λ^s (including the contributions from T and T^{-1}). Then

$$\operatorname{Holant}(\Omega) = \sum_{i+j+k=n} \frac{c_{ijk0}}{2^j}$$

and the value of the Holant on Ω_s, for $s \geq 1$, is

$$\begin{aligned}\operatorname{Holant}(\Omega_s) &= \sum_{i+j+k+\ell=n} \lambda^{(i+j)s} \lambda_3^{ks} \left(s\lambda^{s-1}\right)^\ell \left(\frac{c_{ijk\ell}}{2^{j+\ell}}\right) \\ &= \lambda^{ns} \sum_{i+j+k+\ell=n} \left(\frac{\lambda_3}{\lambda}\right)^{ks} s^\ell \left(\frac{c_{ijk\ell}}{\lambda^\ell 2^{j+\ell}}\right).\end{aligned}$$

If λ_3/λ is a root of unity, then take a t such that $(\lambda_3/\lambda)^t = 1$. Then

$$\operatorname{Holant}(\Omega_{st}) = \lambda^{nst} \sum_{i+j+k+\ell=n} s^\ell \left(\frac{t^\ell c_{ijk\ell}}{\lambda^\ell 2^{j+\ell}}\right).$$

For $s \geq 1$, this gives a coefficient matrix that is Vandermonde. Although this system is not full rank, we can replace all the unknowns $c_{ijk\ell}/2^j$ having $i+j+k=n-\ell$ by their sum to form new unknowns $c'_\ell = \sum_{i+j+k=n-\ell} \frac{c_{ijk\ell}}{2^j}$, where $0 \leq \ell \leq n$. The new unknown c'_0 is the Holant of Ω that we seek. The resulting Vandermonde system

$$\operatorname{Holant}(\Omega_{st}) = \lambda^{nst} \sum_{\ell=0}^{n} s^\ell \left(\frac{t^\ell c'_\ell}{\lambda^\ell 2^\ell}\right)$$

has full rank, so we can solve for the new unknowns and obtain the value of $\operatorname{Holant}(\Omega) = c'_0$. If λ_3/λ is not a root of unity, then we replace all the unknowns $c_{ijk\ell}/(\lambda^\ell 2^{j+\ell})$ having $i+j=m$ with their sum to form new unknowns $c'_{mk\ell}$, for any $0 \leq m, k, \ell$ and $m+k+\ell=n$. The

Holant of Ω is now

$$\operatorname{Holant}(\Omega) = \sum_{m+k=n} c'_{mk0}$$

and the value of the Holant on Ω_s is

$$\begin{aligned}\operatorname{Holant}(\Omega_s) &= \lambda^{ns} \sum_{i+j+k+\ell=n} \left(\frac{\lambda_3}{\lambda}\right)^{ks} s^\ell \left(\frac{c_{ijk\ell}}{\lambda^\ell 2^{j+\ell}}\right) \\ &= \lambda^{ns} \sum_{m+k+\ell=n} \left(\frac{\lambda_3}{\lambda}\right)^{ks} s^\ell c'_{mk\ell}.\end{aligned}$$

After a suitable reordering of the columns, the matrix of coefficients satisfies the hypothesis of Lemma 6.5.3. Therefore, the linear system has full rank. We can solve for the unknowns and obtain the value of $\operatorname{Holant}(\Omega)$.

3. Assume $b_1 = b_2 = 1$. In this case, we have $\lambda_1 = \lambda_2 = \lambda_3$, denoted by λ, and we only need to consider the assignments to Λ^s that assign

   - $(0,0)$ or $(2,2)$ i many times,
   - $(1,1)$ j many times,
   - $(0,1)$ k many times,
   - $(1,2)$ ℓ many times, and
   - $(0,2)$ m many times

   since any other assignment contributes a factor of 0. Let $c_{ijk\ell m}$ be the sum over all such assignments of the products of evaluations of all signatures in Ω_s except for Λ^s (including the contributions from T and T^{-1}). Then

$$\operatorname{Holant}(\Omega) = \sum_{i+j=n} \frac{c_{ij000}}{2^j}$$

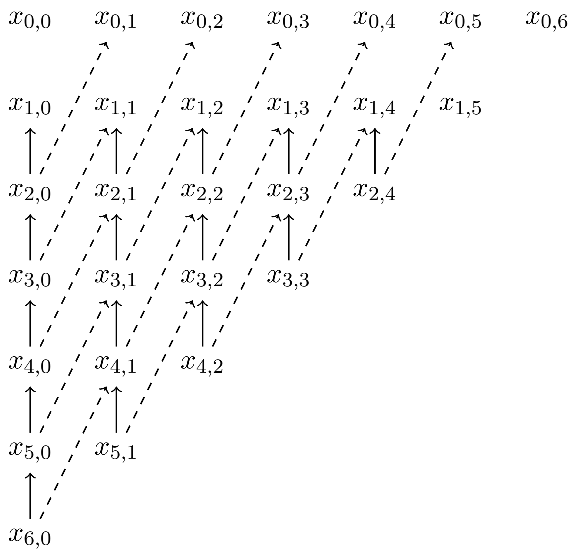


Figure 6.9: An example with $n = 6$ of how to apply update rules (6.5.5) (solid) and (6.5.6) (dashed).

and the value of the Holant on Ω_s, for $s \ge 1$, is

$$\begin{aligned}\text{Holant}(\Omega_s) &= \sum_{i+j+k+\ell+m=n} \lambda^{(i+j)s} \left(s\lambda^{s-1}\right)^{k+\ell} \left(s(s-1)\lambda^{s-2}\right)^m \left(\frac{c_{ijk\ell m}}{2^{j+k+m}}\right) \\ &= \lambda^{ns} \sum_{i+j+k+\ell+m=n} s^{k+\ell+m}(s-1)^m \left(\frac{c_{ijk\ell m}}{\lambda^{k+\ell+2m}2^{j+k+m}}\right).\end{aligned}$$

We replace all the unknowns $c_{ijk\ell m}/(\lambda^{k+\ell+2m}2^{j+k+m})$ having $i+j=p$ and $k+\ell=q$ with their sum to form new unknowns c'_{pqm}, for any $0 \le p, q, m$ and $p+q+m=n$. The Holant of Ω is now c'_{n00}. This new linear system is

$$\text{Holant}(\Omega_s) = \lambda^{ns} \sum_{p+q+m=n} s^{q+m}(s-1)^m c'_{pqm}$$

but is still rank deficient. We now index the columns by (q, m), where $q \ge 0$, $m \ge 0$, and $q+m \le n$. Correspondingly, we rename the variables $x_{q,m} = c'_{pqm}$. Note that $p = n-q-m$ is determined by (q, m). Observe that the column indexed by (q, m) is the sum of the columns indexed by $(q-1, m)$ and $(q-2, m+1)$ provided $q-2 \ge 0$. Namely, $s^{q+m}(s-1)^m = s^{q-1+m}(s-1)^m + s^{q-2+m+1}(s-1)^{m+1}$. Of course this is only meaningful if $q \ge 2$, $m \ge 0$ and

$q + m \leq n$. We write the linear system as

$$\sum_{q\geq 0,\ m\geq 0,\ q+m\leq n} \alpha_{q,m} x_{q,m} = \frac{\text{Holant}(\Omega_s)}{\lambda^{ns}},$$

where $\alpha_{q,m} = s^{q+m}(s-1)^m$ are the coefficients. Hence $\alpha_{q,m}x_{q,m} = \alpha_{q-1,m}x_{q,m} + \alpha_{q-2,m+1}x_{q,m}$, and we define new variables

$$x_{q-1,m} \leftarrow x_{q,m} + x_{q-1,m} \tag{6.5.5}$$

$$x_{q-2,m+1} \leftarrow x_{q,m} + x_{q-2,m+1} \tag{6.5.6}$$

from $q = n - m$ down to 2 for every $0 \leq m \leq n-2$. See Figure 6.9 for an example with $n = 6$.

Observe that in each update, the newly defined variables have a decreased index value for q. A more crucial observation is that the column indexed by $(0, 0)$ is never updated. This is because, in order to be an updated entry, there must be some $q \geq 2$ and $m \geq 0$ such that $(q-1, m) = (0, 0)$ or $(q-2, m+1) = (0, 0)$, which is clearly impossible. Hence $x_{0,0} = c'_{n00}$ is still the Holant value on Ω. The $2n + 1$ unknowns that remain are

$$x_{0,0},\ x_{1,0},\ x_{0,1},\ x_{1,1},\ x_{0,2},\ x_{1,2},\ \ldots,\ x_{0,n-1},\ x_{1,n-1},\ x_{0,n}$$

and their coefficients in row s are

$$1, s, s(s-1), s^2(s-1), s^2(s-1)^2, \ldots, s^{n-1}(s-1)^{n-1}, s^n(s-1)^{n-1}, s^n(s-1)^n.$$

It is clear that the κ-th entry in this row is a monic polynomial in s of degree κ, where $0 \leq \kappa \leq 2n$, and thus s^κ is a linear combination of the first κ entries. It follows that the coefficient matrix is a product of the standard Vandermonde matrix multiplied to its right by an upper triangular matrix with all 1's on the diagonal. Therefore, the linear system has full rank. We can solve for these final unknowns and obtain the value of $\text{Holant}(\Omega) = x_{0,0} = c'_{n00}$. □

Remark. Given such a long and difficult proof, I think some people might miss the basic idea that lead us to find it in the first place.

Given a symmetric signature of arity 4, we want to interpolate another signature of arity 4 using the linear construction in Figure 6.8. Why this construction? Because it is (probably) the simplest arity 4 recursive construction. And what is the worst that can happen? Every nonzero eigenvalue is a root of unity. Then this construction only provides us with a finite amount of information, but the amount of information that we need in order to interpolate grows polynomially with the input size.

And how bad is it? The signature matrix of a symmetric arity 4 signature has 0 for at least one eigenvalue. If the other three eigenvalues are 1 (and the corresponding eigenvectors are linearly independent), then the signature of this signature matrix is the one given in (6.5.3). Is this signature hard? Yes, but we did not know that until after going through this line of reasoning and *then* finding a proof that it is hard.

To simplify the application of Lemma 6.5.4 to a symmetric signature, we state and prove its restriction to a symmetric signature.

Corollary 6.5.5. *For a signature* $[f_0, f_1, f_2, f_3, f_4]$ *with complex weights, if there does* not *exist* $a, b, c \in \mathbb{C}$*, not all zero, such that* $af_k + bf_{k+1} + cf_{k+2} = 0$ *for all* $k \in \{0, 1, 2\}$*, then* Pl-Holant(f) *is* #P*-hard.*

Proof. If the compressed signature matrix $\widetilde{M_f}$ is nonsingular, then Pl-Holant(f) is #P-hard by Lemma 6.5.4, so assume that the rank of $\widetilde{M_f}$ is at most 2. Then we have

$$a' \begin{pmatrix} f_0 \\ f_1 \\ f_2 \end{pmatrix} + 2b' \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} + c' \begin{pmatrix} f_2 \\ f_3 \\ f_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

for some $a', b', c' \in \mathbb{C}$, not all zero. Thus, $a = a'$, $b = 2b'$, and $c = c'$ have the desired property. □

To make Lemma 6.5.4 more applicable, we show that for an arity 4 signature f, the redundancy of M_f and the nonsingularity of $\widetilde{M_f}$ are invariant under an invertible holographic transformation.

Lemma 6.5.6. *Let f be a signature of arity 4 with complex weights, $T \in \mathbb{C}^{2\times 2}$ a matrix, and $\hat{f} = T^{\otimes 4} f$. If M_f is redundant, then $M_{\hat{f}}$ is also redundant and* $\det(\varphi(M_{\hat{f}})) = \det(\varphi(M_f))\det(T)^6$.

Proof. Since $\hat{f} = T^{\otimes 4} f$, we can express $M_{\hat{f}}$ in terms of M_f and T as

$$M_{\hat{f}} = T^{\otimes 2} M_f \left(T^{\mathsf{T}}\right)^{\otimes 2}. \tag{6.5.7}$$

This can be directly checked. Alternatively, this relation is known (and can also be directly checked) had we not introduced the flip of the middle two columns, i.e., if the columns were ordered $00, 01, 10, 11$ by the last two bits in f and $\hat{f}$. Instead, the columns are ordered by $00, 10, 01, 11$ in M_f and $M_{\hat{f}}$. Let $T = (t^i_j)$, where row index i and column index j range from $\{0, 1\}$. Then $T^{\otimes 2} = (t^i_j t^{i'}_{j'})$, with row index ii' and column index jj'. Let

$$\mathcal{E} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

then $\mathcal{E} T^{\otimes 2} \mathcal{E} = T^{\otimes 2}$, i.e., a simultaneous row flip $ii' \leftrightarrow i'i$ and column flip $jj' \leftrightarrow j'j$ keep $T^{\otimes 2}$ unchanged. Then the known relations $M_{\hat{f}} \mathcal{E} = T^{\otimes 2} M_f \mathcal{E} \left(T^{\mathsf{T}}\right)^{\otimes 2}$ and $\mathcal{E} \left(T^{\mathsf{T}}\right)^{\otimes 2} \mathcal{E} = \left(T^{\mathsf{T}}\right)^{\otimes 2}$ imply (6.5.7).

Now $X \in \mathrm{RM}_4(\mathbb{C})$ iff $\mathcal{E} X = X = X \mathcal{E}$. Then it follows that $M_{\hat{f}} \in \mathrm{RM}_4(\mathbb{C})$ if $M_f \in \mathrm{RM}_4(\mathbb{C})$. For the two matrices A and B in the definition of φ, we note that $BA = M_g$, where M_g given in (6.5.3) is the identity element of the semi-group $\mathrm{RM}_4(\mathbb{C})$. Since $M_f \in \mathrm{RM}_4(\mathbb{C})$, we have $BAM_f = M_f = M_f BA$. Then we have

$$\begin{aligned} \varphi(M_{\hat{f}}) &= A M_{\hat{f}} B = A \left(T^{\otimes 2} M_f \left(T^{\mathsf{T}}\right)^{\otimes 2}\right) B \\ &= (A T^{\otimes 2} B)(A M_f B)(A \left(T^{\mathsf{T}}\right)^{\otimes 2} B) \\ &= \varphi(T^{\otimes 2}) \varphi(M_f) \varphi(\left(T^{\mathsf{T}}\right)^{\otimes 2}). \end{aligned} \tag{6.5.8}$$

Another direct calculation shows that

$$\det(\varphi(T^{\otimes 2})) = \det(T)^3 = \det(\varphi((T^{\intercal})^{\otimes 2})).$$

Thus, by applying determinant to both sides of (6.5.8), we have

$$\det(\varphi(M_{\hat{f}})) = \det(\varphi(M_f)) \det(T)^6$$

as claimed. □

In particular, for a matrix $T \in \mathbf{GL}_2(\mathbb{C})$, M_f is redundant and $\widetilde{M_f}$ is nonsingular iff $M_{\hat{f}}$ is redundant and $\widetilde{M_{\hat{f}}}$ is nonsingular. By combining Lemma 6.5.4 and Lemma 6.5.6, we have the following corollary.

Corollary 6.5.7. *Let f be a signature of arity 4 with complex weights. If there exists a matrix $T \in \mathbf{GL}_2(\mathbb{C})$ such that $\hat{f} = T^{\otimes 4} f$ with $M_{\hat{f}}$ is redundant and $\widetilde{M_{\hat{f}}}$ is nonsingular, then* Pl-Holant(f) *is #*P*-hard.*

6.6 Main Result

Before proving our main result for this chapter, we show hardness for a planar tractable case over general graphs. To prove both this result and our main result, we use orthogonal holographic transformations to normalize a class of signatures. We state the general case for symmetric signatures of arity $n \geq 1$ even though we only use the case $n = 4$. Appendix D of [35] considered the case $n = 3$.

Lemma 6.6.1. *Let $f = [f_0, \ldots, f_n]$ be a complex-weighted signature of arity $n \geq 1$. Suppose there exists $c, d \in \mathbb{C}$ such that $f_k = ck\alpha^{k-1} + d\alpha^k$ for all $1 \leq k \leq n$. If $c \neq 0$ and $\alpha \neq \pm i$, then there exists $H \in \mathbf{O}_2(\mathbb{C})$ such that $H^{\otimes n} f = [x, y, 0, \ldots, 0]$ for some $x, y \in \mathbb{C}$ with $y \neq 0$.*

Proof. Let $S = \begin{bmatrix} 1 & \frac{d-1}{n} \\ \alpha & c + \frac{d-1}{n}\alpha \end{bmatrix}$. Note that $\det S = c \neq 0$. Then we can write f as

$$f = S^{\otimes n} [1, 1, 0, \ldots, 0]^{\intercal}.$$

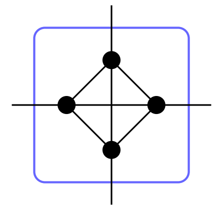


Figure 6.10: The tetrahedron gadget. Each vertex is assigned $[0, 1, 0, 0, 0]$.

This identity can be verified by observing that

$$[1, 1, 0, \ldots, 0]^{\intercal} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes n} + \frac{1}{(n-1)!} \operatorname{Sym}_n^{n-1} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} ; \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right).$$

We consider the value at any entry of Hamming weight k. By considering where the tensor product factor $[0, 1]$ is located among the n possible locations, we get

$$\alpha^k + k\left(c + \frac{d-1}{n}\alpha\right)\alpha^{k-1} + (n-k)\frac{d-1}{n}\alpha^k = ck\alpha^{k-1} + d\alpha^k.$$

Let $H = \frac{1}{\sqrt{1+\alpha^2}} \left[\begin{smallmatrix} 1 & \alpha \\ \alpha & -1 \end{smallmatrix}\right]$, then $H = H^{\intercal} = H^{-1} \in \mathbf{O}_2(\mathbb{C})$ is orthogonal, and $R = HS = \left[\begin{smallmatrix} u & w \\ 0 & v \end{smallmatrix}\right]$ is upper triangular, where $v, w \in \mathbb{C}$ and $u = \sqrt{1+\alpha^2} \neq 0$. Also, $\det R = \det H \det S = (-1)c \neq 0$, so we have $v \neq 0$. Then it follows that

$$\begin{aligned} H^{\otimes n} f &= (HS)^{\otimes n}[1, 1, 0, \ldots, 0] \\ &= R^{\otimes n}[1, 1, 0, \ldots, 0] \\ &= R^{\otimes n}\left([1, 0]^{\otimes n} + \frac{1}{(n-1)!}\operatorname{Sym}_n^{n-1}([1, 0]; [0, 1])\right) \\ &= [u, 0]^{\otimes n} + \frac{1}{(n-1)!}\operatorname{Sym}_n^{n-1}([u, 0]; [w, v]) \\ &= [u^n + nu^{n-1}w, u^{n-1}v, 0, \ldots, 0] \end{aligned}$$

has the required form since $u^{n-1}v \neq 0$. □

Lemma 6.6.2. $\operatorname{Holant}([0, 1, 0, 0, 0])$ *is* #P*-hard.*

Proof. Consider the gadget in Figure 6.10. We assign $[0, 1, 0, 0, 0]$) to each vertex. The signature f of this gadget is $f = [3, 0, 1, 0, 1]$. The compressed signature matrix of this gadget is nonsingular since $\det(\widetilde{M_f}) = 4$, so we are done by Lemma 6.5.4. ☐

Now we prove our main result, a planar Holant dichotomy for a symmetric signature of arity 4.

Theorem 6.6.3. *If f is a non-degenerate, symmetric, complex-valued signature of arity 4 in Boolean variables, then* $\operatorname{Holant}(f)$ *is* #P*-hard unless f is $\mathscr{A}$-transformable or $\mathscr{P}$-transformable or vanishing, in which case, the problem is computable in polynomial time. If the input is restricted to planar graphs, then the problem remains* #P*-hard unless $\mathscr{M}$-transformable, in which case, the problem is computable in polynomial time.*

Proof. Let $f = [f_0, f_1, f_2, f_3, f_4]$. If there do not exist $a, b, c \in \mathbb{C}$, not all zero, such that for all $k \in \{0, 1, 2\}$, $af_k + bf_{k+1} + cf_{k+2} = 0$, then Pl-Holant($f$) is #P-hard by Corollary 6.5.5. Otherwise, there do exist such a, b, c. If $a = c = 0$, then $b \neq 0$, so $f_1 = f_2 = f_3 = 0$. In this case, $f \in \mathscr{P}$ is a generalized equality signature, so f is $\mathscr{P}$-transformable. Now suppose a and c are not both 0. If $b^2 - 4ac \neq 0$, then there exist $\alpha_1, \alpha_2, \beta_1, \beta_1 \in \mathbb{C}$ such that $f_k = \alpha_1^{4-k}\alpha_2^k + \beta_1^{4-k}\beta_2^k$, where $\alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$. A holographic transformation by $\left[\begin{smallmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{smallmatrix}\right]$ transforms f to $=_4$, and we are done by Theorem 6.1.2. Otherwise, $b^2 - 4ac = 0$ and there are two cases. In the first, $f_k = ck\alpha^{k-1} + d\alpha^k$ for any $0 \leq k \leq 2$, where $c \neq 0$. In the second, $f_k = c(4-k)\alpha^{3-k} + d\alpha^{4-k}$ for any $0 \leq k \leq 2$, where $c \neq 0$. These cases map between each other under a holographic transformation by $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, so assume that we are in the first case. If $\alpha = \pm i$, then f is vanishing. Otherwise, there exists an orthogonal matrix H such that $H^{\otimes n} f = [x, y, 0, 0, 0]$ with $y \neq 0$ by Lemma 6.6.1. Then up to a nonzero factor of y, we have $\hat{f} = [v, 1, 0, 0, 0]$, where $v = \frac{x}{y}$. If $v \neq 0$, then Pl-Holant($[v, 1, 0, 0, 0]$) is #P-hard by Corollary 6.3.6. Otherwise $v = 0$ and the problem is counting perfect matchings over 4-regular graphs. Thus $\hat{f} \in \mathscr{M}$, so f is $\mathscr{M}$-transformable, and the problem is tractable by Corollary 4.3.7. Without restricting to planar graphs, Holant($[0, 1, 0, 0, 0]$) is #P-hard by Lemma 6.6.2. ☐

6.7 Closing Thoughts

Superiority of planar pairing Instead of using a planar pairing as was done in the proof of Lemma 6.3.5, there is another approach that one can try. After explaining this approach though, I will argue why I believe that planar pairing is a superior proof technique.

The general situation to which planar pairing applies is that every vertex has a dangling edge and these dangling edges need to be connected together in pairs. To ensure that there are an even number of dangling edges, we can duplicate the graph [46, Lemma IV.2]. Sometimes multiple copies are required [81, Lemma 5.4]. In any case though, we solve the resulting problem using an oracle and want to obtain the solution of the original problem, which is some root of the oracle's answer. The question is, which root is the correct one?

The approach used in the works cited above is to

> observe that in all the #P-hardness proofs in the current paper and all the #P-hardness proofs in all the papers on which the current paper depends, when the target Holant problem for any signature set is proved to be #P-hard, the reduction is from a #P-complete problem (such as counting vertex covers, matchings, perfect matchings, or Eulerian orientations), so the solution is a nonnegative integer, and the only three techniques used in the reduction chains are direct gadget constructions, polynomial interpolations, and holographic transformations.

The proofs then continue and argue why the unique nonnegative root can be recovered when using each of these reduction techniques.

The issue with this approach is the difficulty in checking its correctness. One must read all the #P-hardness proofs in the current paper and all the #P-hardness proofs in all the papers on which the current paper depends in order to determine if the reductions really are structured as claimed. This breaks the modularity intended by the use of lemmas, theorems, and the like.[1] In contrast, the technique of planar pairing only requires one to check the correctness of the proof in which it is used.

[1] This type of argument is also used in the proof of Corollary B.5 in [27], but in this instance, the argument is self contained.

When the reduction must be planar, the dangling edges must be connected in a planar way. There are two equally-good ways to accomplish this. In a planar pairing, there must be an even number of the dangling edges. Then one is able to efficiently pair them up in a planar way. Alternatively, one can “move” all of the dangling edges to the outer face in some planar embedding and then pair them up. Moving the dangling edges to the outer face requires the ability to construct a gadget that functions like as a crossover gadget and allows the dangling edges to move closer to the outer face. See the proof of Lemma IV.2 in [46] for an example.

Chapter 7

Dichotomy for Holant Problems over General Graphs

We prove a dichotomy theorem for Holant problems over an arbitrary set of complex-valued symmetric constraint functions $\mathcal{F}$ on Boolean variables. The dichotomy theorem has an explicit tractability criterion expressible in terms of holographic transformations. A Holant problem defined by a set of constraint functions $\mathcal{F}$ is solvable in polynomial time if it satisfies this tractability criterion, and is #P-hard otherwise.

The tractability criterion can be intuitively stated as follows: A set $\mathcal{F}$ is tractable if (1) every function in $\mathcal{F}$ has arity at most two, or (2) $\mathcal{F}$ is transformable to an affine type, or (3) $\mathcal{F}$ is transformable to a product type, or (4) $\mathcal{F}$ is vanishing, combined with the right type of binary functions, or (5) $\mathcal{F}$ is highly vanishing, combined with any unary functions. The proof of this theorem utilizes previous dichotomy theorems for Holant problems and #CSPs. Holographic transformations play an indispensable role as both a proof technique and in the statement of the tractability criterion. This work was published in [29, 30].

7.1 Background

We prove a dichotomy theorem for Holant problems over an arbitrary set of complex-valued symmetric constraint functions $\mathcal{F}$ on Boolean variables. Compared to previous dichotomy theorems for Holant problems, a significant difficultly is the new tractable case of vanishing signatures.

Vanishing signatures are constraint functions, that when applied to any signature grid, produce a zero Holant value. A simple example is a constraint function of the form $(1, i)^{\otimes k}$ on k variables, which is a tensor product of $(1, i)$. This function on a vertex (of degree k) can be replaced by k copies of the unary function $(1, i)$ on k new vertices, each connected to an incident edge. Whenever two copies of $(1, i)$ meet in the Holant sum, they annihilate each other since they give the value $(1, i) \cdot (1, i) = 0$.

These ghostly constraint functions are like the elusive dark matter. They do not actually contribute any value to the Holant sum. However, in order to give a complete dichotomy for Holant problems, it turns out to be essential that we capture these vanishing signatures. There is another similarity with dark matter. Their contribution to the Holant sum is not directly observed. Yet in terms of the dimension of the algebraic variety they constitute, they make up the vast majority of the tractable symmetric signatures. Furthermore, when combined with others, they provide a large substrate to produce non-vanishing yet tractable signatures. In #CSP, they are invisible due to the presumed inclusion of all the Equality functions; and they lurk beneath the surface when one only considers real-valued Holant problems.

The existence of vanishing signatures have influenced previous dichotomy results, although this influence was not fully recognized at the time. In the dichotomy theorems in [47] and [35], almost all tractable signatures can be transformed into a tractable #CSP problem, except for one special category. The tractability proof for this category used the fact that they are a special case of generalized Fibonacci signatures [52]. However, what went completely unnoticed is that for every input instance using such signatures alone, the Holant value is always zero!

The most significant previous encounter with vanishing signatures was in the parity setting [72]. The authors noticed that a large fraction of signatures always induce an even Holant value, which is vanishing in $\mathbb{Z}_2$. However, the parity dichotomy was achieved using an existential argument without obtaining a complete characterization of the vanishing signatures. Consequently, the dichotomy criterion is non-constructive and is currently not known to be decidable. Nevertheless, this work is important because it was the first to discover nontrivial vanishing signatures in the parity setting and to obtain a dichotomy that was *completed* by vanishing signatures.

We characterize the set of $\mathscr{A}$- and $\mathscr{P}$-transformable signatures after an orthogonal holographic transformation. An orthogonal transformation is natural since the binary EQUALITY $=_2$ is unchanged under such a transformation. With explicit characterizations of these tractable signatures, a complete dichotomy theorem becomes possible. We first prove a dichotomy for a single signature, and then we extend it to an arbitrary set of signatures. The most difficult part is proving a dichotomy for a single signature of arity 4, which we did in Chapter 6 and is given in Theorem 6.6.3.

With this dichotomy, we come to a conclusion on a long series of dichotomies on Holant problems [50, 47, 48, 91, 93, 37, 36, 35, 80]. In particular, this dichotomy extends the dichotomy in [80] that covers all real-valued symmetric signatures. While we do not rely on their real-valued dichotomy itself, we do make important use of one result in [80]: a dichotomy for $\#\mathrm{CSP}^d$. Recall that $\mathcal{T}_k = \left\{ \left[\begin{smallmatrix} 1 & 0 \\ 0 & \omega \end{smallmatrix}\right] \in \mathbb{C}^{2\times 2} \mid \omega^k = 1 \right\}$.

Theorem 7.1.1 (Theorem IV.1 in [80])**.** *Let $d \geq 1$ be an integer. Suppose $\mathcal{F}$ is a set of symmetric, complex-valued signatures in Boolean variables. Then $\#\mathrm{CSP}^d(\mathcal{F})$ is #P-hard unless $\mathcal{F} \subseteq \mathscr{P}$ or $\exists T \in \mathcal{T}_{4d}$ such that $\mathcal{F} \subseteq T\mathscr{A}$, in which case the problem is computable in polynomial time.*

7.2 Statement of Main Result

Here is the main result of this chapter. The sets $\mathscr{V}^\sigma$ and $\mathscr{R}_2^\sigma$ are defined in Section 4.4.

Theorem 7.2.1. *Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures in Boolean variables. Then* Holant$(\mathcal{F})$ *is #P-hard unless $\mathcal{F}$ satisfies one of the following conditions, in which case the problem is computable in polynomial time:*

1. *All non-degenerate signatures in $\mathcal{F}$ are of arity at most 2;*
2. *$\mathcal{F}$ is $\mathscr{A}$-transformable;*
3. *$\mathcal{F}$ is $\mathscr{P}$-transformable;*

4. $\mathcal{F} \subseteq \mathscr{V}^\sigma \cup \{f \in \mathscr{R}_2^\sigma \mid \operatorname{arity}(f) = 2\}$ *for* $\sigma \in \{+, -\}$;

5. *All non-degenerate signatures in* $\mathcal{F}$ *are in* $\mathscr{R}_2^\sigma$ *for* $\sigma \in \{+, -\}$.

Note that any signature in $\mathscr{R}_2^\sigma$ having arity at least 3 is a vanishing signature. Thus all signatures of arity at least 3 in case 5 are vanishing. While both cases 4 and 5 are largely concerned with vanishing signatures, these two cases differ. In case 4, all signatures in $\mathcal{F}$, including unary signatures but excluding binary signatures, must be vanishing of a single type σ; the binary signatures are only required to be in $\mathscr{R}_2^\sigma$. In contrast, case 5 has no requirement placed on degenerate signatures which include all unary signatures. Then all non-degenerate binary signatures are required to be in $\mathscr{R}_2^\sigma$. Finally all non-degenerate signatures of arity at least 3 are also required to be in $\mathscr{R}_2^\sigma$, which is a strong form of vanishing; they must have a large vanishing degree of type σ.

Case 5 is actually a known tractable case [52, 49]. Every signature (after replacing all degenerate signatures with corresponding ones) is a generalized Fibonacci signature with $m = \sigma 2i$, which means that every signature $[f_0, f_1, \ldots, f_n] \in \mathcal{F}$ satisfies $f_{k+2} = m f_{k+1} + f_k$ for $0 \le k \le n-2$. However, we present a unified proof of tractability based on vanishing signatures.

7.2.1 Proof of Tractability

For any signature grid Ω, $\operatorname{Holant}(\Omega)$ is the product of the Holant on each connected component, so we only need to compute over connected components.

For case 1, after decomposing all degenerate signatures into unary ones, a connected component of the graph is either a path or a cycle and the Holant can be computed using matrix product and trace. Case 2 is tractable by Corollary 4.2.7. Case 3 is tractable by Corollary 4.1.6. For case 4, any binary signature $g \in \mathscr{R}_2^\sigma$ has $\operatorname{rd}^\sigma(g) \le 1$, and thus $\operatorname{vd}^\sigma(g) \ge 1 = \operatorname{arity}(g)/2$. Any signature $f \in \mathscr{V}^\sigma$ has $\operatorname{vd}^\sigma(f) > \operatorname{arity}(f)/2$. If $\mathcal{F}$ contains a signature f of arity at least 3, then it must belong to $\mathscr{V}^\sigma$. Then by the combinatorial view, more than half of the unary signatures are $[1, \sigma i]$, so $\operatorname{Holant}(\Omega)$ vanishes. On the other hand, if the arity of every signature in $\mathcal{F}$ is at most 2, then we have reduced to case 1.

Now consider case 5. After decomposing all degenerate signatures into unary ones, recursively absorb any unary signature into its neighboring signature. If it is connected to another unary

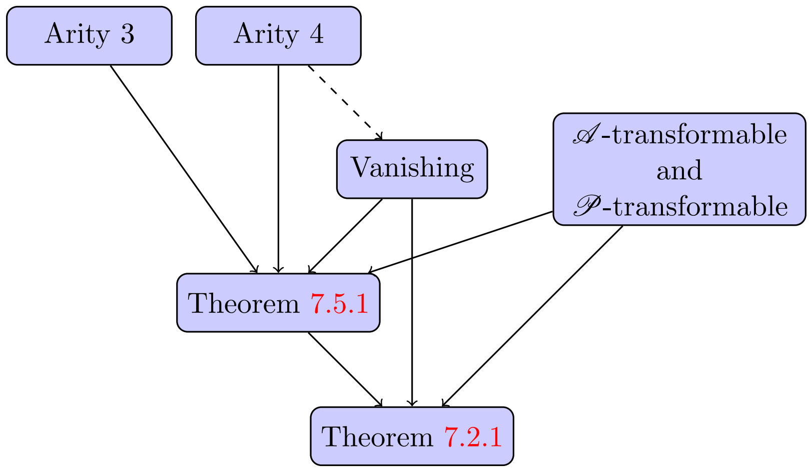


Figure 7.1: Dependency graph of key hardness results for our main dichotomy, Theorem 7.2.1. The dashed edge indicates a dependency in terms of techniques rather than the result itself. "Arity 3(4)" stands for the arity 3(4) single signature dichotomy. "Vanishing" ("$\mathscr{A}$-transformable and $\mathscr{P}$-transformable") stands for the lemmas regarding vanishing ($\mathscr{A}$-transformable and $\mathscr{P}$-transformable) signatures. Not all dependencies on previous dichotomy theorems are shown.

signature, then this produces a global constant factor. If it is connected to a binary signature, then this creates another unary signature. We observe that if $f \in \mathscr{R}_2^\sigma$ has $\operatorname{arity}(f) \geq 2$, then for any unary signature u, after connecting f to u, the signature $\langle f, u \rangle$ still belongs to $\mathscr{R}_2^\sigma$. Hence after recursively absorbing all unary signatures in the above process, we still have a signature grid where all signatures belong to $\mathscr{R}_2^\sigma$. Any remaining signature f that has arity at least 3 belongs to $\mathscr{V}^\sigma$ since $\operatorname{rd}^\sigma(f) \leq 1$ and thus $\operatorname{vd}^\sigma(f) \geq \operatorname{arity}(f) - 1 > \operatorname{arity}(f)/2$. Thus we have reduced to case 4.

7.2.2 Outline of Hardness Proof

The hardness proof of our main dichotomy is more complicated. Our first goal is to prove a dichotomy for a single signature, Theorem 7.5.1. The proof is by induction on the arity of the signature. The induction is done by taking a self loop, which causes the arity to go down by 2. Thus, we need two base cases, a dichotomy for an arity 3 signature and a dichotomy for an arity 4 signature. The dichotomy for an arity 3 is given in Theorem 6.1.3. The dichotomy for an arity 4 was proved in Chapter 6 and is given in Theorem 6.6.3. It is a crucial ingredient in our proof of the full dichotomy. It is not only a base case of the single signature dichotomy but also utilized several times in the inductive step.

We begin our proof by considering what signatures mix with the vanishing signatures to give #P-hardness. When adding unary or binary signatures, the only possible combinations that maintain the tractability of the vanishing signatures are as described in cases 4 and 5 in Theorem 7.2.1. Moreover, combining two vanishing signatures of the opposite type of arity at least 3 implies #P-hardness. The proof of this last statement uses techniques that are similar to those in the proof of the arity 4 dichotomy.

Another important piece of the proof is to understand the signatures that are $\mathscr{A}$-transformable or $\mathscr{P}$-transformable. We obtain new explicit characterizations of these signatures. We use these characterizations to prove dichotomy theorems for any signature set containing an $\mathscr{A}$- or $\mathscr{P}$-transformable signature. Unless every signature in the set is $\mathscr{A}$- or $\mathscr{P}$-transformable, the problem is #P-hard. The proofs of these dichotomy theorems utilize the $\#\mathrm{CSP}^d$ dichotomy.

The main dichotomy, Theorem 7.2.1, depends on Theorem 7.5.1 and the results regarding vanishing signatures as well as the dichotomies when $\mathscr{A}$- or $\mathscr{P}$-transformable signatures appear. Figure 7.1 summarizes the dependencies among these results.

7.3 Mixing with Vanishing Signatures

We know that vanishing signatures, which were characterized in Section 4.4, are tractable. Now we determine what signatures combine with them to give #P-hardness. We begin with unary signatures and their tensor powers.

Lemma 7.3.1. *Let $f \in \mathscr{V}^\sigma$ be a symmetric signature with $\mathrm{rd}^\sigma(f) \geq 2$ where $\sigma \in \{+,-\}$. Suppose $v = u^{\otimes m}$ is a symmetric degenerate signature for some unary signature u and some integer $m \geq 1$. If u is not a multiple of $[1, \sigma i]$, then $\mathrm{Holant}(\{f, v\})$ is #P-hard.*

Proof. We consider $\sigma = +$ since the other case is similar. Since $f \in \mathscr{V}^+$, we have $\mathrm{arity}(f) > 2\,\mathrm{rd}^+(f) \geq 4$, and $\mathrm{vd}^+(f) = \mathrm{arity}(f) - \mathrm{rd}^+(f) > 0$. As $\mathrm{rd}^+(f) \geq 2$, f is a nonzero signature. By Lemma 4.4.14, with zero or more self loops of f, we can construct some f' with $\mathrm{rd}^+(f') = 2$ and arity $n \geq 5$. We can repeatedly apply Lemma 4.4.14, since in each step we reduce the recurrence degree rd^+ by exactly one, which remains positive and thus the resulting signature is nonzero. Being obtained from f by self loops, it remains in $\mathscr{V}^+$. The process can be continued. After two more self loops,

Let $t = \gcd(m, n-4)$. There are integers x and y such that $xm + y(n-4) = t$. By replacing x with $x + z(n-4)$ and y with $y - zm$, for any integer z, we may assume $x > 0$ and $y < 0$. Then if we connect $|y|$ copies of $[1,i]^{\otimes(n-4)}$ to x copies of $v = u^{\otimes m}$, we can realize $u^{\otimes t}$. Since u is not a multiple of $[1,i]$, it follows that $\langle u, [1,i]\rangle$ is a nonzero constant. We can realize $g = u^{\otimes(n-4)}$ by putting $(n-4)/t$ many copies of $u^{\otimes t}$ together.

Now connect this g back to f'. Since the unary u is not a multiple of $[1,i]$, we can directly verify that $g \notin \mathscr{R}^+_{n-4}$ and thus $\mathrm{rd}^+(g) = \mathrm{arity}(g) = n-4$, and $\mathrm{vd}^+(g) = 0$. By Lemma 4.4.13, we get $f'' = \langle f', g\rangle$ of arity 4 and $\mathrm{rd}^+(f'') = 2$. One can verify that $\operatorname{Holant}(f'')$ is #P-hard by Lemma 6.5.4, by writing $f''_k = i^k p(k)$ for some polynomial p of degree exactly 2. A more revealing proof of the #P-hardness of $\operatorname{Holant}(f'')$ is by noticing that this is the problem $\operatorname{Holant}(=_2 \mid f'')$, which is equivalent to $\operatorname{Holant}(\neq_2 \mid \widehat{f''})$ under the holographic transformation $Z = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$. By $\mathrm{rd}^+(f'') = 2$, the signature $\widehat{f''}$ takes the form $[\widehat{f''}_0, \widehat{f''}_1, \widehat{f''}_2, 0, 0]$, where $\widehat{f''}_2 \neq 0$. Then we are done by Corollary 6.5.7 after transforming back to the original setting. □

Next we consider binary signatures, but first we prove a simple interpolate result.

Lemma 7.3.2. *Let $x \in \mathbb{C}$. If $x \neq 0$, then for any set $\mathcal{F}$ containing $[x,1,0]$, we have*

$$\operatorname{Holant}(\neq_2 \mid \mathcal{F} \cup \{[v,1,0]\}) \leq_T \operatorname{Holant}(\neq_2 \mid \mathcal{F})$$

for any $v \in \mathbb{C}$.

Proof. Consider an instance Ω of $\operatorname{Holant}(\neq_2 \mid \mathcal{F} \cup \{[v,1,0]\})$. Suppose that $[v,1,0]$ appears n times in Ω. We stratify the assignments in Ω based on the assignments to $[v,1,0]$. We only need to consider assignments of Hamming weight 0 and 1 since an assignment of Hamming weight 2 contributes a factor of 0. Let i be the number of Hamming weight 0 assignments to $[v,1,0]$ in Ω.

Then there are $n - i$ assignments of Hamming weight 1 and the Holant on Ω is

$$\text{Holant}(\Omega) = \sum_{i=0}^{n} v^i c_i,$$

where c_i is the sum over all such assignments of the product of evaluations of all other signatures on Ω.

We construct from Ω a sequence of instances Ω_s of Holant($\mathcal{F}$) indexed by $s \geq 1$. We obtain Ω_s from Ω by replacing each occurrence of $[v, 1, 0]$ with a gadget g_s created from s copies of $[x, 1, 0]$, connected sequentially but with $(\neq_2) = [0, 1, 0]$ between each sequential pair. The signature of g_s is $[sx, 1, 0]$, which can be verified by the matrix product

$$\left(\begin{bmatrix} x & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right)^{s-1} \begin{bmatrix} x & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^{s-1} \begin{bmatrix} x & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & (s-1)x \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} sx & 1 \\ 1 & 0 \end{bmatrix}.$$

The Holant on Ω_s is

$$\text{Holant}(\Omega_s) = \sum_{i=0}^{n} (sx)^i c_i.$$

For $s \geq 1$, this gives a coefficient matrix that is Vandermonde. Since x is nonzero, sx is distinct for each s. Therefore, the Vandermonde system has full rank. We can solve for the unknowns c_i and obtain the value of Holant(Ω). ☐

Lemma 7.3.3. *Let $f \in \mathscr{V}^\sigma$ be a symmetric non-degenerate signature where $\sigma \in \{+, -\}$. Suppose g is a non-degenerate binary signature. If $g \notin \mathscr{R}_2^\sigma$, then* Holant($\{f, g\}$) *is* #P*-hard.*

Proof. We consider $\sigma = +$ since the other case is similar. A unary signature is degenerate. If f is binary, then $\text{vd}^+(f) > 1$. Hence $\text{vd}^+(f) \geq 2$, and so f is degenerate. Since f is non-degenerate, $\text{arity}(f) \geq 3$.

We prove the lemma by induction on the arity of f. There are two base cases, $\text{arity}(f) = 3$ and $\text{arity}(f) = 4$. However, the arity 3 case is easily reduced to the arity 4 case. We show this first, and then show that the lemma holds in the arity 4 case.

Assume $\text{arity}(f) = 3$. Since $f \in \mathscr{V}^+$, we have $\text{rd}^+(f) < 3/2$, thus $f \in \mathscr{R}_2^+$. From $\text{rd}^+(f) \leq 1$,

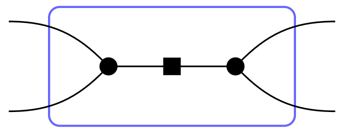


Figure 7.2: The circles are assigned $[t, 1, 0, 0]$ and the square is assigned $\neq_2$.

we have $\operatorname{vd}^+(f) \geq 2$. On the other hand, f is non-degenerate, so $\operatorname{vd}^+(f) < 3$. Thus $\operatorname{vd}^+(f) = 2$ and $\operatorname{rd}^+(f) = 1$.

We connect two copies of f together by one edge to get an arity 4 signature f'. By construction, it may not appear that f' is a symmetric signature. However, we show that f' is in fact *symmetric*, non-degenerate, and vanishing. It is clearly a vanishing signature, since f is vanishing. Consider the Z transformation, under which f is transformed into $\hat{f} = [t, 1, 0, 0]$ for some t up to a nonzero constant. The $=_2$ on the connecting edge between the two copies of f is transformed into $\neq_2$. In the bipartite setting, our construction is the same as the gadget in Figure 7.2. One can verify that the resulting signature is $\hat{f}' = [2t, 1, 0, 0, 0]$. The crucial observation is that it takes the same value 0 on inputs 1010 and 1100, where the left two bits are input to one copy of f and the right two bits are for another. The corresponding signature f' is non-degenerate with $\operatorname{rd}^+(f') = 1$ and vanishing.

Next we consider the base case of $\operatorname{arity}(f) = 4$. Since $f \in \mathscr{V}^+$, we have $\operatorname{vd}^+(f) > 2$ and $\operatorname{rd}^+(f) < 2$. Since f is non-degenerate, we have $\operatorname{rd}^+(f) \neq -1, 0$. Hence $\operatorname{rd}^+(f) = 1$ and $\operatorname{vd}^+(f) = 3$. Also by assumption, the given binary $g \notin \mathscr{R}_2^+$, so we have $\operatorname{rd}^+(g) = 2$. Once again, consider the holographic transformation by Z. This gives

$$\begin{aligned}\operatorname{Holant}(=_2 \mid \{f, g\}) &\equiv_T \operatorname{Holant}([1, 0, 1]Z^{\otimes 2} \mid \{(Z^{-1})^{\otimes 4} f, (Z^{-1})^{\otimes 2} g\}) \\ &\equiv_T \operatorname{Holant}([0, 1, 0] \mid \{\hat{f}, \hat{g}\}),\end{aligned}$$

where up to a nonzero constant, $\hat{f} = [t, 1, 0, 0, 0]$ and $\hat{g} = [a, b, 1]$, for some $t, a, b \in \mathbb{C}$. We have $a \neq b^2$ since g is non-degenerate.

Our next goal is to show that we can realize a signature of the form $[c, 0, 1]$ with $c \neq 0$. If $b = 0$, then $\hat{g}$ is what we want since in this case $a = a - b^2 \neq 0$.

Now we assume $b \neq 0$. By connecting $\hat{g}$ to $\hat{f}$ via $\neq_2$, we get $[t + 2b, 1, 0]$. If $t \neq -2b$, then

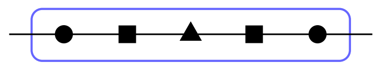


Figure 7.3: A sequence of binary gadgets that forms another binary gadget. The circles are assigned $[v, 1, 0]$, the squares are assigned $\neq_2$, and the triangle is assigned $[a, b, 1]$.

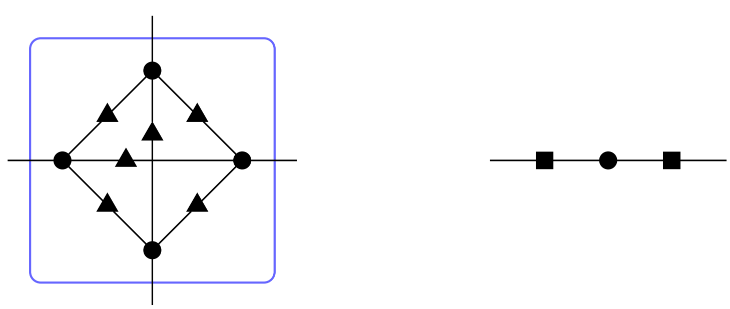
by Lemma 7.3.2, we can interpolate any binary signature of the form $[v, 1, 0]$. Otherwise $t = -2b$. Then we connect two copies of $\hat{g}$ via $\neq_2$, and get $\widehat{g}' = [2ab, a + b^2, 2b]$. By connecting this $\widehat{g}'$ to $\hat{f}$ via $\neq_2$, we get $[2(a - b^2), 2b, 0]$, using $t = -2b$. Since $a \neq b^2$ and $b \neq 0$, we can once again interpolate any $[v, 1, 0]$ by Lemma 7.3.2.

Hence, we have the signature $[v, 1, 0]$, where $v \in \mathbb{C}$ is for us to choose. We construct the gadget in Figure 7.3 with the circles assigned $[v, 1, 0]$, the squares assigned $\neq_2$, and the triangle assigned $[a, b, 1]$. The resulting gadget has signature $[a + 2bv + v^2, b + v, 1]$, which can be verified by the matrix product

$$\begin{bmatrix} v & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} a & b \\ b & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} v & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a + 2bv + v^2 & b + v \\ b + v & 1 \end{bmatrix}.$$

By setting $v = -b$, we get $[c, 0, 1]$, where $c = a - b^2 \neq 0$.

With this signature $[c, 0, 1]$, we construct the gadget in Figure 7.4, where $[c, 0, 1]$ is assigned to



(a) The tetrahedron gadget with edge signatures given in (b).

(b) The gadget representing an edge labeled by a triangle in (a).

Figure 7.4: The tetrahedron gadget with each triangle replaced by the edge in (b), where the circle is assigned $[c, 0, 1]$ and the squares are assigned $\neq_2$. The four circles in (a) are assigned $[t, 1, 0, 0, 0]$.

the circle vertex of arity two in Figure 7.4b and $\hat{f}$ is assigned to the four circle vertices of arity four in Figure 7.4a. We get a signature

$$\hat{h} = [3c^2 + 6ct^2 + t^4, 3ct + t^3, c + t^2, t, 1].$$

We note that this computation is reminiscent of matchgate signatures. The internal edge function $[1, 0, c]$ (which is a flip from $[c, 0, 1]$ since both sides are connected to $\neq_2$) is a generalized equality signature, and the signature $\hat{f}$ on the four circle vertices is a weighted version of the matching function At-Most-One.

The compressed signature matrix of $\hat{h}$ is

$$\widetilde{M_{\hat{h}}} = \begin{bmatrix} 3c^2 + 6ct^2 + t^4 & 2(3ct + t^3) & c + t^2 \\ 3ct + t^3 & 2(c + t^2) & t \\ c + t^2 & 2t & 1 \end{bmatrix}$$

and its determinant is $4c^3 \neq 0$. Thus $\widetilde{M_{\hat{h}}}$ is nonsingular. After a holographic transformation by Z^{-1}, where $Z = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$, the binary disequality $(\neq_2) = [0, 1, 0]$ is transformed to the binary equality $(=_2) = [1, 0, 1]$. Thus $\operatorname{Holant}([0, 1, 0] \mid \hat{h})$ is transformed to $\operatorname{Holant}(=_2 \mid Z^{\otimes 4}\hat{h})$, which is the same as $\operatorname{Holant}(Z^{\otimes 4}\hat{h})$. Then we are done by Corollary 6.5.7.

Now we do the induction step. Assume f is of arity $n \geq 5$. Since f is non-degenerate, $\operatorname{rd}^+(f) \geq 1$. If $\operatorname{rd}^+(f) = 1$, then we connect the binary g to f to get $f' = \langle f, g\rangle$. We have noted that $\operatorname{rd}^+(g) = 2$, so $\operatorname{vd}^+(g) = 0$. By Lemma 4.4.13, we have $\operatorname{rd}^+(f') = 1$ and $\operatorname{arity}(f') = n - 2 \geq 3$. Thus f' is vanishing. Also f' is non-degenerate, for otherwise let $f' = [a, b]^{\otimes(n-2)}$. If $[a, b]$ is a multiple of $[1, i]$, then $\operatorname{rd}^+(f') \leq 0$, which is false. If $[a, b]$ is not a multiple of $[1, i]$, then it can be directly checked that $f' \notin \mathscr{R}^+_{n-2}$, and $\operatorname{rd}^+(f') = n - 2 > 1$, which is also false. Hence f' is a non-degenerate vanishing signature of arity $n - 2$, so we are done by induction hypothesis.

Now suppose $\operatorname{rd}^+(f) = t \geq 2$. Since f is non-degenerate, it is certainly nonzero. Since it is vanishing, certainly $\operatorname{vd}^+(f) > 0$. Hence we can apply Lemma 4.4.14. Let f' be obtained from f by a self loop. Then $\operatorname{rd}^+(f') = t - 1 \geq 1$ and $\operatorname{arity}(f') = n - 2$. Clearly f' is still vanishing.

We claim that f' is non-degenerate. This follows using the same argument as above. If f' were degenerate, then either $\operatorname{rd}^+(f') \leq 0$ or $\operatorname{rd}^+(f') = \operatorname{arity}(f')$, which would contradict f' being a vanishing signature. Therefore, we can apply the induction hypothesis. □

Remark. The gadget in Figure 7.4 is rather complicated. In the process of strengthening this result to planar graphs [27], we found a simpler proof that $\operatorname{Pl-Holant}(\neq_2 \mid [t,1,0,0,0],[c,0,1])$ is #P-hard provided $t \neq 0$, which is planar tractable. I give this proof next combined with an equally simple proof that the $\operatorname{Holant}(\neq_2 \mid [0,1,0,0,0],[c,0,1])$ is #P-hard.

Lemma 7.3.4. *Let $c,t \in \mathbb{C}$. If $c \neq 0$, then $\operatorname{Holant}(\neq_2 \mid [t,1,0,0,0],[c,0,1])$ is #P-hard.*

Proof. By connecting two copies of $\neq_2$ to either side of $[c,0,1]$, we get the signature $[1,0,c]$ on the left. Clearly $\operatorname{Holant}([1,0,c] \mid [t,1,0,0,0]) \leq_T \operatorname{Holant}(\neq_2 \mid [t,1,0,0,0],[c,0,1])$ even when restricted to planar graphs. Then under a holographic transformation by T^{-1}, where $T = \left[\begin{smallmatrix} 1 & 0 \\ 0 & \sqrt{c} \end{smallmatrix}\right]$, we have

$$\begin{aligned} \operatorname{Holant}([1,0,c] \mid [t,1,0,0,0]) &\equiv \operatorname{Holant}([1,0,c](T^{-1})^{\otimes 2} \mid T^{\otimes 4}[t,1,0,0,0]) \\ &\equiv \operatorname{Holant}([1,0,1] \mid [t,\sqrt{c},0,0,0]) \\ &\equiv \operatorname{Holant}([t,\sqrt{c},0,0,0]) \end{aligned}$$

even when restricted to planar graphs. Up to a nonzero facto of $\sqrt{c}$, we have $\operatorname{Holant}([v,1,0,0,0])$ with $v = \frac{t}{\sqrt{c}}$. If $v \neq 0$, then $\operatorname{Pl-Holant}([v,1,0,0,0])$ is #P-hard by Corollary 6.3.6. Otherwise $t = 0$ and $\operatorname{Holant}([0,1,0,0,0])$ is #P-hard by Lemma 6.6.2. □

Next we consider a pair of vanishing signatures of opposite type, both of arity at least 3. We show that opposite types of vanishing signatures cannot mix. More formally, vanishing signatures of opposite types, when put together, lead to #P-hardness. First though we show hardness for a special case.

Lemma 7.3.5. *If $f = [0,1,0,\ldots,0]$ and $g = [0,\ldots,0,1,0]$ are both of arity $n \geq 3$, then the problem $\operatorname{Holant}([0,1,0] \mid \{f,g\})$ is #P-hard.*

Proof. Our goal is to obtain a signature that satisfies the hypothesis of Corollary 6.5.7.

(a) The circle is assigned f, the triangle is assigned g, and the squares are assigned $\neq_2$.

(b) The circle is assigned h', the triangle is assigned h'', and the squares are assigned $\neq_2$.

Figure 7.5: Gadget constructions used to obtain a hard and redundant arity 4 signature.

The gadget in Figure 7.5a, with f assigned to the circle vertex, g assigned to the triangle vertex, and $\neq_2$ assigned to the square vertices, has signature h with signature matrix

$$M_h = \begin{bmatrix} 0 & 0 & 0 & v \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $v = n-2$ is positive since $n \geq 3$. Although this signature matrix is redundant, its compressed form is singular. Rotating this gadget 90° clockwise and 90° counterclockwise yield signatures h' and h'' respectively, with signature matrices

$$M_{h'} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & v & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad \text{and} \qquad M_{h''} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & v & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

The gadget in Figure 7.5b, with h' assigned to the circle vertex, h'' assigned to the triangle vertex, and $\neq_2$ assigned to the square vertices, has a signature r with signature matrix

$$M_r = M_{h'} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} M_{h''} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & v & v^2+1 & 0 \\ 0 & 1 & v & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

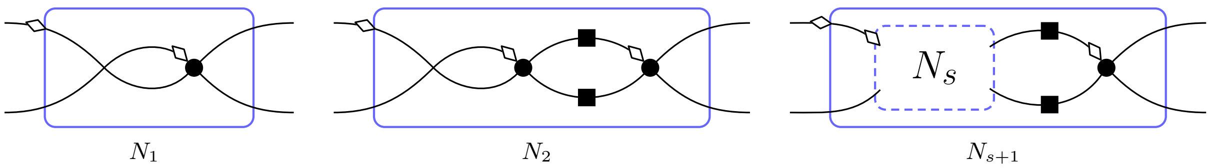


Figure 7.6: Recursive construction to interpolate a signature r' that is only a rotation away from having a redundant signature matrix and nonsingular compressed matrix. The circles are assigned r and the squares are assigned $\neq_2$.

Note that the effect of the $\neq_2$ signatures is to reverse all four rows of $M_{h''}$ before multiplying it to the right of $M_{h'}$. Although this signature matrix is not redundant, every entry of Hamming weight 2 is nonzero since v is positive.

For any nonzero value $t \in \mathbb{C}$, we claim that we can use r to interpolate the following signature r' via the construction in Figure 7.6. Define $p^{\pm} = (v \pm \sqrt{v^2+4})/2$, $P = \left[\begin{smallmatrix} 1 & 1 \\ p^+ & p^- \end{smallmatrix}\right]$, and $T = P \left[\begin{smallmatrix} t & 0 \\ 0 & t^{-1} \end{smallmatrix}\right] P^{-1}$ where $t \in \mathbb{C}$ is any nonzero value. The signature matrix of the target signature r' is

$$M_{r'} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & \multicolumn{2}{c}{\multirow{2}{*}{T}} & 0 \\ 0 & & & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \tag{7.3.1}$$

Consider an instance Ω of Holant($\neq_2 \mid \mathcal{F} \cup \{r'\}$) with $r \in \mathcal{F}$. Suppose that r' appears n times in Ω. We construct from Ω a sequence of instances Ω_s of Holant($\neq_2 \mid \mathcal{F}$) indexed by $s \geq 1$. We obtain Ω_s from Ω by replacing each occurrence of r' with the gadget N_s in Figure 7.6 with r assigned to the circle vertices and $\neq_2$ assigned to the square vertices. In Ω_s, the edge corresponding to the ith significant index bit of N_s connects to the same location as the edge corresponding to the ith significant index bit of r' in Ω.

The signature matrix of N_s is the sth power of the matrix obtained from M_r after reversing all rows, and then switching the first and last rows of the final product, namely

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & v & 0 \\ 0 & v & v^2+1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^s = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & v & 0 \\ 0 & v & v^2+1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & v & 0 \\ 0 & v & v^2+1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{s-1}.$$

The twist of the two input edges on the left side for the first copy of M_r switches the middle two rows, which is equivalent to a total reversal of all rows, followed by the switching of the first and last rows. The total reversals of rows for all subsequent $s-1$ copies of M_r are due to the presence of $\neq_2$ signatures.

After such reversals of rows, it is clear that the matrix is a direct sum of block matrices indexed by $\{00, 11\} \times \{00, 11\}$ and $\{01, 10\} \times \{10, 01\}$. Furthermore, in the final product, the block indexed by $\{00, 11\} \times \{00, 11\}$ is $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. Thus in the gadget N_s, the only entries of M_{N_s} that vary with s are the four entries in the middle. These middle four entries of M_{N_s} form the 2-by-2 matrix $\left[\begin{smallmatrix} 1 & v \\ v & v^2+1 \end{smallmatrix}\right]^s$. Since $\left[\begin{smallmatrix} 1 & v \\ v & v^2+1 \end{smallmatrix}\right] = P \left[\begin{smallmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{smallmatrix}\right] P^{-1}$, where $\lambda_\pm = (v^2 + 2 \pm v\sqrt{v^2+4})/2$ are the eigenvalues, we have

$$\begin{bmatrix} 1 & v \\ v & v^2+1 \end{bmatrix}^s = P \begin{bmatrix} \lambda_+^s & 0 \\ 0 & \lambda_-^s \end{bmatrix} P^{-1}.$$

The determinant is $\lambda_+ \lambda_- = 1$, so the eigenvalues are nonzero. Since v is positive, the ratio of the eigenvalues λ_+/λ_- is not a root of unity, so neither λ_+ nor λ_- is a root of unity.

Now we determine the relationship between Holant(Ω) and Holant(Ω_s). We can view our construction of Ω_s as first replacing $M_{r'}$ with

$$\begin{bmatrix} 1 & 0 \; 0 & 0 \\ 0 & & 0 \\ & P & \\ 0 & & 0 \\ 0 & 0 \; 0 & 1 \end{bmatrix} \Lambda \begin{bmatrix} 1 & 0 \; 0 & 0 \\ 0 & & 0 \\ & P^{-1} & \\ 0 & & 0 \\ 0 & 0 \; 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \; 0 & 0 \\ 0 & & 0 \\ & P & \\ 0 & & 0 \\ 0 & 0 \; 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & t & 0 & 0 \\ 0 & 0 & t^{-1} & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \; 0 & 0 \\ 0 & & 0 \\ & P^{-1} & \\ 0 & & 0 \\ 0 & 0 \; 0 & 1 \end{bmatrix},$$

which does not change the Holant value, and then replacing Λ with

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & \lambda_+^s & 0 & 0 \\ 0 & 0 & \lambda_-^s & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

We stratify the assignments in Ω_s based on the assignments to the n occurrences of Λ. The inputs to this matrix are from $\{0,1\}^2 \times \{0,1\}^2$, which correspond to the four input bits. We only need to consider the assignments that assign

- i many times the bit pattern 0110,
- j many times the bit pattern 1001, and
- k many times the bit patterns 0011 or 1100,

since any other assignment contributes a factor of 0. Let c_{ijk} be the sum over all such assignments of the products of evaluations of all signatures in Ω_s except for Λ. Then

$$\operatorname{Holant}(\Omega) = \sum_{i+j+k=n} t^{i-j} c_{ijk}$$

and the value of the Holant on Ω_s, for $s \geq 1$, is

$$\operatorname{Holant}(\Omega_s) = \sum_{i+j+k=n} \lambda_+^{si} \lambda_-^{sj} c_{ijk} = \sum_{i+j+k=n} \lambda_+^{s(i-j)} c_{ijk}.$$

This Vandermonde system does not have full rank. However, we can define for $-n \leq \ell \leq n$,

$$c'_\ell = \sum_{\substack{i-j=\ell \\ i+j+k=n}} c_{ijk}.$$

Then

$$\operatorname{Holant}(\Omega) = \sum_{-n \leq \ell \leq n} t^\ell c'_\ell \qquad \text{and} \qquad \operatorname{Holant}(\Omega_s) = \sum_{-n \leq \ell \leq n} \lambda_+^{s\ell} c'_\ell.$$

Now this Vandermonde has full rank because λ_+ is neither 0 nor a root of unity. Therefore, we can solve for the unknowns c'_ℓ and obtain the value of $\operatorname{Holant}(\Omega)$. This completes our claim that we can interpolate the signature r' in (7.3.1), for any nonzero $t \in \mathbb{C}$.

Let $t = (\sqrt{v^2+8} + \sqrt{v^2+4})/2$ so $t^{-1} = (\sqrt{v^2+8} - \sqrt{v^2+4})/2$. Let $a = (\sqrt{v^2+8} - v)/2$ and $b = (\sqrt{v^2+8} + v)/2$, so $ab = 2 \neq 0$. One can verify that

$$P \begin{bmatrix} t & 0 \\ 0 & t^{-1} \end{bmatrix} P^{-1} = \begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix}.$$

Thus, the signature matrix for r' is

$$M_{r'} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & a & 1 & 0 \\ 0 & 1 & b & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

After a quarter rotation on the edges of r', we have a signature r'' with a redundant signature matrix

$$M_{r''} = \begin{bmatrix} 0 & 0 & 0 & a \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ b & 0 & 0 & 0 \end{bmatrix},$$

and its compressed signature matrix is nonsingular. After a holographic transformation by Z^{-1}, where $Z = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$, the binary disequality $(\neq_2) = [0,1,0]$ is transformed to the binary equality $(=_2) = [1,0,1]$. Thus the problem $\operatorname{Holant}([0,1,0] \mid r'')$ is transformed to $\operatorname{Holant}(=_2 \mid Z^{\otimes 4} r'')$, which is the same as $\operatorname{Holant}(Z^{\otimes 4} r'')$. Then we are done by Corollary 6.5.7. □

Lemma 7.3.6. *Let $f \in \mathscr{V}^+$ and $g \in \mathscr{V}^-$ be symmetric non-degenerate signatures of arity at least* 3. *Then* $\operatorname{Holant}(\{f, g\})$ *is* #P*-hard.*

Proof. Let $\operatorname{rd}^+(f) = d$, $\operatorname{rd}^-(g) = d'$, $\operatorname{arity}(f) = n$ and $\operatorname{arity}(g) = n'$, then $2d < n$ and $2d' < n'$. We can apply Lemma 4.4.14 zero or more times to construct a signature obtained from g by adding a certain number of self loops. The resulting signature is a tensor power of $[1, -i]$ up to a nonzero scalar. To see this, note that we start with $\mathrm{rd}^-(g) < \mathrm{vd}^-(g)$ with their sum being arity(g). We can apply Lemma 4.4.14 if the signature is nonzero and its vd^- is positive. Each time we apply Lemma 4.4.14, we reduce rd^- and vd^- by one, and the arity by two. Thus $\mathrm{rd}^- < \mathrm{vd}^-$ is maintained until rd^- becomes 0, at which point the signature is a tensor power of $[1, -i]$ up to a nonzero scalar. The initial signature g is non-degenerate by assumption, so it is certainly nonzero. While rd^- is positive, the signature is nonzero, thus Lemma 4.4.14 applies. If $d \geq 2$, then by Lemma 7.3.1, $\operatorname{Holant}(\{f, g\})$ is #P-hard. Similarly, the problem is #P-hard if $d' \geq 2$. Thus we may assume that $d = d' = 1$.

We perform the $Z = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$ transformation

$$\begin{aligned} \operatorname{Holant}(=_2 \mid \{f, g\}) &\equiv_T \operatorname{Holant}([1,0,1]Z^{\otimes 2} \mid \{(Z^{-1})^{\otimes n} f, (Z^{-1})^{\otimes n'} g\}) \\ &\equiv_T \operatorname{Holant}([0,1,0] \mid \{\hat{f}, \hat{g}\}). \end{aligned}$$

Since $\mathrm{rd}(f) = d = 1$, by Lemma 4.4.16, we have $(Z^{-1})^{\otimes n} f = \hat{f} = [\hat{f}_0, \hat{f}_1, 0, \ldots, 0]$, where $\hat{f}_1 \neq 0$. Similarly, for g with $\mathrm{rd}^-(g) = d' = 1$, we have $(Z^{-1})^{\otimes n'} g = \hat{g} = [0, \ldots, 0, \hat{g}_1, \hat{g}_0]$, where $\hat{g}_1 \neq 0$.

So up to nonzero constants, we have $\hat{f} = [a, 1, 0, \ldots, 0]$ and $\hat{g} = [0, \ldots, 0, 1, b]$ for some $a, b \in \mathbb{C}$. We show that it is always possible to get two such signatures of the same arity $\min\{n, n'\}$. Suppose $n > n'$. We form a loop from $\hat{f}$, where the loop is really a path consisting of one vertex and two edges, with the vertex assigned the signature $\neq_2$. It is easy to see that this signature is the degenerate signature $2[1,0]^{\otimes(n-2)}$. Similarly, we can form a loop from $\hat{g}$ and can get $2[0,1]^{\otimes(n'-2)}$. Thus we have both $[1,0]^{\otimes(n-2)}$ and $[0,1]^{\otimes(n'-2)}$. We can connect all $n'-2$ edges of the second to the first, connected by $\neq_2$. This gives $[1,0]^{\otimes(n-n')}$. We can continue subtracting the smaller arity from the larger one. We continue this process in a subtractive version of the Euclidean algorithm, and end up with both $[1,0]^{\otimes t}$ and $[0,1]^{\otimes t}$, where $t = \gcd(n-2, n'-2) = \gcd(n-n', n'-2)$. In particular, $t \mid n - n'$ and by taking $(n-n')/t$ many copies of $[0,1]^{\otimes t}$, we can get $[0,1]^{\otimes(n-n')}$. Connecting this back to $\hat{f}$ via $\neq_2$, we get a symmetric signature of arity n' consisting of the first $n'+1$ entries of $\hat{f}$. A similar proof works when $n' > n$.

Thus without loss of generality, we may assume $n = n'$. If $a \neq 0$, then connect $[0,1]^{\otimes(n-2)}$ to $\hat{f} = [a,1,0,\ldots,0]$ via $\neq_2$ we get $\hat{h} = [a,1,0]$. For $a \neq 0$, translating this back by Z, we have a binary signature $h \notin \mathscr{R}_2^-$ together with the given $g \in \mathscr{V}^-$. By Lemma 7.3.3, $\operatorname{Holant}(\{f,g\})$ is #P-hard. A similar proof works for the case $b \neq 0$.

The only case left is when $\hat{f} = [0,1,0,\ldots,0]$ of arity n, and $\hat{g} = [0,\ldots,0,1,0]$ of arity n. This is #P-hard by Lemma 7.3.5. □

7.4 $\mathscr{A}$- and $\mathscr{P}$-transformable Signatures

In this section, we investigate the properties of $\mathscr{A}$- and $\mathscr{P}$-transformable signatures. Throughout, we define $\alpha = \frac{1+i}{\sqrt{2}} = \sqrt{i} = e^{\frac{\pi i}{4}}$ Recall that $\mathscr{F}_{123} = \mathscr{F}_1 \cup \mathscr{F}_2 \cup \mathscr{F}_3$, where $\mathscr{F}_1$, $\mathscr{F}_2$, and $\mathscr{F}_3$ are defined in Section 4.2. While the main results in this section assume that the signatures involved are symmetric, we note that some of the lemmas also hold without this assumption.

7.4.1 Characterization of $\mathscr{A}$- and $\mathscr{P}$-transformable Signatures

Recall that by definition, if a set of signatures $\mathcal{F}$ is $\mathscr{A}$-transformable (resp. $\mathscr{P}$-transformable), then the binary equality $=_2$ must be simultaneously transformed into $\mathscr{A}$ (resp. $\mathscr{P}$) along with $\mathcal{F}$. We first characterize the possible matrices of such a transformation by just considering the transformation of the binary equality. While there are many binary signatures in $\mathscr{A} \cup \mathscr{P}$, it turns out that it is sufficient to consider only three signatures.

Proposition 7.4.1. *Let $T \in \mathbb{C}^{2\times 2}$ be a matrix. Then the following hold:*

1. $[1,0,1]T^{\otimes 2} = [1,0,1]$ *iff* $T \in \mathbf{O}_2(\mathbb{C})$*;*
2. $[1,0,1]T^{\otimes 2} = [1,0,i]$ *iff there exists an* $H \in \mathbf{O}_2(\mathbb{C})$ *such that* $T = H \left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right]$*;*
3. $[1,0,1]T^{\otimes 2} = [0,1,0]$ *iff there exists an* $H \in \mathbf{O}_2(\mathbb{C})$ *such that* $T = \frac{1}{\sqrt{2}} H \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$.

Proof. Case 1 is clear since

$$[1,0,1]T^{\otimes 2} = [1,0,1] \iff T^{\intercal} I_2 T = I_2 \iff T^{\intercal} T = I_2,$$

the definition of a (2-by-2) orthogonal matrix. Now we use this case to prove the others.

For $M_2 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right]$ and $M_3 = Z = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$, let $T_j = HM_j$ (for $j = 2, 3$), where $H \in \mathbf{O}_2(\mathbb{C})$. Then

$$[1,0,1]T_j^{\otimes 2} = [1,0,1](HM_j)^{\otimes 2} = [1,0,1]M_j^{\otimes 2} = f_j,$$

where f_j is the binary signature in case j.

On the other hand, suppose that $[1,0,1](T_j)^{\otimes 2} = f_j$. Then we have

$$[1,0,1](T_j M_j^{-1})^{\otimes 2} = f_j (M_j^{-1})^{\otimes 2} = [1,0,1],$$

so $H = T_j M_j^{-1} \in \mathbf{O}_2(\mathbb{C})$ by case 1. Thus $T_j = HM_j$ as desired. □

We also need the following lemma; the proof is direct.

Lemma 7.4.2. *If a symmetric signature $f = [f_0, f_1, \ldots, f_n]$ can be expressed in the form $f = a[1,\lambda]^{\otimes n} + b[1,\mu]^{\otimes n}$, for some $a, b, \lambda, \mu \in \mathbb{C}$, then the f_k's satisfy the recurrence relation $f_{k+2} = (\lambda+\mu) f_{k+1} - \lambda\mu f_k$ for $0 \le k \le n-2$.*

To simplify the proof of the characterization of the $\mathscr{A}$-transformable signatures, we introduce the left and right stabilizer groups of $\mathscr{A}$:

$$\mathrm{LStab}(\mathscr{A}) = \{T \in \mathbf{GL}_2(\mathbb{C}) \mid T\mathscr{A} \subseteq \mathscr{A}\};$$
$$\mathrm{RStab}(\mathscr{A}) = \{T \in \mathbf{GL}_2(\mathbb{C}) \mid \mathscr{A}T \subseteq \mathscr{A}\}.$$

In fact, these two groups are equal and coincide with the group of nonsingular signature matrices of binary affine signatures. More precisely, for a binary signature $f = (f^{00}, f^{01}, f^{10}, f^{11})$, we define its signature matrix M_f to be

$$M_f = \begin{bmatrix} f^{00} & f^{01} \\ f^{10} & f^{11} \end{bmatrix}.$$

Let

$$\mathscr{A}^{2\times 2} = \{M_f \mid f \in \mathscr{A},\ \text{arity}(f) = 2,\ \text{and}\ \det(M_f) \neq 0\}$$

be the set of nonsingular signature matrices of the binary affine signatures. It is straightforward to verify that $\mathscr{A}^{2\times 2}$ is closed under multiplication and inverses. Therefore $\mathscr{A}^{2\times 2}$ forms a group.

Let $D = \left[\begin{smallmatrix}1 & 0\\ 0 & i\end{smallmatrix}\right]$ and $H_2 = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right]$. Also let $X = \left[\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right]$ and $Z = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1 & 1\\ i & -i\end{smallmatrix}\right]$. Note that $Z = DH_2$ and that $D^2 Z = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1 & 1\\ -i & i\end{smallmatrix}\right] = ZX$, hence $X = Z^{-1}D^2 Z$. Furthermore, $D, H_2, X, Z \in \operatorname{LStab}(\mathscr{A}) \cap \operatorname{RStab}(\mathscr{A}) \cap \mathscr{A}^{2\times 2}$, as well as all nonzero scalar multiples of these matrices.

Not only are the groups $\operatorname{LStab}(\mathscr{A})$, $\operatorname{RStab}(\mathscr{A})$, and $\mathscr{A}^{2\times 2}$ equal, they are generated by D and H_2 with a nonzero scalar multiple.

Lemma 7.4.3. $\operatorname{LStab}(\mathscr{A}) = \operatorname{RStab}(\mathscr{A}) = \mathscr{A}^{2\times 2} = \mathbb{C}^* \cdot \langle D, H_2\rangle$.

Proof. Let

$$\mathbf{S} = \{S \in \mathbf{GL}_2(\mathbb{C}) \mid \mathscr{F}_{123} S \subseteq \mathscr{F}_{123}\}$$

be the right stabilizer group of $\mathscr{F}_{123}$. It is easy to verify that $\mathscr{A}^{2\times 2} \subseteq \operatorname{RStab}(\mathscr{A}) \subseteq \mathbf{S}$. Together with the fact that $D, H_2 \in \mathscr{A}^{2\times 2}$, we have $\mathbb{C}^* \cdot \langle D, H_2\rangle \subseteq \mathscr{A}^{2\times 2} \subseteq \operatorname{RStab}(\mathscr{A}) \subseteq \mathbf{S}$. To finish the proof, we show that $\mathbf{S} \subseteq \mathbb{C}^* \cdot \langle D, H_2\rangle$. For $\operatorname{LStab}(\mathscr{A})$, the proof is similar.

Consider some $T \in \mathbf{S}$. For $f = (=_3)$, we have $fT^{\otimes 3} \in \mathscr{F}_{123}$. Then by the form of $\mathscr{F}_{123}$, for some $M \in \langle D, H_2\rangle$, chosen to be either I, or $H_2^{\intercal} = H_2$, or $Z^{\intercal} = H_2 D$, we have $f(TM^{-1})^{\otimes 3} \in \mathscr{F}_1$, which is a generalized equality signature. Then either TM^{-1} or $TM^{-1}X$ is a diagonal matrix $T' = \lambda\left[\begin{smallmatrix}1 & 0\\ 0 & d\end{smallmatrix}\right]$. Furthermore, by applying T' to $=_4$, we conclude that $(=_4)T'^{\otimes 4} \in \mathscr{F}_1$, since it is in $\mathscr{F}_{123}$ but not in $\mathscr{F}_2 \cup \mathscr{F}_3$ because T' is diagonal. It follows that d is a power of i, and hence $\left[\begin{smallmatrix}1 & 0\\ 0 & d\end{smallmatrix}\right]$ is a power of D. Thus $T \in \mathbb{C}^* \cdot \langle D, H_2\rangle$. □

Since $\operatorname{LStab}(\mathscr{A}) = \operatorname{RStab}(\mathscr{A})$, we simply write $\operatorname{Stab}(\mathscr{A})$ for this group. Of course each T under which $\mathcal{F}$ is $\mathscr{A}$-transformable is just a particular solution that can be extended by any element in $\operatorname{Stab}(\mathscr{A})$.

Lemma 7.4.4. *Let $\mathcal{F}$ be a set of signatures. Then $\mathcal{F}$ is $\mathscr{A}$-transformable under T iff $\mathcal{F}$ is $\mathscr{A}$-transformable under any $T' \in T\operatorname{Stab}(\mathscr{A})$.*

Proof. Sufficiency is trivial since $I_2 \in \operatorname{Stab}(\mathscr{A})$. If $\mathcal{F}$ is $\mathscr{A}$-transformable under T, then by definition, we have $(=_2)T^{\otimes 2} \in \mathscr{A}$ and $\mathcal{F}' = T^{-1}\mathcal{F} \subseteq \mathscr{A}$. Let $T' = TM \in T\operatorname{Stab}(\mathscr{A})$ for any $M \in \operatorname{Stab}(\mathscr{A})$. It then follows that $(=_2)T'^{\otimes 2} = (=_2)T^{\otimes 2}M^{\otimes 2} \in \mathscr{A}M = \mathscr{A}$ and $T'^{-1}\mathcal{F} = M^{-1}\mathcal{F}' \subseteq M^{-1}\mathscr{A} = \mathscr{A}$. Therefore $\mathcal{F}$ is $\mathscr{A}$-transformable under any $T' \in T\operatorname{Stab}(\mathscr{A})$. □

After restricting by Proposition 7.4.1 and normalizing by Lemma 7.4.4, one only needs to check a small subset of $\mathbf{GL}_2(\mathbb{C})$ to determine if $\mathcal{F}$ is $\mathscr{A}$-transformable.

Lemma 7.4.5. *Let $\mathcal{F}$ be a set of signatures. Then $\mathcal{F}$ is $\mathscr{A}$-transformable iff there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $\mathcal{F} \subseteq H\mathscr{A}$ or $\mathcal{F} \subseteq H \left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right] \mathscr{A}$.*

Proof. Sufficiency is easily verified by checking that $=_2$ is transformed into $\mathscr{A}$ in both cases. In particular, H leaves $=_2$ unchanged.

If $\mathcal{F}$ is $\mathscr{A}$-transformable, then by definition, there exists a matrix T such that $(=_2)T^{\otimes 2} \in \mathscr{A}$ and $T^{-1}\mathcal{F} \subseteq \mathscr{A}$. Since $=_2$ is non-degenerate and symmetric, $(=_2)T^{\otimes 2} \in \mathscr{A}$ is equivalent to $(=_2)T^{\otimes 2} \in \mathscr{F}_{123}$.

Any signature in $\mathscr{F}_{123}$ is expressible as $c(v_1^{\otimes n} + i^t v_2^{\otimes n})$, where $t \in \{0, 1, 2, 3\}$ and (v_1, v_2) is a pair of vectors in the set

$$\left\{ \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right), \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right), \left(\begin{bmatrix} 1 \\ i \end{bmatrix}, \begin{bmatrix} 1 \\ -i \end{bmatrix} \right) \right\}.$$

We use $\operatorname{Stab}(\mathscr{A})$ to further normalize these three sets by Lemma 7.4.4. In particular, $\mathscr{F}_1 = H_2\mathscr{F}_2$ and $\mathscr{F}_1 = (DH_2)^{-1}\mathscr{F}_3$. Furthermore, the binary signatures in $\mathscr{F}_1$ are just the four signatures $[1, 0, 1]$, $[1, 0, i]$, $[1, 0, -1]$, and $[1, 0, -i]$ up to a scalar. We also normalize these four as $[1, 0, 1] = [1, 0, -1]D^{\otimes 2}$ and $[1, 0, i] = [1, 0, -i]D^{\otimes 2}$. Hence $\mathcal{F}$ being $\mathscr{A}$-transformable implies that there exists a matrix T such that $(=_2)T^{\otimes 2} \in \{[1, 0, 1], [1, 0, i]\}$ and $T^{-1}\mathcal{F} \subseteq \mathscr{A}$. Now we apply Proposition 7.4.1.

1. If $(=_2)T^{\otimes 2} = [1, 0, 1]$, then by case 1 of Proposition 7.4.1, we have $T \in \mathbf{O}_2(\mathbb{C})$. Therefore $\mathcal{F} \subseteq H\mathscr{A}$ where $H = T \in \mathbf{O}_2(\mathbb{C})$.
2. If $(=_2)T^{\otimes 2} = [1, 0, i]$, then by case 2 of Proposition 7.4.1, there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $T = H \left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right]$. Therefore $\mathcal{F} \subseteq T\mathscr{A} = H \left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha \end{smallmatrix}\right] \mathscr{A}$ where $H \in \mathbf{O}_2(\mathbb{C})$.

This completes the proof. □

Using these two lemmas, we can characterize all $\mathscr{A}$-transformable signatures. We first define the three sets $\mathscr{A}_1$, $\mathscr{A}_2$, and $\mathscr{A}_3$.

Definition 7.4.6. A symmetric signature f of arity n is in $\mathscr{A}_1$ if there exists an $H \in \mathbf{O}_2(\mathbb{C})$ and a nonzero constant $c \in \mathbb{C}$ such that $f = cH^{\otimes n}\left(\left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]^{\otimes n} + \beta\left[\begin{smallmatrix}1\\-1\end{smallmatrix}\right]^{\otimes n}\right)$, where $\beta = \alpha^{tn+2r}$ for some $r \in \{0,1,2,3\}$ and $t \in \{0,1\}$.

When such an H exists, we say that $f \in \mathscr{A}_1$ with transformation H. If $f \in \mathscr{A}_1$ with I_2, then we say f is in the canonical form of $\mathscr{A}_1$. If f is in the canonical form of $\mathscr{A}_1$, then by Lemma 7.4.2, for any $0 \le k \le n-2$, we have $f_{k+2} = f_k$ and one of the following holds:

- $f_0 = 0$, or
- $f_1 = 0$, or
- $f_1 = \pm i f_0 \neq 0$, or
- n is odd and $f_1 = \pm(1 \pm \sqrt{2}) i f_0 \neq 0$ (all four sign choices are permissible).

Notice that when n is odd and $t = 1$ in Definition 7.4.6, it has some complication as described by the factor α^{tn+2r}.

Definition 7.4.7. A symmetric signature f of arity n is in $\mathscr{A}_2$ if there exists an $H \in \mathbf{O}_2(\mathbb{C})$ and a nonzero constant $c \in \mathbb{C}$ such that $f = cH^{\otimes n}\left(\left[\begin{smallmatrix}1\\i\end{smallmatrix}\right]^{\otimes n} + \left[\begin{smallmatrix}1\\-i\end{smallmatrix}\right]^{\otimes n}\right)$.

Similarly, when such an H exists, we say that $f \in \mathscr{A}_2$ with transformation H. If $f \in \mathscr{A}_2$ with I_2, then we say f is in the canonical form of $\mathscr{A}_2$. If f is in the canonical form of $\mathscr{A}_2$, then by Lemma 7.4.2, for any $0 \le k \le n-2$, we have $f_{k+2} = -f_k$. Non-degeneracy of f implies $f_1 \neq \pm i f_0$.

It is worth noting that $\{\left[\begin{smallmatrix}1\\i\end{smallmatrix}\right], \left[\begin{smallmatrix}1\\-i\end{smallmatrix}\right]\}$ is setwise invariant up to scale under any transformation in $\mathbf{O}_2(\mathbb{C})$ up to nonzero constants. That is, these vectors are the eigenvectors of orthogonal matrices. Thus for any $H \in \mathbf{O}_2(\mathbb{C})$, we can write $\left[\begin{smallmatrix}1 & 1\\ i & -i\end{smallmatrix}\right]^{-1} H \left[\begin{smallmatrix}1 & 1\\ i & -i\end{smallmatrix}\right] = D$, where D is either a diagonal or anti-diagonal matrix. It is also helpful to view this equation as $H\left[\begin{smallmatrix}1 & 1\\ i & -i\end{smallmatrix}\right] = \left[\begin{smallmatrix}1 & 1\\ i & -i\end{smallmatrix}\right] D$.

Using this fact, the following lemma gives a characterization of $\mathscr{A}_2$. It says that any signature in $\mathscr{A}_2$ is essentially in canonical form.

Lemma 7.4.8. *Let f be a symmetric arity n signature. Then $f \in \mathscr{A}_2$ iff $f = c\left(\left[\begin{smallmatrix}1\\i\end{smallmatrix}\right]^{\otimes n} + \beta\left[\begin{smallmatrix}1\\-i\end{smallmatrix}\right]^{\otimes n}\right)$ for some nonzero constants $c, \beta \in \mathbb{C}$.*

Proof. Assume $f = c\left(\left[\begin{smallmatrix}1\\i\end{smallmatrix}\right]^{\otimes n} + \beta\left[\begin{smallmatrix}1\\-i\end{smallmatrix}\right]^{\otimes n}\right)$ for some $c, \beta \neq 0$. Consider the orthogonal transformation $H = \left[\begin{smallmatrix}a & b\\ b & -a\end{smallmatrix}\right]$, where $a = \frac{1}{2}\left(\beta^{\frac{1}{2n}} + \beta^{-\frac{1}{2n}}\right)$ and $b = \frac{1}{2i}\left(\beta^{\frac{1}{2n}} - \beta^{-\frac{1}{2n}}\right)$. We pick a and b so that

$a+bi=\beta^{\frac{1}{2n}}$, $a-bi=\beta^{-\frac{1}{2n}}$, and $(a+bi)(a-bi)=a^2+b^2=1$. Also $\left(\frac{a+bi}{a-bi}\right)^n=\beta$. Then

$$\begin{aligned} H^{\otimes n}f &= c\left(\begin{bmatrix} a+bi \\ -ai+b \end{bmatrix}^{\otimes n} + \beta \begin{bmatrix} a-bi \\ ai+b \end{bmatrix}^{\otimes n}\right) \\ &= c\left((a+bi)^n \begin{bmatrix} 1 \\ -i \end{bmatrix}^{\otimes n} + (a-bi)^n\beta \begin{bmatrix} 1 \\ i \end{bmatrix}^{\otimes n}\right) \\ &= c\sqrt{\beta}\left(\begin{bmatrix} 1 \\ -i \end{bmatrix}^{\otimes n} + \begin{bmatrix} 1 \\ i \end{bmatrix}^{\otimes n}\right), \end{aligned}$$

so f can be written as

$$f = c\sqrt{\beta}(H^{-1})^{\otimes n}\left(\begin{bmatrix} 1 \\ i \end{bmatrix}^{\otimes n} + \begin{bmatrix} 1 \\ -i \end{bmatrix}^{\otimes n}\right).$$

Therefore $f \in \mathscr{A}_2$.

On the other hand, the desired form $f = c([\begin{smallmatrix} 1 \\ i \end{smallmatrix}]^{\otimes n} + \beta [\begin{smallmatrix} 1 \\ -i \end{smallmatrix}]^{\otimes n})$ follows from the fact that $\{[\begin{smallmatrix} 1 \\ i \end{smallmatrix}], [\begin{smallmatrix} 1 \\ -i \end{smallmatrix}]\}$ is fixed setwise under any orthogonal transformation up to nonzero constants. $\square$

Definition 7.4.9. A symmetric signature f of arity n is in $\mathscr{A}_3$ if there exists an $H \in \mathbf{O}_2(\mathbb{C})$ and a nonzero constant $c \in \mathbb{C}$ such that $f = cH^{\otimes n}\left([\begin{smallmatrix} 1 \\ \alpha \end{smallmatrix}]^{\otimes n} + i^r [\begin{smallmatrix} 1 \\ -\alpha \end{smallmatrix}]^{\otimes n}\right)$ for some $r \in \{0, 1, 2, 3\}$.

Again, when such an H exists, we say that $f \in \mathscr{A}_3$ with transformation H. If $f \in \mathscr{A}_3$ with I_2, then we say f is in the canonical form of $\mathscr{A}_3$. If f is in the canonical form of $\mathscr{A}_3$, then by Lemma 7.4.2, for any $0 \le k \le n-2$, we have $f_{k+2} = i f_k$ and one of the following holds:

- $f_0 = 0$, or
- $f_1 = 0$, or
- $f_1 = \pm\alpha i f_0 \neq 0$.

Now we characterize the $\mathscr{A}$-transformable signatures.

Lemma 7.4.10. *Let f be a non-degenerate symmetric signature. Then f is $\mathscr{A}$-transformable iff $f \in \mathscr{A}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$.*

Proof. Assume that f is $\mathscr{A}$-transformable of arity n. By applying Lemma 7.4.5 to $\{f\}$, there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $f \in H\mathscr{A}$ or $f \in H\left[\begin{smallmatrix}1 & 0\\ 0 & \alpha\end{smallmatrix}\right]\mathscr{A}$. This is equivalent to $(H^{-1})^{\otimes n} f \in \mathscr{A}$ or $(H^{-1})^{\otimes n} f \in \left[\begin{smallmatrix}1 & 0\\ 0 & \alpha\end{smallmatrix}\right]\mathscr{A}$. Since f is non-degenerate and symmetric, we can replace $\mathscr{A}$ in the previous expressions with $\mathscr{F}_{123}$. Now we consider the possible cases.

1. If $(H^{-1})^{\otimes n} f \in \mathscr{F}_1$, then a further transformation by $\frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right] \in \mathbf{O}_2(\mathbb{C})$ puts it into the canonical form of $\mathscr{A}_1$.
2. If $(H^{-1})^{\otimes n} f \in \mathscr{F}_2$, then it is already in the canonical form of $\mathscr{A}_1$.
3. If $(H^{-1})^{\otimes n} f \in \mathscr{F}_3$, then it is already of the equivalent form of $\mathscr{A}_2$ given by Lemma 7.4.8.
4. If $(H^{-1})^{\otimes n} f \in \left[\begin{smallmatrix}1 & 0\\ 0 & \alpha\end{smallmatrix}\right]\mathscr{F}_1$, then a further transformation by $\frac{1}{\sqrt{2}}\left[\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right] \in \mathbf{O}_2(\mathbb{C})$ puts it into the canonical form of $\mathscr{A}_1$.
5. If $(H^{-1})^{\otimes n} f \in \left[\begin{smallmatrix}1 & 0\\ 0 & \alpha\end{smallmatrix}\right]\mathscr{F}_2$, then it is already in the canonical form of $\mathscr{A}_3$.
6. If $(H^{-1})^{\otimes n} f \in \left[\begin{smallmatrix}1 & 0\\ 0 & \alpha\end{smallmatrix}\right]\mathscr{F}_3$, then it is of the form $\left[\begin{smallmatrix}1\\ \alpha^3\end{smallmatrix}\right]^{\otimes n} + i^r \left[\begin{smallmatrix}1\\ -\alpha^3\end{smallmatrix}\right]^{\otimes n}$ and a further transformation by $\left[\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right] \in \mathbf{O}_2(\mathbb{C})$ puts it into the canonical form of $\mathscr{A}_3$. To see this,

$$\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}^{\otimes n}\left(\begin{bmatrix}1\\ \alpha^3\end{bmatrix}^{\otimes n} + i^r \begin{bmatrix}1\\ -\alpha^3\end{bmatrix}^{\otimes n}\right) = \begin{bmatrix}-\alpha^3\\ 1\end{bmatrix}^{\otimes n} + i^r \begin{bmatrix}\alpha^3\\ 1\end{bmatrix}^{\otimes n}$$
$$= \left(-\alpha^3\right)^n \left(\begin{bmatrix}1\\ -\frac{1}{\alpha^3}\end{bmatrix}^{\otimes n} + (-1)^n i^r \begin{bmatrix}1\\ \frac{1}{\alpha^3}\end{bmatrix}^{\otimes n}\right)$$
$$= \left(-\alpha^3\right)^n \left(\begin{bmatrix}1\\ \alpha\end{bmatrix}^{\otimes n} + i^{2n+r} \begin{bmatrix}1\\ -\alpha\end{bmatrix}^{\otimes n}\right).$$

Let $\hat{f} = H^{\otimes n} f$ be f after the claimed orthogonal transformation. By examining each case separately, where $\hat{f}$ is expressed as the sum of two tensor powers, up to a global factor $\lambda \neq 0$, the following forms are possible:

1. $\hat{f} = \left[\begin{smallmatrix}1\\ 1\end{smallmatrix}\right]^{\otimes n} + \beta \left[\begin{smallmatrix}1\\ -1\end{smallmatrix}\right]^{\otimes n}$ where $\beta = \alpha^{tn+2r}$ for some $r \in \{0,1,2,3\}$ and $t \in \{0,1\}$ (type $\mathscr{A}_1$);
2. $\hat{f} = \left[\begin{smallmatrix}1\\ i\end{smallmatrix}\right]^{\otimes n} + \beta \left[\begin{smallmatrix}1\\ -i\end{smallmatrix}\right]^{\otimes n}$ where $\beta = i^r$ (type $\mathscr{A}_2$);
3. $\hat{f} = \left[\begin{smallmatrix}1\\ \alpha\end{smallmatrix}\right]^{\otimes n} + i^r \left[\begin{smallmatrix}1\\ -\alpha\end{smallmatrix}\right]^{\otimes n}$ for some $r \in \{0,1,2,3\}$ (type $\mathscr{A}_3$).

Conversely, if there exists a matrix $H \in \mathbf{O}_2(\mathbb{C})$ such that $H^{\otimes n} f$ is in one of the canonical forms of $\mathscr{A}_1$, $\mathscr{A}_2$, or $\mathscr{A}_3$, then one can directly check that f is $\mathscr{A}$-transformable. In fact, the transformations that we applied above are all invertible. $\square$

We also have a similar characterization for $\mathscr{P}$-transformable signatures. We define the stabilizer group of $\mathscr{P}$ similar to Stab($\mathscr{A}$). It is easy to see the left and right stabilizers coincide, which we denote by Stab($\mathscr{P}$). Furthermore, Stab($\mathscr{P}$) is generated by nonzero scalar multiples of matrices of the form $\left[\begin{smallmatrix} 1 & 0 \\ 0 & \nu \end{smallmatrix}\right]$ for any nonzero $\nu \in \mathbb{C}$ and $X = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$.

Lemma 7.4.11. *Let $\mathcal{F}$ be a set of signatures. Then $\mathcal{F}$ is $\mathscr{P}$-transformable iff there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $\mathcal{F} \subseteq H\mathscr{P}$ or $\mathcal{F} \subseteq H \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right] \mathscr{P}$.*

Proof. Sufficiency is easily verified by checking that $=_2$ is transformed into $\mathscr{P}$ in both cases. In particular, H leaves $=_2$ unchanged.

If $\mathcal{F}$ is $\mathscr{P}$-transformable, then by definition, there exists a matrix T such that $(=_2)T^{\otimes 2} \in \mathscr{P}$ and $T^{-1}\mathcal{F} \subseteq \mathscr{P}$. The non-degenerate binary signatures in $\mathscr{P}$ are either $[0,1,0]$ or of the form $[1,0,\nu]$, up to a scalar. However, notice that $[1,0,1] = [1,0,\nu] \left[\begin{smallmatrix} 1 & 0 \\ 0 & \nu^{-\frac{1}{2}} \end{smallmatrix}\right]^{\otimes 2}$ and $\left[\begin{smallmatrix} 1 & 0 \\ 0 & \nu^{-\frac{1}{2}} \end{smallmatrix}\right] \in \text{Stab}(\mathscr{P})$. Thus, we only need to consider $[1,0,1]$ and $[0,1,0]$. Now we apply Proposition 7.4.1.

1. If $(=_2)T^{\otimes 2} = [1,0,1]$, then by case 1 of Proposition 7.4.1, we have $T \in \mathbf{O}_2(\mathbb{C})$. Therefore $\mathcal{F} \subseteq H\mathscr{P}$ where $H = T \in \mathbf{O}(\mathbb{C})$.
2. If $(=_2)T^{\otimes 2} = [0,1,0]$, then by case 3 of Proposition 7.4.1, there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $T = \frac{1}{\sqrt{2}} H \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$. Therefore $\mathcal{F} \subseteq H \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right] \mathscr{P}$ where $H \in \mathbf{O}_2(\mathbb{C})$. $\square$

We also have similar definitions of the sets $\mathscr{P}_1$ and $\mathscr{P}_2$.

Definition 7.4.12. A symmetric signature f of arity n is in $\mathscr{P}_1$ if there exists $H \in \mathbf{O}_2(\mathbb{C})$ and a nonzero constant $c \in \mathbb{C}$ such that $f = cH^{\otimes n}\left(\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]^{\otimes n} + \beta \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]^{\otimes n}\right)$, where $\beta \neq 0$.

When such an H exists, we say that $f \in \mathscr{P}_1$ with transformation H. If $f \in \mathscr{P}_1$ with I_2, then we say f is in the canonical form of $\mathscr{P}_1$. If f is in the canonical form of $\mathscr{P}_1$, then by Lemma 7.4.2, for any $0 \le k \le n-2$, we have $f_{k+2} = f_k$. Since f is non-degenerate, $f_1 \neq \pm f_0$ is implied.

It is easy to check that $\mathscr{A}_1 \subset \mathscr{P}_1$. The corresponding definition for $\mathscr{P}_2$ coincides with Definition 7.4.7 for $\mathcal{A}_2$. In other words, we define $\mathscr{P}_2 = \mathcal{A}_2$.

Now we characterize the $\mathscr{P}$-transformable signatures as we did for the $\mathscr{A}$-transformable signatures in Lemma 7.4.10.

Lemma 7.4.13. *Let f be a non-degenerate symmetric signature. Then f is $\mathscr{P}$-transformable iff $f \in \mathscr{P}_1 \cup \mathscr{P}_2$.*

Proof. Assume that f is $\mathscr{P}$-transformable of arity n. By applying Lemma 7.4.11 to $\{f\}$, there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $f \in H\mathscr{P}$ or $f \in H \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right] \mathscr{P}$. This is equivalent to $(H^{-1})^{\otimes n} f \in \mathscr{P}$ or $(H^{-1})^{\otimes n} f \in \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right] \mathscr{P}$.

The symmetric signatures in $\mathscr{P}$ take the form $[0, 1, 0]$, or $[a, 0, \ldots, 0, b] = a[1, 0]^{\otimes n} + b[0, 1]^{\otimes n}$, where $ab \neq 0$ since f is non-degenerate. Now we consider the possible cases.

1. If $(H^{-1})^{\otimes n} f = \left[\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right] = \frac{1}{2i}\left(\left[\begin{smallmatrix} 1 \\ i \end{smallmatrix}\right]^{\otimes 2} - \left[\begin{smallmatrix} 1 \\ -i \end{smallmatrix}\right]^{\otimes 2}\right)$, then it is already of the equivalent form of $\mathscr{P}_2 = \mathscr{A}_2$ given by Lemma 7.4.8.
2. If $(H^{-1})^{\otimes n} f = a\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]^{\otimes n} + b\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]^{\otimes n}$, then a further transformation by $\frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right] \in \mathbf{O}_2(\mathbb{C})$ puts it into the canonical form of $\mathscr{P}_1$.
3. If $(H^{-1})^{\otimes n} f = \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]^{\otimes 2}\left[\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right] = 2\left[\begin{smallmatrix} 1 \\ 0 \\ 1 \end{smallmatrix}\right] = \left(\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]^{\otimes 2} + \left[\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right]^{\otimes 2}\right)$, then it is already in the canonical form of $\mathscr{P}_1$.
4. If $(H^{-1})^{\otimes n} f = \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]^{\otimes n}\left(a\left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]^{\otimes n} + b\left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right]^{\otimes n}\right)$, then it is already of the equivalent form of $\mathscr{P}_2 = \mathscr{A}_2$ given by Lemma 7.4.8.

Conversely, if there exists a matrix $H \in \mathbf{O}_2(\mathbb{C})$ such that $H^{\otimes n} f$ is in one of the canonical forms of $\mathscr{P}_1$ or $\mathscr{P}_2$, then one can directly check that f is $\mathscr{P}$-transformable. In fact, the transformations that we applied above are all invertible. □

Combining Lemma 7.4.10 and Lemma 7.4.13, we have a necessary and sufficient condition for a single non-degenerate signature to be $\mathscr{A}$- or $\mathscr{P}$-transformable.

Corollary 7.4.14. *Let f be a non-degenerate signature. Then f is $\mathscr{A}$- or $\mathscr{P}$-transformable iff $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$.*

Notice that our definitions of $\mathscr{P}_1$, $\mathscr{P}_2$, and $\mathscr{A}_3$ each involve an orthogonal transformation. For any single signature $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$, Holant$(f)$ is tractable. However, this does not imply that

Holant($\mathscr{P}_1$), Holant($\mathscr{P}_2$), or Holant($\mathscr{A}_3$) is tractable. One can check, using Theorem 7.2.1, that Holant($\mathscr{P}_2$) is tractable while Holant($\mathscr{P}_1$) and Holant($\mathscr{A}_3$) are #P-hard.

7.4.2 Dichotomies when $\mathscr{A}$- or $\mathscr{P}$-transformable Signatures Appear

In this subsection, we prove three dichotomies when $\mathscr{A}$- or $\mathscr{P}$-transformable signatures appear. First though, we prove an interpolation result.

Lemma 7.4.15. *Let $a, b \in \mathbb{C}$. If $ab \neq 0$, then for any set $\mathcal{F}$ of complex-weighted signatures containing $[a, 0, \dots, 0, b]$ of arity $r \geq 3$,*

$$\operatorname{Holant}(\mathcal{F} \cup \mathcal{EQ}_2) \leq_T \operatorname{Holant}(\mathcal{F}).$$

The reduction also holds when restricted to planar graphs.

Proof. Since $a \neq 0$, we can normalize the first entry to get $[1, 0, \dots, 0, x]$, where $x \neq 0$. First, we show how to obtain an arity 4 generalized equality signature. If $r = 3$, then we connect two copies together by a single edge to get an arity 4 signature. For larger arities, we form self-loops until realizing a signature of arity 3 or 4. By this process, we have a signature $g = [1, 0, 0, 0, y]$, where $y \neq 0$. If y is a pth root of unity, then we can directly realize $=_4$ by connecting p copies of g together, two edges at a time as in Figure 6.8. Otherwise, y is not a root of unity and we can interpolate $=_4$ as follows.

Consider an instance Ω of Holant($\mathcal{F} \cup \{=_4\}$). Suppose that $=_4$ appears n times in Ω. We stratify the assignments in Ω based on the assignments to $=_4$. We only need to consider the all-zero and all-one assignments since any other assignment contributes a factor of 0. Let i be the number of all-one assignments to $=_4$ in Ω. Then there are $n - i$ all-zero assignments and the Holant on Ω is

$$\operatorname{Holant}(\Omega) = \sum_{i=0}^{n} c_i,$$

where c_i is the sum over all such assignments of the product of evaluations of all other signatures on Ω.

We construct from Ω a sequence of instances Ω_s of $\text{Holant}(\mathcal{F})$ indexed by $s \geq 1$. We obtain Ω_s from Ω by replacing each occurrence of $=_4$ with a gadget g_s created from s copies of $[1,0,0,0,y]$, connecting two edges together at a time as in Figure 6.8. The Holant on Ω_s is

$$\text{Holant}(\Omega_s) = \sum_{i=0}^{n} (y^s)^i c_i.$$

For $s \geq 1$, this gives a coefficient matrix that is Vandermonde. Since y is neither 0 nor a root of unity, y^s is distinct for each s. Therefore, the Vandermonde system has full rank. We can solve for the unknowns c_i and obtain the value of $\text{Holant}(\Omega)$.

With $=_4$, it is easy to construct all equality signatures of even arity, so we are done. □

Our characterizations of $\mathscr{A}$-transformable signatures in Lemma 7.4.10 and $\mathscr{P}$-transformable signatures in Lemma 7.4.13 are up to transformations in $\mathbf{O}_2(\mathbb{C})$. Since an orthogonal transformation never changes the complexity of the problem, in the proofs of following lemmas, we assume any signature in $\mathscr{A}_i$ for $i = 1,2,3$, or $\mathscr{P}_j$ for $j = 1,2$, is already in the canonical form.

Lemma 7.4.16. *Let $\mathcal{F}$ be a set of symmetric signatures. Suppose $\mathcal{F}$ contains a non-degenerate signature $f \in \mathscr{P}_1$ of arity $n \geq 3$. Then* $\text{Holant}(\mathcal{F})$ *is #P-hard unless $\mathcal{F}$ is $\mathscr{P}$-transformable or $\mathscr{A}$-transformable.*

Proof. By assumption, for any $0 \leq k \leq n-2$, $f_{k+2} = f_k$ and $f_1 \neq \pm f_0$ since f is not degenerate. We can express f as

$$f = a_0 \begin{bmatrix} 1 \\ 1 \end{bmatrix}^{\otimes n} + a_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}^{\otimes n},$$

where $a_0 = (f_0 + f_1)/2$ and $a_1 = (f_0 - f_1)/2$. For this f, we can further perform an orthogonal transformation by $H_2 = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$ so that f is transformed into the generalized equality signature $2^{n/2}[a_0, 0, \ldots, 0, a_1]$ of arity n, where $a_0 a_1 \neq 0$. By Lemma 7.4.15, we can realize any equality signature of even arity. Thus, $\#\text{CSP}^2(H_2\mathcal{F}) \leq_T \text{Holant}(\mathcal{F})$.

Now we apply Theorem 7.1.1, the $\#\text{CSP}^d$ dichotomy, to the set $H_2\mathcal{F}$. If this problem is #P-hard, then $\text{Holant}(\mathcal{F})$ is #P-hard as well. Otherwise, this problem is $\#\text{CSP}^2$ tractable. Therefore, there exists some T of the form $\left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha^k \end{smallmatrix}\right]$, where the integer $k \in \{0, 1, \ldots, 7\}$, such that $TH_2\mathcal{F}$ is a subset of $\mathscr{A}$ or $\mathscr{P}$.

If $TH_2\mathcal{F} \subseteq \mathscr{P}$, then we have $H_2\mathcal{F} \subseteq T^{-1}\mathscr{P}$. Notice that $T \in \operatorname{Stab}(\mathscr{P})$, so $T^{-1}\mathscr{P} = \mathscr{P}$. Thus, $\mathcal{F}$ is $\mathscr{P}$-transformable under this H_2 transformation. Otherwise, $TH_2\mathcal{F} \subseteq \mathscr{A}$. It is easy to verify that $(=_2)((TH_2)^{-1})^{\otimes 2}$ is $[1, 0, i^{-k}] \in \mathscr{A}$. Thus, $\mathcal{F}$ is $\mathscr{A}$-transformable under this TH_2 transformation. $\square$

Lemma 7.4.17. *Let $\mathcal{F}$ be a set of symmetric signatures. Suppose $\mathcal{F}$ contains a non-degenerate signature $f \in \mathscr{P}_2$ of arity $n \geq 3$. Then* Holant($\mathcal{F}$) *is* #P-*hard unless $\mathcal{F}$ is $\mathscr{P}$-transformable or $\mathscr{A}$-transformable.*

Proof. By assumption, for any $0 \leq k \leq n-2$, $f_{k+2} = -f_k$ and $f_1 \neq \pm i f_0$ since f is not degenerate. We can express f as

$$f = a_0 \begin{bmatrix} 1 \\ i \end{bmatrix}^{\otimes n} + a_1 \begin{bmatrix} 1 \\ -i \end{bmatrix}^{\otimes n},$$

where $a_0 = (f_0 + i f_1)/2$ and $a_1 = (f_0 - i f_1)/2$, and $a_0 a_1 \neq 0$. Then under the holographic transformation $Z' = \left[\begin{smallmatrix} \sqrt[n]{a_0} & \sqrt[n]{a_1} \\ \sqrt[n]{a_0} i & -\sqrt[n]{a_1} i \end{smallmatrix}\right]^{-1}$, we have

$$Z'^{\otimes n} f = (=_n) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes n} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}^{\otimes n}$$

and

$$\begin{aligned} \operatorname{Holant}(=_2 \mid \mathcal{F} \cup \{f\}) &\equiv_T \operatorname{Holant}([1, 0, 1](Z'^{-1})^{\otimes 2} \mid Z'\mathcal{F} \cup \{Z'^{\otimes n} f\}) \\ &\equiv_T \operatorname{Holant}(2a_0a_1[0, 1, 0] \mid Z'\mathcal{F} \cup \{=_n\}). \end{aligned}$$

Thus, we have a bipartite graph with $=_n$ on the right and $(\neq_2) = [0, 1, 0]$ on the left up to a nonzero scalar, so all equality signatures of arity a multiple of n are realizable on the right side. To see this, first notice that we can move equality signatures from the right side to the left side using the binary disequality because the binary disequality just reverses signatures (i.e. exchanges

(a) Vertices assigned $[1, \alpha i, i, -\alpha]$. (b) Vertices assigned $[1, \alpha i, i, -\alpha, -1]$.

Figure 7.7: Constructions to realize $[1, 0, i]$.

the 0 and 1 input bits), which leaves the equality signatures unchanged. Now we do an induction. Suppose we can realize the equality $=_{(k-1)n}$ on the right side for some integer $k > 1$. We create a new signature on the right using one $=_{(k-1)n}$ and two $=_n$ on the right and one $=_n$ on the left. Since $n \geq 3$, we can connect one wire of the left $=_n$ to each of the three equality signatures on the right. The remaining wires of the left $=_n$ can be connected arbitrarily to the signatures on the right. The resulting signature is an equality of arity $(k-1)n + 2n - n = kn$. Since we have $=_{kn}$ on both sides for any integer $k \geq 1$, $\#\text{CSP}^n(Z'\mathcal{F}) \leq_T \text{Holant}(\mathcal{F})$.

Now we apply Theorem 7.1.1, the $\#\text{CSP}^d$ dichotomy, to the set $Z'\mathcal{F}$. If this problem is #P-hard, then $\text{Holant}(\mathcal{F})$ is #P-hard as well. Otherwise, this problem is $\#\text{CSP}^n$ tractable. Let ω be a primitive $4n$-th root of unity. Then under the holographic transformation $T = \left[\begin{smallmatrix} 1 & 0 \\ 0 & \omega^k \end{smallmatrix}\right]$ for some integer k, $TZ'\mathcal{F}$ is a subset of $\mathscr{A}$ or $\mathscr{P}$. If $TZ'\mathcal{F} \subseteq \mathscr{P}$, then we have $Z'\mathcal{F} \subseteq T^{-1}\mathscr{P}$. Notice that $T \in \text{Stab}(\mathscr{P})$, so $T^{-1}\mathscr{P} = \mathscr{P}$. Thus, $\mathcal{F}$ is $\mathscr{P}$-transformable under this Z' transformation.

Otherwise, $TZ'\mathcal{F} \subseteq \mathscr{A}$. It is easy to verify that $(=_2)((TZ')^{-1})^{\otimes 2}$ is $[0, 1, 0] \in \mathscr{A}$ up to a scalar. Thus, $\mathcal{F}$ is $\mathscr{A}$-transformable under this TZ' transformation. □

Lemma 7.4.18. *Let $\mathcal{F}$ be a set of symmetric signatures. Suppose $\mathcal{F}$ contains a non-degenerate signature $f \in \mathscr{A}_3$ of arity $n \geq 3$. Then* $\text{Holant}(\mathcal{F})$ *is #P-hard unless $\mathcal{F}$ is $\mathscr{A}$-transformable.*

Proof. By assumption, for any $0 \leq k \leq n-2$, we have $f_{k+2} = i f_k$. We can express f as

$$f = \lambda \left(\begin{bmatrix} 1 \\ \alpha \end{bmatrix}^{\otimes n} + i^r \begin{bmatrix} 1 \\ -\alpha \end{bmatrix}^{\otimes n} \right),$$

for some integer r.

A self loop on f yields f', where $f'_k = f_k + f_{k+2} = (1+i) f_k$. Thus up to the constant $(1+i)$,

f' is just the first $n-2$ entries of f. By doing more self loops, we eventually obtain a quaternary signature when n is even or a ternary one when n is odd. There are eight cases depending on the first two entries of f and the parity of n. However, for any case, we can realize the signature $[1,0,i]$. We list them here. (In the calculations below, we omit certain nonzero constant factors without explanation.)

- $[0,1,0,i]$: Another self loop gives $[0,1]$. Connect it back to the ternary to get $[1,0,i]$.
- $[1,0,i,0]$: Another self loop gives $[1,0]$. Connect it back to the ternary to get $[1,0,i]$.
- $[1,\alpha i,i,-\alpha]$: Another self loop gives $[1,\alpha i]$. Connect two copies of it to the ternary to get $[1,-\alpha]$. Then connect this back to the ternary to finally get $[1,0,i]$. See Figure 7.7a.
- $[1,-\alpha i,i,\alpha]$: Same construction as the previous case.
- $[0,1,0,i,0]$: Another self loop gives $[0,1,0]$. Connect it back to the quaternary to get $[1,0,i]$.
- $[1,0,i,0,-1]$: Another self loop gives $[1,0,i]$ directly.
- $[1,\alpha i,i,-\alpha,-1]$: Another self loop gives $[1,\alpha i,i]$. Connect two copies of it together to get $[1,-\alpha,-i]$. Connect this back to the quaternary to get $[1,0,i]$. See Figure 7.7b.
- $[1,-\alpha i,i,\alpha,-1]$: Same construction as the previous case.

With $[1,0,i]$ in hand, we can connect three copies to get $[1,0,-i]$. Now we construct a bipartite graph, with $\mathcal{F}\cup\{=_2\}$ on the right side and $[1,0,-i]$ on the left, and do a holographic transformation by $Z=\left[\begin{smallmatrix}\alpha & 1\\ -\alpha & 1\end{smallmatrix}\right]$ to get

$$\begin{aligned}\operatorname{Holant}([1,0,-i]\mid\mathcal{F}\cup\{f,=_2\}) &\equiv_T \operatorname{Holant}([1,0,-i](Z^{-1})^{\otimes 2}\mid Z\mathcal{F}\cup\{Z^{\otimes n}f, Z^{\otimes 2}(=_2)\})\\ &\equiv_T \operatorname{Holant}(\tfrac{1}{2i}[1,0,1]\mid Z\mathcal{F}\cup\{[1,0,\ldots,0,i^k],[1,-i,1]\})\\ &\equiv_T \operatorname{Holant}\left(Z\mathcal{F}\cup\{[1,0,\ldots,0,i^k],[1,-i,1]\}\right).\end{aligned}$$

Notice that f becomes $[1,0,\ldots,0,i^k]$ where $k=r+2n$ (after normalizing the first entry) and $=_2$ becomes $[1,-i,1]$. On the other side, $[1,0,-i]$ becomes $[1,0,1]$. By Lemma 7.4.15, we can realize any equality signature of even arity. Thus, $\#\mathrm{CSP}^2(Z\mathcal{F}\cup\{[1,-i,1]\})\le_T \operatorname{Holant}(\mathcal{F})$.

Now we apply Theorem 7.1.1, the $\#\mathrm{CSP}^d$ dichotomy, to the set $Z\mathcal{F}\cup\{[1,-i,1]\}$. If this problem is #P-hard, then $\operatorname{Holant}(\mathcal{F})$ is #P-hard as well. Otherwise, this problem is $\#\mathrm{CSP}^2$ tractable.

Therefore, there exists some T of the form $\left[\begin{smallmatrix} 1 & 0 \\ 0 & \alpha^d \end{smallmatrix}\right]$, where the integer $d \in \{0, 1, \ldots, 7\}$, such that $TZ\mathcal{F} \cup \{T^{\otimes 2}[1, -i, 1]\}$ is a subset of $\mathscr{A}$ or $\mathscr{P}$.

However, $T^{\otimes 2}[1, -i, 1]$ can never be in $\mathscr{P}$. Thus $TZ\mathcal{F} \cup \{T^{\otimes 2}[1, -i, 1]\} \subseteq \mathscr{A}$. Further notice that if $d \in \{1, 3, 5, 7\}$ in the expression of T, then $T^{\otimes 2}[1, -i, 1]$ is not in $\mathscr{A}$. Hence, T must be of the form $\left[\begin{smallmatrix} 1 & 0 \\ 0 & i^d \end{smallmatrix}\right]$, where the integer $d \in \{0, 1, 2, 3\}$. For such T, $T^{\otimes 2}[1, -i, 1] \in \mathscr{A}$, and $T^{-1}\mathscr{A} = \mathscr{A}$ as $T \in \operatorname{Stab}(\mathscr{A})$. Thus, $TZ\mathcal{F} \cup \{T^{\otimes 2}[1, -i, 1]\} \subseteq \mathscr{A}$ simply becomes $Z\mathcal{F} \subseteq \mathscr{A}$. Moreover, $(=_2)(Z^{-1})^{\otimes 2}$ is $[1, i, 1] \in \mathscr{A}$. Therefore, $\mathcal{F}$ is $\mathscr{A}$-transformable under this Z transformation. □

7.5 Main Result

In this section, we prove our main dichotomy theorem. We begin with a dichotomy for a single signature, which we prove by induction on its arity.

Theorem 7.5.1. *If f is a non-degenerate symmetric signature of arity at least 3 with complex weights in Boolean variables, then* Holant(f) *is* #P*-hard unless $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$ or f is vanishing, in which case the problem is computable in polynomial time.*

Recall that $\mathscr{A}_1 \subset \mathscr{P}_1$ and $\mathscr{A}_2 = \mathscr{P}_2$, and $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$ iff f is $\mathscr{A}$-transformable or $\mathscr{P}$-transformable by Corollary 7.4.14.

Proof. Let the arity of f be n. The base cases of $n = 3$ and $n = 4$ are proved in Theorem 6.1.3 and Theorem 6.6.3 respectively. Now assume $n \geq 5$.

With the signature f, we form a self loop to get a signature f' of arity at least 3. We consider the cases separately whether f' is degenerate or not.

- Suppose $f' = [a, b]^{\otimes(n-2)}$ is degenerate. There are three cases to consider.
    1. If $a = b = 0$, then f' is the all zero signature. For f, this means $f_{k+2} = -f_k$ for $0 \leq k \leq n - 2$, so $f \in \mathscr{P}_2$ by Lemma 7.4.8, and therefore Holant(f) is tractable.
    2. If $a^2 + b^2 \neq 0$, then f' is nonzero and $[a, b]$ is not a constant multiple of either $[1, i]$ or $[1, -i]$. We may normalize so that $a^2 + b^2 = 1$. Then the orthogonal transformation $\left[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right]$ transforms the column vector $[a, b]$ to $[1, 0]$. Let $\hat{f}$ be the transformed signature from f, and $\widehat{f'} = [1, 0]^{\otimes(n-2)}$ the transformed signature from f'.

Since an orthogonal transformation keeps $=_2$ invariant, this transformation commutes with the operation of taking a self loop, i.e., $\widehat{f'} = (\hat{f})'$. Here $(\hat{f})'$ is the function obtained from $\hat{f}$ by taking a self loop. So $\hat{f}_0 + \hat{f}_2 = 1$ and for every integer $1 \le k \le n-2$, we have $\hat{f}_k = -\hat{f}_{k+2}$. With one or more self loops, we eventually obtain either $[1,0]$ when n is odd or $[1,0,0]$ when n is even. In either case, we connect an appropriate number of copies of this signature to $\hat{f}$ to get a arity 4 signature $\hat{g} = [\hat{f}_0, \hat{f}_1, \hat{f}_2, -\hat{f}_1, -\hat{f}_2]$. We show that Holant$(\hat{g})$ is #P-hard. To see this, we first compute $\det(\widetilde{M_g}) = -2(\hat{f}_0 + \hat{f}_2)(\hat{f}_1^2 + \hat{f}_2^2) = -2(\hat{f}_1^2 + \hat{f}_2^2)$, since $\hat{f}_0 + \hat{f}_2 = 1$. Therefore if $\hat{f}_1^2 + \hat{f}_2^2 \neq 0$, Holant$(\hat{g})$ is #P-hard by Lemma 6.5.4. Otherwise $\hat{f}_1^2 + \hat{f}_2^2 = 0$, and we consider $\hat{f}_2 = i\hat{f}_1$ since the other case is similar. Since f is non-degenerate, $\hat{f}$ is non-degenerate, which implies $\hat{f}_2 \neq 0$. We can express $\hat{g}$ as $[1,0]^{\otimes 4} - \hat{f}_2[1,i]^{\otimes 4}$. Under the holographic transformation by $T = \left[\begin{smallmatrix} 1 & (-\hat{f}_2)^{1/4} \\ 0 & i(-\hat{f}_2)^{1/4} \end{smallmatrix}\right]$, we have

$$\begin{aligned} \text{Holant}(=_2 \mid \hat{g}) &\equiv_T \text{Holant}([1,0,1]T^{\otimes 2} \mid (T^{-1})^{\otimes 4}\hat{g}) \\ &\equiv_T \text{Holant}(\hat{h} \mid =_4), \end{aligned}$$

where

$$\hat{h} = [1,0,1]T^{\otimes 2} = [1, (-\hat{f}_2)^{1/4}, 0]$$

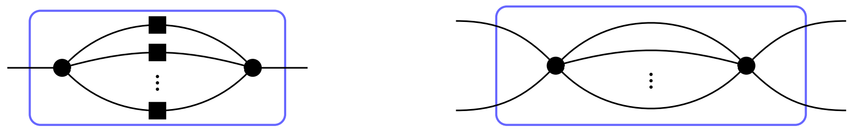
and $\hat{g}$ is transformed by T^{-1} into the arity 4 equality $=_4$, since

$$T^{\otimes 4}\left(\begin{bmatrix}1\\0\end{bmatrix}^{\otimes 4} + \begin{bmatrix}0\\1\end{bmatrix}^{\otimes 4}\right) = \begin{bmatrix}1\\0\end{bmatrix}^{\otimes 4} - \hat{f}_2\begin{bmatrix}1\\i\end{bmatrix}^{\otimes 4} = \hat{g}.$$

By Theorem 6.1.1, Holant$(\hat{h} \mid =_4)$ is #P-hard as $\hat{f}_2 \neq 0$.

3. If $a^2 + b^2 = 0$ but $(a,b) \neq (0,0)$, then $[a,b]$ is a nonzero multiple of $[1, \pm i]$. Ignoring the constant multiple, we have $f' = [1,i]^{\otimes(n-2)}$ or $[1,-i]^{\otimes(n-2)}$. We consider the first case since the other case is similar.

   In the first case, the characteristic polynomial of the recurrence relation of f' is $x - i$,



(a) The circles are assigned $\hat{f}$ and the squares are assigned $\neq_2$.

(b) The circles are assigned f.

Figure 7.8: Two gadgets used when $f' = [1, \pm i]^{\otimes(n-2)}$.

so that of f is $(x-i)(x^2+1) = (x-i)^2(x+i)$. Hence there exist a_0, a_1, and c such that

$$f_k = (a_0 + a_1 k) i^k + c(-i)^k$$

for every integer $0 \leq k \leq n$. Let f^+ and f^- be two signatures of arity n such that $f_k^+ = (a_0 + a_1 k) i^k$ and $f_k^- = c(-i)^k$ for every $0 \leq k \leq n$. Hence $f_k = f_k^+ + f_k^-$ and we write $f = f^+ + f^-$. If $a_1 = 0$, then f' is the all zero signature, a contradiction. If $c = 0$, then f is vanishing, one of the tractable cases. Now we assume $a_1 c \neq 0$ and show that $\operatorname{Holant}(f)$ is #P-hard. Hence $\operatorname{rd}^+(f^+) = 1$ and $\operatorname{rd}^-(f^-) = 0$. Under the holographic transformation $Z = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$, we have

$$\begin{aligned} \operatorname{Holant}(=_2 \mid f) &\equiv_T \operatorname{Holant}([1,0,1] Z^{\otimes 2} \mid (Z^{-1})^{\otimes n} f) \\ &\equiv_T \operatorname{Holant}([0,1,0] \mid \hat{f}), \end{aligned}$$

where $\hat{f}$ takes the form $[\hat{f}_0, \hat{f}_1, 0, \ldots, 0, c']$ with $c' = 2^{n/2} c \neq 0$ and $\hat{f}_1 \neq 0$, since $\hat{f}$ is the Z^{-1}-transformation of the sum of f^+ and f^-, with $\operatorname{rd}^+(f^+) = 1$ and $\operatorname{rd}^-(f^-) = 0$ respectively. On the other side, $(=_2) = [1,0,1]$ is transformed into $(\neq_2) = [0,1,0]$. Now consider the gadget in Figure 7.8a with $\hat{f}$ assigned to both vertices. This gadget has the binary signature $[0, c\hat{f}_0, 2c\hat{f}_1]$, which is equivalent to $[0, \hat{f}_0, 2\hat{f}_1]$ since $c \neq 0$. Translating back by Z to the original setting, this signature is $g = [\hat{f}_0 + \hat{f}_1, -i\hat{f}_1, \hat{f}_0 - \hat{f}_1]$. This can be verified as

$$\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} 0 & \hat{f}_0 \\ \hat{f}_0 & 2\hat{f}_1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}^{\intercal} = 2 \begin{bmatrix} \hat{f}_0 + \hat{f}_1 & -i\hat{f}_1 \\ -i\hat{f}_1 & \hat{f}_0 - \hat{f}_1 \end{bmatrix}.$$

Since $\hat{f}_1 \neq 0$, it can be directly checked that $g \notin \mathscr{R}_2^+$.

If $\hat{f}_0 \neq 0$, then g is non-degenerate. In this case we construct some function in $\mathscr{V}^+$. We connect f' back to f, getting a binary signature $p = Z^{\otimes 2}[0, 0, c']$. Then we connect p to f, the resulting signature is $p' = Z^{\otimes n-2}[\hat{f}_0, \hat{f}_1, 0, \ldots, 0]$ of arity $n - 2 \geq 3$ up to the constant factor of $c' \neq 0$. Notice that p' is non-degenerate and $p' \in \mathscr{V}^+$. By Lemma 7.3.3, Holant$(\{p', g\})$ is #P-hard, hence Holant(f) is also #P-hard.

Otherwise suppose $\hat{f}_0 = 0$. Then we have $g = [1, -i]^{\otimes 2}$ after ignoring the nonzero factor $\hat{f}_1$. Connecting this degenerate signature to f, we get a signature $h = \langle f, g \rangle$. We note that g annihilates the signature $f^- = c[1, -i]^{\otimes n}$, and thus $h = \langle f^+, g \rangle$. Then $\text{rd}^+(f^+) = 1$, $\text{vd}^+(g) = 0$, and we can apply Lemma 4.4.13. It follows that $\text{rd}^+(h) = 1$ and $\text{arity}(h) \geq 3$. This implies that h is non-degenerate and $h \in \mathscr{V}^+$.

Moreover, assigning f to both vertices in the gadget of Figure 7.8b, we get a non-degenerate signature $h' \in \mathscr{V}^-$ of arity 4. To see this, consider this gadget after a holographic transformation by Z. In this bipartite setting, it is the same as assigning $\hat{f} = [0, \hat{f}_1, 0, \ldots, 0, c]$ (or equivalently $[0, 1, 0, \ldots, 0, c'']$, where $c'' = c/\hat{f}_1 \neq 0$) to both the circle and triangle vertices in the gadget of Figure 7.5a. The square vertices there are still assigned $(\neq_2) = [0, 1, 0]$. While it is not apparent from the gadget's geometry, this signature is in fact symmetric. In particular, its values on inputs 1010 and 1100 are both 0. The resulting signature is $\widehat{h'} = (Z^{-1})^{\otimes 4} h' = [0, 0, 0, c'', 0]$. Hence $\text{rd}^-(h') = 1$, and therefore h' is non-degenerate and $h' \in \mathscr{V}^-$.

By Lemma 7.3.6, Holant$(\{h, h'\})$ is #P-hard, hence Holant(f) is also #P-hard.

- Suppose f' is non-degenerate. If f' is not in one of the tractable cases, then Holant(f') is #P-hard and so is Holant(f). We now assume Holant(f') is not #P-hard. Then, by inductive hypothesis, $f' \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$ or f' is vanishing. If $f' \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$, then applying Lemma 7.4.16, Lemma 7.4.17, or Lemma 7.4.18 to f' and the set $\{f, f'\}$, we either have that $\operatorname{Holant}(\{f, f'\})$ is #P-hard, so $\operatorname{Holant}(f)$ is #P-hard as well, or that f is $\mathscr{A}$- or $\mathscr{P}$-transformable, so by Corollary 7.4.14, $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$.

Otherwise, f' is vanishing, so $f' \in \mathscr{V}^\sigma$ for $\sigma \in \{+, -\}$ by Theorem 4.4.12. For simplicity, assume that $f' \in \mathscr{V}^+$. The other case is similar. Let $\operatorname{rd}^+(f') = d - 1$, where $2d < n$ and $d \geq 2$ since f' is non-degenerate. Then the entries of f' can be expressed as

$$f'_k = i^k q(k),$$

where $q(x)$ is a polynomial of degree exactly $d - 1$. However, notice that if f' satisfies some recurrence relation with characteristic polynomial $t(x)$, then f satisfies a recurrence relation with characteristic polynomial $(x^2 + 1)t(x)$. In this case, $t(x) = (x - i)^d$. Then the corresponding characteristic polynomial of f is $(x - i)^{d+1}(x + i)$, and thus the entries of f are

$$f_k = i^k p(k) + c(-i)^k$$

for some constant c and a polynomial $p(x)$ of degree at most d. However, the degree of $p(x)$ is exactly d, otherwise the polynomial $q(x)$ for f' would have degree less than $d - 1$. If $c = 0$, then f is vanishing, a tractable case. Otherwise, $c \neq 0$, and we show the problem is #P-hard. Under the transformation $Z = \frac{1}{\sqrt{2}} \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$, we have

$$\begin{aligned} \operatorname{Holant}(=_2 \mid f) &\equiv_T \operatorname{Holant}([1, 0, 1] Z^{\otimes 2} \mid (Z^{-1})^{\otimes n} f) \\ &\equiv_T \operatorname{Holant}([0, 1, 0] \mid \hat{f}), \end{aligned}$$

where $\hat{f} = [\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0, c]$, with $\hat{f}_d \neq 0$. Taking a self loop in the original setting is equivalent to connecting $[0, 1, 0]$ to a signature after this transformation. Thus, doing this once on $\hat{f}$, we can get $\widehat{f'} = [\hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0]$ corresponding to f', and doing this $d - 2$ times on $\hat{f}$, we get a signature $\hat{h} = [\hat{f}_{d-2}, \hat{f}_{d-1}, \hat{f}_d, 0, \ldots, 0, 0/c]$ of arity $n - 2(d - 2) = n - 2d + 4$. The last entry is c when $d = 2$ and is 0 when $d > 2$.

As $n > 2d$, we may do two more self loops and get $[\hat{f}_d, 0, \ldots, 0]$ of arity $k = n - 2d$. Now connect this signature back to $\hat{f}$ via $[0, 1, 0]$. It is the same as getting the last $n - k + 1 = 2d + 1$ signature entries of $\hat{f}$. We may repeat this operation zero or more times until the arity k' of the resulting signature is less than or equal to k. We claim that this signature has the form $\hat{g} = [0, \ldots, 0, c]$. In other words, the $k' + 1$ entries of $\hat{g}$ consist of the last c and k' many 0's in the signature $\hat{f}$, all appearing after $\hat{f}_d$. This is because there are $n - d - 1$ many 0 entries in the signature $\hat{f}$ after $\hat{f}_d$, and $n - d - 1 \geq k \geq k'$.

Translating back by the Z transformation, having both $[\hat{f}_d, 0, \ldots, 0]$ of arity k and $\hat{g} = [0, \ldots, 0, c]$ of arity k' is equivalent to, in the original setting, having both $[1, i]^{\otimes k}$ and $[1, -i]^{\otimes k'}$. If $k > k'$, then we can connect $[1, -i]^{\otimes k'}$ to $[1, i]^{\otimes k}$ and get $[1, i]^{\otimes (k-k')}$. Replacing k by $k - k'$, we can repeat this process until the new $k \leq k'$. If the new $k < k'$, then we can continue as in the subtractive Euclid algorithm. We continue this procedure and eventually we get $[1, i]^{\otimes t}$ and $[1, -i]^{\otimes t}$, where $t = \gcd(k, k')$, where $k = n - 2d$ and $k' \leq k$, as defined in the previous paragraph. Now putting k/t many copies of $[1, -i]^{\otimes t}$ together, we get $[1, -i]^{\otimes k}$.

In the transformed setting, $[1, -i]^{\otimes k}$ is $[0, \ldots, 0, 1]$ of arity k. Then we connect this back to $\hat{h}$ via $[0, 1, 0]$. Doing this is the same as forcing k connected edges of h to be assigned 0, because $[0, 1, 0]$ flips the assigned value 1 in $[0, \ldots, 0, 1]$ to 0. Thus we get a signature of arity $n - 2d + 4 - k = 4$, which is $[\hat{f}_{d-2}, \hat{f}_{d-1}, \hat{f}_d, 0, 0]$. Note that the last entry is 0 (and not c). Then we are done by Corollary 6.5.7 after transforming back to the original setting. □

Now we are ready to prove of our main theorem.

Proof of hardness for Theorem 7.2.1. Assume that Holant($\mathcal{F}$) is not #P-hard. If all of the non-degenerate signatures in $\mathcal{F}$ are of arity at most 2, then the problem is tractable case 1. Otherwise we have some non-degenerate signatures of arity at least 3. For any such f, by Theorem 7.5.1, $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$ or f is vanishing. If any of them is in $\mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$, then by Lemma 7.4.16, Lemma 7.4.17, or Lemma 7.4.18, we have that $\mathcal{F}$ is $\mathscr{A}$- or $\mathscr{P}$-transformable, which are tractable cases 2 and 3.

Now we assume that all non-degenerate signatures of arity at least 3 in $\mathcal{F}$ are vanishing, and there is a nonempty set of such signatures in $\mathcal{F}$. By Lemma 7.3.6, they must all be in $\mathscr{V}^\sigma$ with the same $\sigma \in \{+, -\}$. By Lemma 7.3.3, we know that any non-degenerate binary signature in $\mathcal{F}$ has to be in $\mathscr{R}_2^\sigma$. Furthermore, if $\mathcal{F}$ contains an $f \in \mathscr{V}^\sigma$ such that $\mathrm{rd}^\sigma(f) \geq 2$, then by Lemma 7.3.1, the only unary signatures allowed in $\mathcal{F}$ are some multiple of $[1, \sigma i]$, and all degenerate signatures in $\mathcal{F}$ are a tensor product of some multiple of $[1, \sigma i]$. Thus, all non-degenerate signatures of arity at least 3 as well as all degenerate signatures belong to $\mathscr{V}^\sigma$, and all non-degenerate binary signatures belong to $\mathscr{R}_2^\sigma$. This is tractable case 4.

Finally, we have the following: (i) all non-degenerate signatures of arity at least 3 in $\mathcal{F}$ belong to $\mathscr{V}^\sigma$; (ii) all signatures $f \in \mathcal{F} \cap \mathscr{V}^\sigma$ have $\mathrm{rd}^\sigma(f) \leq 1$, which implies that $f \in \mathscr{R}_2^\sigma$; and (iii) all non-degenerate binary signatures in $\mathcal{F}$ belong to $\mathscr{R}_2^\sigma$. Hence all non-degenerate signatures in $\mathcal{F}$ belong to $\mathscr{R}_2^\sigma$. All unary signatures also belong to $\mathscr{R}_2^\sigma$ by definition. This is indeed tractable case 5. The proof is complete. □

If $\mathcal{F}$ is finite, then the criterion of Theorem 7.2.1 is decidable in polynomial time [31, 34].

7.6 Closing Thoughts

A planar dichotomy implies a nonplanar dichotomy This dichotomy was generalized by restricting to planar graphs [27]. We were surprised to find new cases that are planar tractable but *not* by holographic reductions to matchgates and ultimately the FKT algorithm.

A dichotomy for Pl-Holant($\mathcal{F}$) implies a dichotomy for Holant($\mathcal{F}$). This is because Holant($\mathcal{F}$) can be viewed as the special case of Pl-Holant($\mathcal{F}$) when $\mathcal{F}$ contains this "crossover signature" X, which has signature matrix

$$M_X = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Clearly Pl-Holant($\mathcal{F}$) $\leq_T$ Holant($\mathcal{F}$). For the other direction, take an instance Ω of Holant($\mathcal{F}$) and embed the underlying graph in the plane so that at most two edges cross at any point. At each of these points, add a vertex assigned X to replace these crossed edges. This does not change the Holant value and is now an instance of Pl-Holant($\mathcal{F}$) since $X \in \mathcal{F}$.

Chapter 8

Dichotomy for #CSP over Planar Graphs

We prove a dichotomy theorem for symmetric complex-weighted Boolean #CSP when the constraint graph of the input must be planar. The problems that are #P-hard over general graphs but tractable over planar graphs are precisely those with a holographic reduction to matchgates. This generalizes a theorem [46] to the case of real weights. The main idea is to use the popular pinning technique, but because of the planarity restriction, this only possible in the Hadamard basis. This work was published in [73, 74].

8.1 Background

In 1979, Valiant [125] defined the class #P to explain the apparent intractability of counting perfect matchings in a graph. Yet over a decade earlier, Kasteleyn [89, 88] gave a polynomial-time algorithm to compute this quantity for planar graphs. This was an important milestone in a decades-long research program by physicists in statistical mechanics to determine what problems the restriction to the planar setting renders tractable [5, 83, 109, 145, 146, 97, 120, 87, 89, 68, 88, 101, 102, 141]. More recently, Valiant introduced matchgates [128, 127] and *holographic* algorithms [132, 131] that rely on Kasteleyn's algorithm to solve certain counting problems over planar graphs. Subsequently, the signatures with holographic transformations to matchgates were characterized (as described in Section 4.3).

From the viewpoint of computational complexity, we seek to understand exactly which intractable problems the planarity restriction enable us to efficiently compute. Partial answers to this question have been given in the context of various counting frameworks [138, 36], including the dichotomy from Chapter 5 and a dichotomy for symmetric complex-weighted Boolean #CSP in [46]. In every case, the problems that are #P-hard over general graphs but tractable over planar graphs are essentially those given in Section 4.3. In this chapter, we give more evidence for this phenomenon by extending the dichotomy in [46] from real to complex weights. This extension also generalizes the dichotomy in [53] by restricting to planar graphs.

Our main result is stated as follows.

Theorem 8.1.1. *Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures in Boolean variables. Then* Pl-#CSP($\mathcal{F}$) *is* #P*-hard unless $\mathcal{F}$ satisfies one of the following conditions, in which case it is tractable:*

1. #CSP($\mathcal{F}$) *is tractable (cf. [53]); or*
2. *$\widehat{\mathcal{F}}$ is realizable by matchgates (cf. Section 4.3).*

A more explicit description of the tractable cases can be found in Theorem 8.5.3.

Although this theorem is stated for the framework of counting constraint satisfaction problems (#CSP), our proof is in the more general framework of Holant problems. From the Holant perspective, the set $\mathcal{EQ} = \{=_n \mid n \geq 1\}$ of Equality signatures is always available in #CSP($\mathcal{F}$). By the signature theory of Cai and Lu [43], the Hadamard matrix $H = \left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$ defines the only[1] holographic transformation under which $\mathcal{EQ}$ becomes matchgate realizable. Let $\widehat{\mathcal{F}}$ denote $H\mathcal{F}$ for any set $\mathcal{F}$ of signatures. Then $\widehat{\mathcal{EQ}} = \{[1,0],[1,0,1],[1,0,1,0],\dots\}$ is the set of unweighted signatures with even support. Therefore, #CSP($\mathcal{F}$), which is equivalent to Holant($\mathcal{F} \cup \mathcal{EQ}$) by Lemma 2.2.1, is further equivalent to Holant($\widehat{\mathcal{F}} \cup \widehat{\mathcal{EQ}}$) by Lemma 3.2.2.

In many previous dichotomy theorems for Boolean #CSP($\mathcal{F}$), the proof of hardness began by pinning. The goal of this technique is to realize the constant functions [1, 0] and [0, 1] and was

[1] Up to transformations under which matchgates are closed.

always achieved by a *nonplanar* reduction. In the nonplanar setting, $[1, 0]$ and $[0, 1]$ are contained in each of the maximal tractable sets. Therefore, pinning in this setting does not imply the collapse of any complexity classes. However, the signatures in $\mathcal{EQ} \cup \{[1, 0], [0, 1]\}$ are not simultaneously realizable by matchgates. If we are to prove our main theorem, one should not expect to be able to pin for Pl-#CSP($\mathcal{F}$), since otherwise #P would collapse to P. Instead, apply the Hadamard transformation and consider Pl-Holant($\widehat{\mathcal{F}} \cup \widehat{\mathcal{EQ}}$). In this Hadamard basis, pinning becomes possible again since $[1, 0]$ and $[0, 1]$ are included in every maximal tractable set. Indeed, we prove our pinning result in this Hadamard basis in Section 8.4. The pinned version of Pl-Holant($\mathcal{F} \cup \widehat{\mathcal{EQ}}$), which assumes that the signatures $[1, 0]$ and $[0, 1]$ are available, is denoted by Pl-Holant^c($\mathcal{F} \cup \widehat{\mathcal{EQ}}$) (where c stands for *constants*).

Our other main technique is domain pairing. This technique is used to overcome a parity restriction. If all signatures have even arity, then the signature of any gadget will also have even arity. Domain pairing is a type of reduction that allows us to realize a signature of odd arity given only signatures of even arity.

We use Theorem 6.1.1 with $\mathcal{G} = \mathcal{EQ}$, which is the special case of Pl-#CSP($\mathcal{F}$) when $\mathcal{F}$ contains a single binary signature. Furthermore, we perform a holographic transformation by the Hadamard matrix $H = \left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$. Under this transformation, it is easy to see that the conditions $f_0 f_2 = f_1^2$ and $f_0 f_2 = -f_1^2 \wedge f_0 = -f_2$ are invariant while the conditions $f_1 = 0$ and $f_0 = f_2$ map to each other. Therefore, by an apparent coincidence, the tractability conditions remain the same. To be clear, we restate Theorem 6.1.1 when restricted to planar graphs both before and after a holographic transformation by H with $\mathcal{G} = \mathcal{EQ}$.

Theorem 8.1.2 (Special case of Theorem 6.1.1)**.** *For any $f_0, f_1, f_2 \in \mathbb{C}$, both* Pl-Holant($[f_0, f_1, f_2] \mid \mathcal{EQ}$) *and* Pl-Holant($[f_0, f_1, f_2] \mid \widehat{\mathcal{EQ}}$) *are* #P*-hard unless one of the following conditions hold, in which case both problems are computable in polynomial time:*

1. $f_0 f_2 = f_1^2$*;*
2. $f_1 = 0$*;*
3. $f_0 f_2 = -f_1^2$ *and* $f_0 = -f_2$*;*
4. $f_0 = f_2$*.*

In the standard basis of the Pl-#CSP framework, the set $\widehat{\mathscr{M}}$ of signatures is tractable and consists of signatures with the following expressions.[2]

Theorem 8.1.3 (Special case of Theorem 4 in [42]). *A symmetric signature $[f_0, f_1, \ldots, f_n]$ is realizable under the basis $\left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$ iff it takes one of the following forms:*

1. *there exists constants $\lambda, \alpha, \beta \in \mathbb{C}$ and $\varepsilon = \pm 1$, such that for all ℓ, $0 \le \ell \le n$,*

$$f_\ell = \lambda[(\alpha+\beta)^{n-\ell}(\alpha-\beta)^\ell + \varepsilon(\alpha-\beta)^{n-\ell}(\alpha+\beta)^\ell];$$

2. *there exists a constant $\lambda \in \mathbb{C}$, such that for all ℓ, $0 \le \ell \le n$,*

$$f_\ell = \lambda(n-2\ell)(-1)^\ell;$$

3. *there exists a constant $\lambda \in \mathbb{C}$, such that for all ℓ, $0 \le \ell \le n$,*

$$f_\ell = \lambda(n-2\ell).$$

We note that case 1 corresponds to the general case ($\varepsilon = +1$ for signatures with even parity and $\varepsilon = -1$ for signatures with odd parity) while case 3 corresponds to the perfect matching signatures $[0, 1, 0, \ldots, 0]$ and case 2 corresponds to their reversals.

The tractability results for the Pl-#CSP framework are already known and given in Chapter 4. We state this tractability in the Hadamard basis with $[1, 0]$ and $[0, 1]$ present.

Theorem 8.1.4. *Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures in Boolean variables. Then* Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is tractable if $\mathcal{F} \subseteq \mathscr{A}$, $\mathcal{F} \subseteq \widehat{\mathscr{P}}$, or $\mathcal{F} \subseteq \mathscr{M}$.*

8.2 Domain Pairing

In this section, we discuss a technique called domain pairing, which pairs input variables to simulate a problem on a domain of size four and then reduces a problem in the Boolean domain to it. As

[2] Even though Theorem 8.1.3 is technically about generator signatures, neither generators nor recognizers are mentioned because Theorems 3 and 4 in [42] coincide when the basis is an orthogonal transformation.

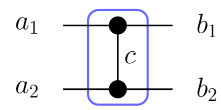


Figure 8.1: Gadget designed for the paired domain. One vertex is assigned $[1, 0, 1, 0]$ and the other is assigned $[x, 0, y, 0]$.

explained in the previous section, we work in the Hadamard basis instead of the standard basis. The goal then becomes a dichotomy for Pl-Holant($\mathcal{F} \cup \widehat{\mathcal{EQ}}$).

By a simple parity argument, gadgets constructed with signatures of even arity can only realize other signatures of even arity. In particular, this means that $=_4$ cannot by itself be used to construct $=_3$. Nevertheless, there is a clever argument that can realize $=_3$ using $=_4$. The catch is the domain changes from individual elements to pairs of elements. Thus, we call this reduction technique *domain pairing*. This technique was first used in the proof of Lemma III.2 in [46] with real weights. It was also used in the proof of Lemma 4.6 in [72] in the parity case and in Lemma IV.5 in [80] with real weights as well as grouping more than just two domain elements.

We prove a generalization of the domain pairing lemma for complex weights.

Lemma 8.2.1 (Domain pairing)**.** *Let $a, b, x, y \in \mathbb{C}$. Suppose $f = [x, 0, y, 0]$ and $g = [a, 0, \ldots, 0, b]$ with arity at least 3. If $aby \neq 0$ and $x^2 \neq y^2$, then* Pl-Holant($\{f, g\} \cup \widehat{\mathcal{EQ}}$) *is* #P*-hard.*

Proof. We reduce from Pl-Holant($[x, y, y] \mid \mathcal{EQ}$) to Pl-Holant($\{f, g\} \cup \widehat{\mathcal{EQ}}$). This first problem is #P-hard by Theorem 8.1.2 since $y \neq 0$ and $x^2 \neq y^2$. By Lemma 7.4.15, we have $\mathcal{EQ}_2$.

An instance of Pl-Holant($[x, y, y] \mid \mathcal{EQ}$) is a planar bipartite graph $G = (U, V, E)$ in which every vertex in U has degree 2. We replace every vertex in V of degree k (which is assigned $=_k \in \mathcal{EQ}$) with a vertex of degree $2k$ and assign $=_{2k} \in \mathcal{EQ}_2$. Furthermore, we bundle two adjacent variables to form k bundles of 2 edges each. The k bundles correspond to the k incident edges of the original vertex with degree k.

If the inputs to these Equality signatures are restricted to $\{(0, 0), (1, 1)\}$ on each bundle, then these Equality signatures take value 1 on $((0, 0), \ldots, (0, 0))$ and $((1, 1), \ldots, (1, 1))$ and take value 0 elsewhere. Thus, if we restrict the domain to $\{(0, 0), (1, 1)\}$, it is the Equality signature $=_k$.

To simulate $[x, y, y]$, we connect $f = [x, 0, y, 0]$ to $e = [1, 0, 1, 0] \in \widehat{\mathcal{EQ}}$ by a single edge as shown in Figure 8.1 to form a gadget with signature

$$h(a_1, a_2, b_1, b_2) = \sum_{c=0,1} f(a_1, b_1, c) \cdot e(a_2, b_2, c).$$

We replace every (degree 2) vertex in U (which is assigned $[x, y, y]$) by a degree 4 vertex assigned h, where the variables of h are bundled as (a_1, a_2) and (b_1, b_2).

The vertices in this new graph G' are connected as in the original graph G, except that every original edge is replaced by two edges that connect to the same side of the gadget in Figure 8.1. Notice that h is only connected by (a_1, a_2) and (b_1, b_2) to some bundle of two incident edges of an Equality signature. Since this Equality signature enforces that the value on each bundle is either $(0, 0)$ or $(1, 1)$, we only need to consider the restriction of h to the domain $\{(0, 0), (1, 1)\}$. On this domain, $h = [x, y, y]$ is a *symmetric* signature of arity 2. Therefore, the Holant of G' has the same Holant value as the original graph G. □

There are two scenarios that lead to Lemma 8.2.1. The proof of the first is immediate.

Corollary 8.2.2. *Let $a, b, x, y \in \mathbb{C}$. Suppose $f = [x, 0, y]$ and $g = [a, 0, \ldots, 0, b]$ with arity at least 3. If $abxy \neq 0$ and $x^4 \neq y^4$, then* Pl-Holant$(\{f, g\} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard.*

Proof. Connect three copies of $f = [x, 0, y]$ to $[1, 0, 1, 0]$, with one on each edge, to get $x[x^2, 0, y^2, 0]$ and apply Lemma 8.2.1. □

The second scenario that leads to Lemma 8.2.1 is Lemma 8.2.4. The proof of Lemma 8.2.4 applies Corollary 8.2.2 after interpolating a unary signature in one of two ways. The next lemma considers one of those ways.

Lemma 8.2.3. *Let $x \in \mathbb{C}$. Suppose $\mathcal{F}$ is a set of signatures containing $f = [1, x, 1]$. If $x \notin \{0, \pm 1\}$ and M_f has infinite order modulo a scalar, then*

$$\text{Pl-Holant}(\mathcal{F} \cup \{[a, b]\} \cup \widehat{\mathcal{EQ}}) \leq_T \text{Pl-Holant}(\mathcal{F} \cup \widehat{\mathcal{EQ}})$$

for any $a, b \in \mathbb{C}$.

Proof. Consider the unary recursive construction (M_f, s) in Figure 6.1, where $s = \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$. The determinant of M_f is $1 - x^2 \neq 0$. The determinant of $[s\ M_f s]$ is $x \neq 0$. By assumption, M_f has infinite order modulo a scalar. Therefore, we can interpolate any unary signature by Lemma 6.2.4. □

Lemma 8.2.4. *Let $a, b \in \mathbb{C}$. Suppose $f = [a, 0, \ldots, 0, b]$ with arity at least 3. If $ab \neq 0$ and $a^4 \neq b^4$, then* Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard.*

Proof. Since $a \neq 0$, we normalize f to $[1, 0, \ldots, 0, x]$, where $x \neq 0$ and $x^4 \neq 1$. If the arity of f is even, then after some number of self-loops, we have $[1, 0, x]$ and are done by Corollary 8.2.2. Otherwise, the arity of f is odd. After some number of self-loops, we have $g = [1, 0, 0, x]$. If we had the signature $[1, 1]$, then we could connect this to g to get $[1, 0, x]$ and be done by Corollary 8.2.2. We now show how to interpolate $[1, 1]$ in one of two ways. In either case, we use the signature $[1, x]$, which we obtain via a self-loop on g.

Suppose $\Re(x)$, the real part of x, is nonzero. Connecting $[1, x]$ to $[1, 0, 1, 0]$ gives $h = [1, x, 1]$. The eigenvalues of M_h are $\lambda_\pm = 1 \pm x$. Since $\Re(x) \neq 0$ iff $|\frac{\lambda_+}{\lambda_-}| \neq 1$, the ratio of the eigenvalues is not a root of unity, so M_h has infinite order modulo a scalar. Therefore, we can interpolate $[1, 1]$ by Lemma 8.2.3.

Otherwise, $\Re(x) = 0$ but x is not a root of unity since $x \neq \pm i$. Connecting $[1, x]$ to g gives $h = [1, 0, x^2]$. Consider the unary recursive construction (M_h, s) in Figure 6.1, where $s = \left[\begin{smallmatrix}1\\x\end{smallmatrix}\right]$. The determinant of M_h is $x^2 \neq 0$, so its eigenvalues are nonzero. Also, the determinant of $[s\ M_h s]$ is $x(x^2 - 1) \neq 0$. The ratio of the eigenvalues of M_h is x^2, which is not a root of unity since x is not a root of unity. Therefore M_h has infinite order modulo a scalar and we can interpolate $[1, 1]$ by Lemma 6.2.4. □

8.3 Mixing of Tractable Signatures

In this section, we determine which tractable signatures combine to give #P-hardness. To help understand the various cases considered in the lemmas, Figure 8.2 contains a Venn diagram of the signatures in $\mathscr{A}$, $\widehat{\mathscr{P}}$, and $\mathscr{M}$.

The first two lemmas consider the case when one of the signatures is unary.

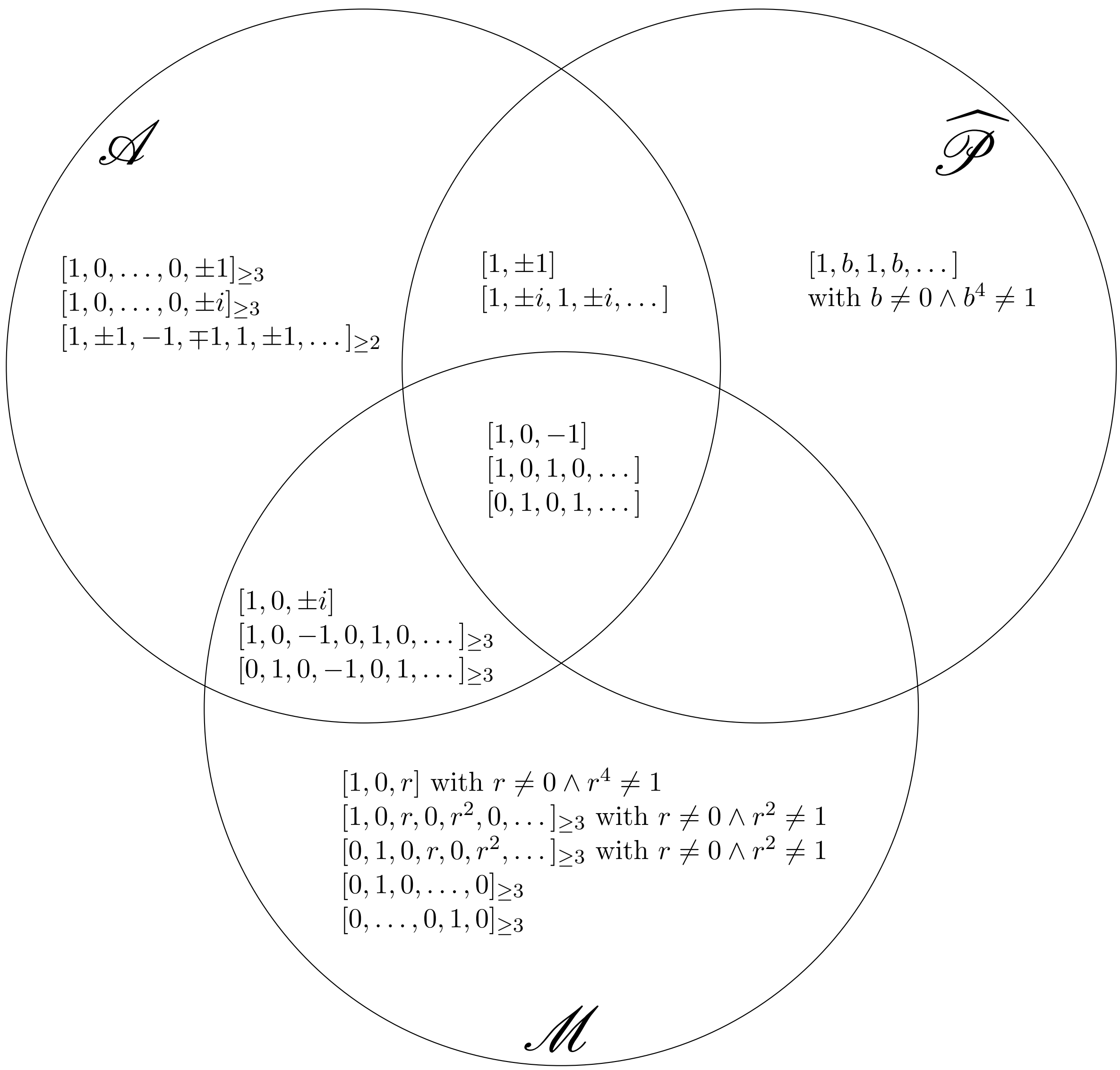


Figure 8.2: Venn diagram of the tractable Pl-#CSP signature sets in the Hadamard basis. Each signature has been normalized for simplicity of presentation. For a signature f, the notation "$f_{\geq k}$" is short for "arity$(f) \geq k$". Notice that $\mathscr{M} \cap \widehat{\mathscr{P}} - \mathscr{A}$ is empty.

Lemma 8.3.1. *Let $a, b \in \mathbb{C}$. Suppose $f \in \mathscr{A} - \widehat{\mathscr{P}}$. If $ab \neq 0$ and $a^4 \neq b^4$, then the problem* $\operatorname{Pl-Holant}(\{[a,b], f\} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard.*

Proof. Up to a nonzero scalar, the possibilities for f are

- $[1, 0, \pm i]$;
- $[1, 0, \ldots, 0, x]$ of arity at least 3 with $x^4 = 1$;
- $[1, \pm 1, -1, \mp 1, 1, \pm 1, -1, \mp 1, \ldots]$ of arity at least 2;
- $[1, 0, -1, 0, 1, 0, -1, 0, \ldots, 0 \text{ or } 1 \text{ or } (-1)]$ of arity at least 3;
- $[0, 1, 0, -1, 0, 1, 0, -1, \ldots, 0 \text{ or } 1 \text{ or } (-1)]$ of arity at least 3.

We handle these cases below.

1. Suppose $f = [1, 0, \pm i]$. Connecting $[a, b]$ to $[1, 0, 1, 0]$ gives $[a, b, a]$, and connecting two copies of $[1, 0, \pm i]$ to $[a, b, a]$, one on each edge, gives $g = [a, \pm ib, -a]$. Then $\operatorname{Pl-Holant}(g \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2.

2. Suppose $f = [1, 0, \ldots, 0, x]$ of arity at least 3 with $x^4 = 1$. Connecting $[a, b]$ to f gives $g = [a, 0, \ldots, 0, bx]$ of arity at least 2. Note that $(bx)^4 = b^4 \neq a^4$. If the arity of g is exactly 2, then $\operatorname{Pl-Holant}(\{f, g\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Corollary 8.2.2, so we are done. Otherwise, the arity of g is at least 3 and $\operatorname{Pl-Holant}(\{g\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.2.4.

3. Suppose $f = [1, \pm 1, -1, \ldots]$ of arity at least 2. Connecting some number of $[1, 0]$ gives $[1, \pm 1, -1]$ of arity exactly 2. Connecting $[a, b]$ to $[1, 0, 1, 0]$ gives $[a, b, a]$ and connecting two copies of $[a, b, a]$ to $[1, \pm 1, -1]$, one on each edge, gives $g = [a^2 \pm 2ab - b^2, \pm(a^2 + b^2), -a^2 \pm 2ab + b^2]$. This is easily verified by

$$\begin{bmatrix} a & b \\ b & a \end{bmatrix} \begin{bmatrix} 1 & \pm 1 \\ \pm 1 & -1 \end{bmatrix} \begin{bmatrix} a & b \\ b & a \end{bmatrix} = \begin{bmatrix} a^2 \pm 2ab - b^2 & \pm(a^2 + b^2) \\ \pm(a^2 + b^2) & -a^2 \pm 2ab + b^2 \end{bmatrix}.$$

   Then $\operatorname{Pl-Holant}(g \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2.

4. Suppose $f = [1, 0, -1, 0, \ldots]$ of arity at least 3. Connecting some number of $[1, 0]$ gives $g = [1, 0, -1, 0]$ of arity exactly 3. Connecting $[a, b]$ to g gives $h = [a, -b, -a]$. Then $\operatorname{Pl-Holant}(h \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2.

5. The argument for $f = [0, 1, 0, -1, \dots]$ is similar to the previous case. $\square$

Lemma 8.3.2. *Let $a, b \in \mathbb{C}$. If $f \in \mathscr{M} - \mathscr{A}$ and $ab \neq 0$, then* Pl-Holant$(\{[a,b], f\} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard.*

Proof. Up to a nonzero scalar, the possibilities for f are as follows:

- $[1, 0, r]$ with $r \neq 0$ and $r^4 \neq 1$;
- $[1, 0, r, 0, r^2, 0, \dots]$ of arity at least 3 with $r \neq 0$ and $r^2 \neq 1$;
- $[0, 1, 0, r, 0, r^2, \dots]$ of arity at least 3 with $r \neq 0$ and $r^2 \neq 1$;
- $[0, 1, 0, \dots, 0]$ of arity at least 3;
- $[0, \dots, 0, 1, 0]$ of arity at least 3.

We handle these cases below.

1. Suppose $f = [1, 0, r]$ with $r^4 \neq 1$ and $r \neq 0$. Connecting $[a, b]$ to $[1, 0, 1, 0]$ gives $[a, b, a]$ and connecting two copies of $[1, 0, r]$ to $[a, b, a]$, one on each edge, gives $g = [a, br, ar^2]$. If $a^2 \neq b^2$, then Pl-Holant$(g \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2.

   Otherwise, $a^2 = b^2$ and we begin by connecting $[a, b]$ to $[1, 0, r]$ to get $[a, br]$. Then by the same construction, we have $g = [a, br^2, ar^2]$ and Pl-Holant$(g \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2.
2. Suppose $f = [1, 0, r, 0, \dots]$ of arity at least 3 with $r^2 \neq 1$ and $r \neq 0$. Connecting some number of $[1, 0]$ gives $g = [1, 0, r, 0]$ of arity exactly 3. Connecting $[a, b]$ to g gives $h = [a, br, a]$. If $a^2 \neq b^2 r$, then Pl-Holant$(h \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2.

   Otherwise, $a^2 = b^2 r$ and we begin by connecting $[1, 0]$ and $[a, b]$ to $[1, 0, r, 0]$ to get $[a, br]$. Then by the same construction, we have $g = [a, br^2, ar]$ and Pl-Holant$(g \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2.
3. The argument for $f = [0, 1, 0, r, \dots]$ is similar to the previous case.
4. Suppose $f = [0, 1, 0, \dots, 0]$ of arity $k \geq 3$. Connecting $k - 2$ copies of $[a, b]$ to f gives $g = a^{k-3}[(k-2)b, a, 0]$. Then Pl-Holant$(g \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2.
5. The argument for $f = [0, \dots, 0, 1, 0]$ is similar to the previous case. $\square$

Now we consider the general case of two signatures from two different tractable sets. Three tractable sets give rise to three pairs of tractable sets to consider, each of which is covered in one of the next three lemmas.

Lemma 8.3.3. *If* $f \in \mathscr{A} - \widehat{\mathscr{P}}$ *and* $g \in \widehat{\mathscr{P}} - \mathscr{A}$, *then* Pl-Holant$(\{f, g\} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard.*

Proof. The only possibility for g is $[a, b, a, b, \dots]$, where $ab \neq 0$ and $a^4 \neq b^4$. Connecting some number of $[1, 0]$ to g gives $[a, b]$ and we are done by Lemma 8.3.1. □

Lemma 8.3.4. *If* $f \in \mathscr{A} - \mathscr{M}$ *and* $g \in \mathscr{M} - \mathscr{A}$, *then* Pl-Holant$(\{f, g\} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard.*

Proof. Suppose f does not contain a 0 entry. Then after connecting some number of $[1, 0]$ to f, we have a unary signature $[a, b]$ with $ab \neq 0$, and are done by Lemma 8.3.2.

Otherwise, f contains a 0 entry. Then $f = [x, 0, \dots, 0, y]$ of arity at least 3 with $xy \neq 0$ (and $x^4 = y^4$). Up to a nonzero scalar, the possibilities for g are as follows:

- $[1, 0, r]$ with $r \neq 0$ and $r^4 \neq 1$;
- $[1, 0, r, 0, r^2, 0, \dots]$ of arity at least 3 with $r \neq 0$ and $r^2 \neq 1$;
- $[0, 1, 0, r, 0, r^2, \dots]$ of arity at least 3 with $r \neq 0$ and $r^2 \neq 1$;
- $[0, 1, 0, 0, \dots, 0]$ of arity at least 3;
- $[0, \dots, 0, 0, 1, 0]$ of arity at least 3.

We handle these cases below.

1. Suppose $g = [1, 0, r]$ with $r \neq 0$ and $r^4 \neq 1$. Then we are done by Corollary 8.2.2.
2. Suppose $g = [1, 0, r, 0, \dots]$ of arity at least 3 with $r \neq 0$ and $r^2 \neq 1$. After connecting some number of $[1, 0]$ to g, we have $h = [1, 0, r, 0]$ of arity exactly 3. Then Pl-Holant$(\{f, h\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.2.1.
3. Suppose $g = [0, 1, 0, r, \dots]$ of arity at least 3 with $r \neq 0$ and $r^2 \neq 1$. After connecting some number of $[1, 0]$ to g, we have $h = [0, 1, 0, r]$ of arity exactly 3. Connecting two more copies of $[1, 0]$ to h gives $[0, 1]$. Then we apply a holographic transformation by $T = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, so f is transformed to $\hat{f} = [y, 0, \dots, 0, x]$ and h is transformed to $\hat{h} = [r, 0, 1, 0]$. Every even arity signature in $\widehat{\mathcal{EQ}}$ remains unchanged after a holographic transformation by T. By attaching $[0, 1]T = [1, 0]$ to every even arity signature in $T\widehat{\mathcal{EQ}}$, we obtain all of the odd arity signatures in $\widehat{\mathcal{EQ}}$ again. Then Pl-Holant$(\{\hat{f}, \hat{h}\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.2.1.

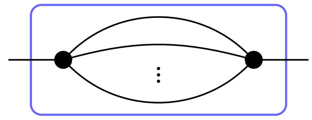


Figure 8.3: The vertices are assigned $g = [0, 1, 0, \ldots, 0]$.

4. Suppose $g = [0, 1, 0, \ldots, 0]$ of arity $k \geq 3$. The gadget in Figure 8.3 with g assigned to both vertices has signature $h = [k-1, 0, 1]$. Then Pl-Holant$(\{f, h\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Corollary 8.2.2.
5. The argument for $g = [0, \ldots, 0, 1, 0]$ is similar to the previous case. □

Lemma 8.3.5. *If $f \in \mathscr{M} - \widehat{\mathscr{P}}$ and $g \in \widehat{\mathscr{P}} - \mathscr{M}$ and $\{f, g\} \not\subseteq \mathscr{A}$,* Pl-Holant$(\{f, g\} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard.*

Proof. The only possibility for g is $[a, b, a, b, \ldots]$ with $ab \neq 0$. Connecting some number of $[1, 0]$ to g gives $h = [a, b]$. If $f \notin \mathscr{A}$, then Pl-Holant$(\{f, h\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.3.2. Otherwise, $f \in \mathscr{A}$, so $g \notin \mathscr{A}$ and Pl-Holant$(\{f, g\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.3.3. □

We summarize this section with the following theorem, which says that the tractable signature sets cannot mix. Signatures from different tractable sets, when put together, lead to #P-hardness.

Theorem 8.3.6 (No Mixing)**.** *Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures in Boolean variables. If $\mathcal{F} \subseteq \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$, then* Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard unless $\mathcal{F} \subseteq \mathscr{A}$, $\mathcal{F} \subseteq \widehat{\mathscr{P}}$, or $\mathcal{F} \subseteq \mathscr{M}$, in which case* Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is tractable.*

Proof. If $\mathcal{F}$ is a subset of $\mathscr{A}$, $\widehat{\mathscr{P}}$, or $\mathscr{M}$, then the tractability is given in Theorem 8.1.4. Otherwise $\mathcal{F}$ is not a subset of $\mathscr{A}$, $\widehat{\mathscr{P}}$, or $\mathscr{M}$. Then $\mathcal{F}$ contains a signature $g \in (\widehat{\mathscr{P}} \cup \mathscr{M}) - \mathscr{A}$ since $\mathcal{F} \not\subseteq \mathscr{A}$. Suppose $\mathcal{F}$ contains a signature $f \in \mathscr{A} - \widehat{\mathscr{P}} - \mathscr{M}$. If $g \in \widehat{\mathscr{P}} - \mathscr{A}$, then Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.3.3. Otherwise, $g \in \mathscr{M} - \mathscr{A}$ and Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.3.4.

Now assume that $\mathcal{F} \subseteq \widehat{\mathscr{P}} \cup \mathscr{M}$. Since $(\widehat{\mathscr{P}} \cap \mathscr{M}) - \mathscr{A}$ is empty (see Figure 8.2), either $g \in \widehat{\mathscr{P}} - \mathscr{M} - \mathscr{A}$ or $g \in \mathscr{M} - \widehat{\mathscr{P}} - \mathscr{A}$. If $g \in \widehat{\mathscr{P}} - \mathscr{M} - \mathscr{A}$, then there exists a signature $f \in \mathscr{M} - \widehat{\mathscr{P}}$ since $\mathcal{F} \not\subseteq \widehat{\mathscr{P}}$. In which case, Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.3.5. Otherwise, $g \in \mathscr{M} - \widehat{\mathscr{P}} - \mathscr{A}$ and there exists a signature $f \in \widehat{\mathscr{P}} - \mathscr{M}$ since $\mathcal{F} \not\subseteq \mathscr{M}$. In which case, Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.3.5. □

8.4 Pinning for Planar Graphs

The idea of "pinning" is a common reduction technique between counting problems. For the #CSP framework, pinning fixes some variables to specific values of the domain by means of the constant functions [17, 58, 13, 80]. In particular, for counting graph homomorphisms, pinning is used when the input graph is connected and the target graph is disconnected. In this case, pinning a vertex of the input graph to a vertex of the target graph forces all the vertices of the input graph to map to the same connected component of the target graph [59, 14, 70, 122, 23]. For the Boolean domain, the constant 0 and constant 1 functions are the signatures $[1, 0]$ and $[0, 1]$ respectively.

From these works, the most relevant pinning lemma for the Pl-#CSP framework is by Dyer, Goldberg, and Jerrum in [58], where they show how to pin in the #CSP framework. However, the proof of this pinning lemma is highly nonplanar. Cai, Lu, and Xia [46] overcame this difficultly in the proof of their dichotomy theorem for the real-weighted Pl-#CSP framework by first undergoing a holographic transformation by the Hadamard matrix $H = \left[\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right]$ and then pinning in this Hadamard basis.[3] We stress that this holographic transformation is necessary. Indeed, if one were able to pin in the standard basis of the Pl-#CSP framework, then P = #P would follow since Pl-#CSP($\widehat{\mathscr{M}}$) is tractable but Pl-#CSP($\widehat{\mathscr{M}} \cup \{[1, 0], [0, 1]\}$) is #P-hard by our main dichotomy in Theorem 8.5.3 (or, more specifically, by Lemma 8.3.2).

Since Pl-#CSP($\mathcal{F}$) is equivalent to Pl-Holant($\mathcal{F} \cup \mathcal{EQ}$), the expression of Pl-#CSP($\mathcal{F}$) in the Hadamard basis is Pl-Holant($H\mathcal{F} \cup \widehat{\mathcal{EQ}}$). Then we already have $[1, 0] \in \widehat{\mathcal{EQ}}$, so pinning in the Hadamard basis of Pl-#CSP($\mathcal{F}$) amounts to obtaining the missing signature $[0, 1]$.

8.4.1 The Road to Pinning

We begin the road to pinning with a lemma that assumes the presence of $[0, 0, 1] = [0, 1]^{\otimes 2}$, which is the tensor product of two copies of $[0, 1]$. In our pursuit to realize $[0, 1]$, this may be as close as we can get, such as when every signature has even arity. Another roadblock to realizing $[0, 1]$ is when every signature has even parity. Recall that a signature has even parity if its support is on entries

[3] The pinning in [46], which is accomplished in Section IV, is not summarized in a single statement but is implied by the combination of all the results in that section.

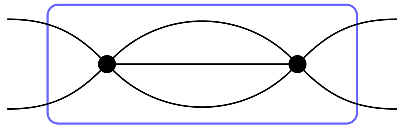


Figure 8.4: The circles are assigned $[a, 0, 0, 0, b, c]$.

of even Hamming weight. By a simple parity argument, gadgets constructed with signatures of even parity can only realize signatures of even parity. However, if every signature has even parity and $[0, 0, 1]$ is present, then we can already prove a dichotomy. To prove this, we use the following lemma.

Lemma 8.4.1. *Let $a, b, c \in \mathbb{C}$. If $ab \neq 0$, then* Pl-Holant$([a, 0, 0, 0, b, c])$ *is* #P*-hard.*

Proof. Let f be the signature of the gadget in Figure 8.4 with $[a, 0, 0, 0, b, c]$ assigned to both vertices. The signature matrix of f is

$$\begin{bmatrix} a^2 & 0 & 0 & 0 \\ 0 & b^2 & b^2 & bc \\ 0 & b^2 & b^2 & bc \\ 0 & bc & bc & 3b^2 + c^2 \end{bmatrix},$$

which is redundant. Its compressed form is nonsingular since its determinant is $6a^2b^4 \neq 0$. Thus, we are done by Lemma 6.5.4. □

Lemma 8.4.2. *Suppose $\mathcal{F}$ is a set of symmetric signatures with complex weights containing* $[0, 0, 1]$*. If every signature in $\mathcal{F}$ has even parity, then either* Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard or $\mathcal{F}$ is a subset of $\mathscr{A}$, $\widehat{\mathscr{P}}$, or $\mathscr{M}$, in which case* Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is tractable.*

Proof. The tractability is given in Theorem 8.1.4. If every non-degenerate signature in $\mathcal{F}$ is of arity at most 3, then $\mathcal{F} \subseteq \mathscr{M}$ since all signatures in $\mathcal{F}$ satisfy the (even) parity condition.

Otherwise $\mathcal{F}$ contains some non-degenerate signature of arity at least 4. For every signature $f \in \mathcal{F}$ with $f = [f_0, f_1, \dots, f_m]$ and $m \geq 4$, using $[0, 0, 1]$ and $[1, 0]$, we can obtain all subsignatures of the form $[f_{k-2}, 0, f_k, 0, f_{k+2}]$ for any even k such that $2 \leq k \leq m - 2$. If any subsignature g of this form satisfies $f_{k-2}f_{k+2} \neq f_k^2$ and $f_k \neq 0$, then we are done by Lemma 6.5.4.

Otherwise all subsignatures of signatures in $\mathcal{F}$ of the above form satisfy $f_{k-2}f_{k+2} = f_k^2$ or $f_k = 0$. There are two types of signatures with this property. In the first type, the signature entries of even Hamming weight form a geometric progression. More specifically, the signatures of the first type have the form

$$[\alpha^n, 0, \alpha^{n-1}\beta, 0, \dots, 0, \alpha\beta^{n-1}, 0, \beta^n] \qquad \text{or} \qquad [\alpha^n, 0, \alpha^{n-1}\beta, 0, \dots, 0, \alpha\beta^{n-1}, 0, \beta^n, 0]$$

for some $\alpha, \beta \in \mathbb{C}$, which are in $\mathscr{M}$. In the second type, the signatures have arity at least 4 or 5 and are of the form $[x, 0, \dots, 0, y]$ or $[x, 0, \dots, 0, y, 0]$ respectively, with $xy \neq 0$ and an odd number of 0's between x and y (since they have even parity). If all signatures in $\mathcal{F}$ are of the first type, then $\mathcal{F} \subseteq \mathscr{M}$.

Otherwise $\mathcal{F}$ contains a signature f of the second type. Suppose $f = [x, 0, \dots, 0, y, 0]$ of arity at least 5 with $xy \neq 0$. After some number of self-loops, we have $g = [x, 0, 0, 0, y, 0]$ of arity exactly 5. Then we are done by Lemma 8.4.1.

Otherwise $f = [x, 0, \dots, 0, y]$ of arity at least 4 with $xy \neq 0$. If $x^4 \neq y^4$, then we are done by Lemma 8.2.4.

Otherwise $x^4 = y^4$. This puts every signature of the second type in $\mathscr{A}$. Therefore $\mathcal{F} \subseteq \mathscr{A} \cup \mathscr{M}$ and we are done by Theorem 8.3.6. □

The conclusion of every result in the rest of this section states that we are able to pin (under various assumptions on $\mathcal{F}$). Formally speaking, we repeatedly prove that Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard (or in P) if and only if Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard (or in P). The difference between these two counting problems is the presence of $[0, 1]$ in Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$. We always prove this statement in one of three ways:

1. either we show that Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is tractable (so Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is as well);
2. or we show that Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard (so Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is as well);
3. or we show how to reduce Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ to Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ by realizing $[0, 1]$ using signatures in $\mathcal{F} \cup \widehat{\mathcal{EQ}}$.

Lemma 8.4.3. *Let $\mathcal{F}$ be any set of complex-weighted symmetric signatures containing* $[0, 0, 1]$.

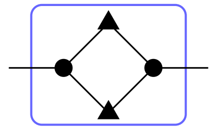


Figure 8.5: The circles are assigned $[1, 0, 1, 0]$ and the triangles are assigned $[1, 0, x]$.

Then Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard (or in P) iff* Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard (or in P).*

Proof. If we had a unary signature $[a, b]$ where $b \neq 0$, then connecting $[a, b]$ to $[0, 0, 1]$ gives the signature $[0, b]$, which is $[0, 1]$ after normalizing. Thus, in order to reduce Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ to Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ by constructing $[0, 1]$, it suffices to construct a unary signature $[a, b]$ with $b \neq 0$.

For every signature $f \in \mathcal{F}$ with $f = [f_0, f_1, \ldots, f_m]$, using $[0, 0, 1]$ and $[1, 0]$, we can obtain all subsignatures of the form $[f_{k-1}, f_k]$ for any odd k such that $1 \leq k \leq m$. If any subsignature satisfies $f_k \neq 0$, then we can construct $[0, 1]$.

Otherwise all signatures in $\mathcal{F}$ have even parity and we are done by Lemma 8.4.2. □

There are two scenarios that lead to Lemma 8.4.3, which are the focus of the next two lemmas.

Lemma 8.4.4. *For $x \in \mathbb{C}$, let $\mathcal{F}$ be any set of complex-weighted symmetric signatures containing $[1, 0, x]$ such that $x \notin \{0, \pm 1\}$. Then* Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard (or in P) iff* Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard (or in P).*

Proof. There are two cases. In either case, we realize $[0, 0, 1]$ and finish by applying Lemma 8.4.3.

First we claim that the conclusion holds provided $|x| \notin \{0, 1\}$. Combining k copies of $[1, 0, x]$ gives $[1, 0, x^k]$. Since $|x| \notin \{0, 1\}$, x is neither zero nor a root of unity, so we can use polynomial interpolation to realize $[a, 0, b]$ for any $a, b \in \mathbb{C}$, including $[0, 0, 1]$.

Otherwise $|x| = 1$. The gadget in Figure 8.5 has signature $[f_0, f_1, f_2] = [1 + x^2, 0, 2x]$. If $x = \pm i$, then we have $[0, 0, \pm 2i]$, which is $[0, 0, 1]$ after normalizing.

Otherwise $x \neq \pm i$, so $f_0 \neq 0$. Since $x \neq 0$, we have $f_2 \neq 0$. Since $x \neq \pm 1$, we have $|f_0| < 2$. However, $|f_2| = 2$. Therefore, after normalizing, the signature $[1, 0, y]$ with $y = \frac{2x}{1+x^2}$ has $|y| > 1$, so it can interpolate $[0, 0, 1]$ by our initial claim since $|y| \notin \{0, 1\}$. □

Lemma 8.4.5. *Let $\mathcal{F}$ be any set of complex-weighted symmetric signatures containing a signature $[f_0, f_1, \ldots, f_n]$ that is not identically zero but has $f_0 = 0$. Then $\text{Pl-Holant}^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard (or in P) iff $\text{Pl-Holant}(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard (or in P).*

Proof. If $f_1 \neq 0$, then we connect $n-1$ copies of $[1, 0]$ to f to get $[0, f_1]$, which is $[0, 1]$ after normalizing. If $f_1 = 0$, then $n \geq 2$. If $f_2 \neq 0$, then we connect $n-2$ copies of $[1, 0]$ to f to get $[0, 0, f_2]$, which is $[0, 0, 1]$ after normalizing. Then we are done by Lemma 8.4.3. If $f_1 = f_2 = 0$, then $n \geq 3$. After some number of self-loops, we get a signature with exactly one or two initial 0's, which is one of the above scenarios. □

As a significant step toward pinning for any signature set $\mathcal{F}$, we show how to pin given any binary signature. Some cases resist pinning and are excluded.

Lemma 8.4.6. *Let $\mathcal{F}$ be any set of complex-weighted symmetric signatures containing a binary signature f. Then $\text{Pl-Holant}^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard (or in P) iff $\text{Pl-Holant}(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard (or in P) unless $f \in \{[0, 0, 0], [1, 0, -1], [1, r, r^2], [1, b, 1]\}$, up to a nonzero scalar, for any $b, r \in \mathbb{C}$.*

Proof. Let $f = [f_0, f_1, f_2]$. If $f_0 = 0$ and either $f_1 \neq 0$ or $f_2 \neq 0$, then we are done by Lemma 8.4.5. Otherwise, $f = [0, 0, 0]$ or $f_0 \neq 0$, in which case we normalize f_0 to 1. If $\text{Pl-Holant}(f \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2, then $\text{Pl-Holant}(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is also #P-hard. Otherwise, f is one of the tractable cases, which implies that

$$f \in \{[0, 0, 0], [1, r, r^2], [1, 0, x], [1, \pm 1, -1], [1, b, 1]\}.$$

If $f = [1, \pm 1, -1]$, then we connect f to $[1, 0, 1, 0]$ to get $[0, \pm 2]$, which is $[0, 1]$ after normalizing. If $f = [1, 0, x]$, then we are done by Lemma 8.4.4 unless $x \in \{0, \pm 1\}$. The remaining cases are all excluded by assumption, so we are done. □

8.4.2 Pinning in the Hadamard Basis

Before we show how to pin in the Hadamard basis, we handle two simple cases.

Lemma 8.4.7. *If $\mathcal{F}$ is a set of signatures containing $[1, \pm i]$, then we have* $\text{Pl-Holant}^c(\mathcal{F} \cup \widehat{\mathcal{EQ}}) \leq_T \text{Pl-Holant}(\mathcal{F} \cup \widehat{\mathcal{EQ}})$.

Proof. Connect two copies of $[1, \pm i]$ to $[1, 0, 1, 0]$ to get $[0, \pm 2i]$, which is $[0, 1]$ after normalizing. □

The next lemma considers the signature $[1, b, 1, b^{-1}]$, which we also encounter in Theorem 8.5.1, the single signature dichotomy.

Lemma 8.4.8. *Let $b \in \mathbb{C}$. If $b \notin \{0, \pm 1\}$, then* $\text{Pl-Holant}(\{[1, b, 1, b^{-1}]\} \cup \widehat{\mathcal{EQ}})$ *is #P-hard.*

Proof. Connect two copies of $[1, 0]$ to $f = [1, b, 1, b^{-1}]$ to get $[1, b]$. Connecting this back to f gives $g = [1 + b^2, 2b, 2]$. Then $\text{Pl-Holant}(g \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2. □

Now we are ready to prove our pinning result.

Theorem 8.4.9 (Pinning)**.** *Let $\mathcal{F}$ be any set of complex-weighted symmetric signatures. Then* $\text{Pl-Holant}^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is #P-hard (or in P) iff* $\text{Pl-Holant}(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is #P-hard (or in P).*

This theorem does not exclude the possibility that either framework can express a problem of intermediate complexity. It merely says that if one framework cannot express a problem of intermediate complexity, then neither can the other. Our goal is to prove a dichotomy for $\text{Pl-Holant}(\mathcal{F} \cup \widehat{\mathcal{EQ}})$. By Theorem 8.4.9, this is equivalent to proving a dichotomy for $\text{Pl-Holant}^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$.

Proof of Theorem 8.4.9. For simplicity, we normalize the first nonzero entry of every signature in $\mathcal{F}$ to 1. If $\mathcal{F}$ contains the degenerate signature $[0, 1]^{\otimes n}$ for some $n \geq 1$, then we take self-loops on this signature until we have either $[0, 1]$ or $[0, 0, 1]$ (depending on the parity of n). If we have $[0, 1]$, we are done. Otherwise, we have $[0, 0, 1]$ and are done by Lemma 8.4.3.

Now assume that any degenerate signature in $\mathcal{F}$ is not of the form $[0, 1]^{\otimes n}$. Then we can replace these degenerate signatures in $\mathcal{F}$ by their unary versions using $[1, 0]$. This does not change the complexity of the problem. If $\mathcal{F}$ contains only unary signatures, then $\mathcal{F} \subseteq \widehat{\mathscr{P}}$ and $\text{Pl-Holant}^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is tractable by Theorem 8.1.4.

Otherwise $\mathcal{F}$ contains a signature f of arity at least two. We connect some number of $[1, 0]$ to f until we obtain a signature with arity exactly two. We call the resulting signature the binary prefix of f. If this binary prefix is not one of the exceptional forms in Lemma 8.4.6, then we are done, so assume that it is one of the exceptional forms.

Now we perform case analysis according to the exceptional forms in Lemma 8.4.6. There are five cases below because we consider $[1, r, r^2]$ as $[1, 0, 0]$ and $[1, r, r^2]$ with $r \neq 0$ as separate cases. In each case, we either show that the conclusion of the theorem holds or that $f \in \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$. After the case analysis, we then handle all of these tractable f together.

1. Suppose the binary prefix of f is $[0, 0, 0]$. If f is not identically zero, then we are done by Lemma 8.4.5.

   Thus, in this case, we may assume $f = [0, 0, \ldots, 0]$ is identically zero.

2. Suppose the binary prefix of f is $[1, 0, -1]$. If f is not of the form

   $$[1, 0, -1, 0, 1, 0, -1, 0, \ldots, 0 \text{ or } 1 \text{ or } (-1)], \tag{8.4.1}$$

   then after one self-loop, we have a signature of arity at least one with 0 as its first entry but is not identically zero, so we are done by Lemma 8.4.5.

   Thus, in this case, we may assume f has the form given in (8.4.1).

3. Suppose the binary prefix of f is $[1, 0, 0]$. If f is not of the form $[1, 0, \ldots, 0]$, then after connecting some number of $[1, 0]$, we have $[1, 0, \ldots, 0, x]$ of arity at least 3, where $x \neq 0$. If $x^4 \neq 1$, then Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.2.4, so Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

   Otherwise, $x^4 = 1$. Suppose that x is not the last entry in f. Then connecting one fewer $[1, 0]$ than before, we have $g = [1, 0, \ldots, 0, x, y]$ and there are two cases to consider. If the index of x in g is odd, then after some number of self-loops, we have $h = [1, 0, 0, x, y]$. The determinant of the compressed signature matrix of h is $-2x^2 \neq 0$. Thus, Holant(h) is #P-hard by Lemma 6.5.4, so Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

   Otherwise, the index of x in g is even. After some number of self-loops, we have $h = [1, 0, 0, 0, x, y]$. Then by Lemma 8.4.1, Holant(h) is #P-hard, so Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

Thus, in this case, we may assume either $f = [1, 0, \dots, 0]$ or $f = [1, 0, \dots, 0, x]$ with $x^4 = 1$.

4. Suppose the binary prefix of f is $[1, r, r^2]$, where $r \neq 0$. If f is not of the form $[1, r, \dots, r^n]$, then after connecting some number of $[1, 0]$, we have $[1, r, \dots, r^m, y]$, where $y \neq r^{m+1}$ and $m \geq 2$. Using $[1, 0]$, we can get $[1, r]$. If $r = \pm i$, then we are done by Lemma 8.4.7, so assume that $r \neq \pm i$. Then we can attach $[1, r]$ back to the initial signature some number of times to get $g = [1, r, r^2, x]$ after normalizing, where $x \neq r^3$. We connect $[1, r]$ once more to get $h = [1+r^2, r(1+r^2), r^2+rx]$. If h does not have one of the exceptional forms in Lemma 8.4.6, then we are done, so assume that it does.

   Since the second entry of h is not 0 and $x \neq r^3$, the only possibility is that h has the form $[1, b, 1]$ up to a scalar. This gives $x = r^{-1}$. Note that $r \neq \pm 1$ since $x \neq r^3$. A self-loop on $g = [1, r, r^2, r^{-1}]$ gives $[1 + r^2, r + r^{-1}]$, which is $[1, r^{-1}]$ after normalizing. Connecting this back to g gives $h = [2, 2r, r^2 + r^{-2}]$. We assume that h has one of the exceptional forms in Lemma 8.4.6 since we are done otherwise. If h has the form $[1, r, r^2]$ up to a scalar, then $r^4 = 1$, a contradiction, so it must have the form $[1, b, 1]$ up to a scalar. But then $r^2 = 1$, which is also a contradiction.

   Thus, in this case, we may assume $f = [1, r, \dots, r^n]$.

5. Suppose the binary prefix of f is $[1, b, 1]$. If $b = \pm 1$, then this binary prefix is degenerate and was considered in the previous case, so assume that $b \neq \pm 1$. If f is not of the form $[1, b, 1, b, \dots]$, then suppose that the index of the first entry in f to break the pattern is even. Then after connecting some number of $[1, 0]$, we have $[1, b, 1, \dots, b, y]$, where $y \neq 1$. Then after some number of self-loops and normalizing, we have $g = [1, b, 1, b, x]$, where $x \neq 1$. The determinant of its compressed signature matrix is $(b^2 - 1)(1 - x) \neq 0$. Thus, Holant(g) is #P-hard by Lemma 6.5.4, so Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

   Otherwise, the index of the first entry in f to break the pattern is odd. Then after connecting some number of $[1, 0]$, we have $[1, b, 1, \dots, 1, y]$, where $y \neq b$. Then after some number of self-loops and normalizing, we have $[1, b, 1, x]$, where $x \neq b$. We do a self-loop to get $g = [2, b+x]$. If $b = 0$, then connecting g to $[1, 0, 1, x]$ gives $h = [2, x, 2 + x^2]$. We assume that h has one of the exceptional forms in Lemma 8.4.6 since we are done otherwise. Because $x \neq 0$, the only possibility is that h has the form $[1, r, r^2]$ up to a scalar. Then we get $x^2 = -4$, so $g = [2, x] = 2[1, \pm i]$ and we are done by Lemma 8.4.7. We use the signature g again below.

Otherwise, $b \neq 0$. Using $[1, 0]$, we can get $h = [1, b, 1]$. If the signature matrix M_h of h has finite order modulo a scalar, then $M_h^\ell = \beta I_2$ for some positive integer ℓ and some nonzero complex value β. Thus after normalizing, we can construct the anti-gadget $[1, -b, 1]$ by connecting $\ell - 1$ copies of h together. Connecting $[1, 0]$ to $[1, -b, 1]$ gives $[1, -b]$ and connecting this to $[1, b, 1, x]$ gives $[1 - b^2, 0, 1 - bx]$. If $\frac{1-bx}{1-b^2} \notin \{0, \pm 1\}$, then we are done by Lemma 8.4.4.

Otherwise, $y = \frac{1-bx}{1-b^2} \in \{0, \pm 1\}$. For $y = 0$, we get $x = b^{-1}$ and are done by Lemma 8.4.8 since $b \notin \{0, \pm 1\}$. For $y = 1$, we get $b = x$, a contradiction. For $y = -1$, we get $2 - b^2 - bx = 0$. Then connecting $[1, -b, 1]$ to $g = [2, b + x]$ gives $[2 - b^2 - bx, x - b] = [0, x - b]$, which is $[0, 1]$ after normalizing.

Otherwise, M_h has infinite order modulo a scalar, so we can interpolate $[0, 1]$ by Lemma 8.2.3 since $b \notin \{0, \pm 1\}$.

Thus, in this case, we may assume $f = [1, b, 1, b, \dots]$.

At this point, every signature in $\mathcal{F}$ (including the unary signatures) must be of one of the following forms:

- $[0, \dots, 0]$, which is in $\mathscr{A} \cap \widehat{\mathscr{P}} \cap \mathscr{M}$;
- $[1, 0, -1, 0, 1, 0, -1, 0, \dots, 0$ or 1 or $(-1)]$, which is in $\mathscr{A} \cap \mathscr{M}$;
- $[1, 0, \dots, 0, x]$, where $x^4 = 1$, which is in $\mathscr{A}$;
- $[1, b, 1, b, \dots, 1$ or $b]$, which is in $\widehat{\mathscr{P}}$.

In particular, every possible unary signature either fits into the first case or the last case. Therefore $\mathcal{F} \subseteq \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$ and we are done by Theorem 8.3.6. □

8.5 Main Result

In this section, we prove our main dichotomy theorem. But first a dichotomy for a single signature.

Theorem 8.5.1. *If f is a non-degenerate symmetric signature of arity at least 2 with complex* *weights in Boolean variables, then* Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard unless* $f \in \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$*, in which case the problem is computable in polynomial time.*

Proof. When $f \in \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$, the problem is tractable by Theorem 8.1.4. When $f \notin \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$, we prove that Pl-Holant$^c(\{f\} \cup \widehat{\mathcal{EQ}})$ is #P-hard, which is sufficient because of pinning (Theorem 8.4.9). Using $[1, 0]$ and $[0, 1]$, we can obtain any subsignature of f.

Notice that once we have $[0, 1]$ and $\widehat{\mathcal{EQ}}$, we can realize every signature in $T\widehat{\mathcal{EQ}}$, where $T = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. In fact, every even arity signature in $\widehat{\mathcal{EQ}}$ is also in $T\widehat{\mathcal{EQ}}$, and we obtain all the odd arity signatures in $T\widehat{\mathcal{EQ}}$ by attaching $[0, 1]$ to all the even arity signatures in $\widehat{\mathcal{EQ}}$. Therefore, a holographic transformation by T does not change the complexity of the problem. Furthermore, $\mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$ is closed under T. We use these facts later.

The possibilities for f can be divided into three cases:

- f satisfies the parity condition;
- f does not satisfy the parity condition but does contain a 0 entry;
- f does not contain a 0 entry.

We handle these cases below.

1. Suppose that f satisfies the parity condition. If f has even parity, then we are done by Lemma 8.4.2.

   Otherwise, f has odd parity. If f has odd arity, then under a holographic transformation by $T = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, f is transformed to $\hat{f}$, which has even parity. Then either Pl-Holant$^c(\{\hat{f}\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.4.2 (and thus Pl-Holant$^c(\{f\} \cup \widehat{\mathcal{EQ}})$ is also #P-hard), or $\hat{f} \in \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$ (and thus $f \in \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$).

   Otherwise, the arity of f is even. Connect $[0, 1]$ to f to get a signature g with even parity and odd arity. Then either Pl-Holant$^c(\{g\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.4.2 (and thus Pl-Holant$^c(\{f\} \cup \widehat{\mathcal{EQ}})$ is also #P-hard), or $g \in \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$. In the latter case, it must be that $g \in \mathscr{M}$ since non-degenerate generalized equality signatures cannot have both even parity and odd arity. (See Figure 8.2 at the end of the Appendix, which contains a Venn diagram of the signatures in $\mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$, up to constant factors.) In particular, the even parity entries of g form a geometric progression. Therefore $f \in \mathscr{M}$ since f has odd parity and the same geometric progression among its odd parity entries.

2. Suppose that f contains a 0 entry but does not satisfy the parity condition. Since f does not satisfy the parity condition, there must be at least two nonzero entries separated by an even number of 0 entries. Thus, f contains a subsignature $g = [a, 0, \ldots, 0, b]$ of arity $n = 2k+1 \geq 1$, where $ab \neq 0$.

   If $k = 0$, then $n = 1$ and we can shift either to the right or to the left and find the 0 entry in f and obtain a binary subsignature h of the form $[c, d, 0]$ or $[0, c, d]$, where $cd \neq 0$. Then Pl-Holant$(h \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2, so Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

   Otherwise $k \geq 1$, so $n \geq 3$. If $a^4 \neq b^4$, then Pl-Holant$(\{g\} \cup \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.2.4, so Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

   Otherwise, $a^4 = b^4$, so $g \in \mathscr{A}$. If $f = g$, then we are done, so assume that $f \neq g$, which implies that there is another entry just before a or just after b. If this entry is nonzero, then f has a subsignature h of the form $[c, a, 0]$ or $[0, b, d]$, where $cd \neq 0$. Then Pl-Holant$(h \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2, so Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

   Otherwise, this entry is 0 and f has a subsignature h of the form

   $$[0, a, 0, \ldots, 0, b] \qquad \text{or} \qquad [a, 0, \ldots, 0, b, 0]$$

   of arity at least 4. If the arity of h is even, then after some number of self-loops, we have a signature h' of the form $[0, a, 0, 0, b]$ or $[a, 0, 0, b, 0]$ of arity exactly 4. Then Pl-Holant(h') is #P-hard by Lemma 6.5.4 since $ab \neq 0$, so Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

   Otherwise, the arity of h is odd. After some number of self-loops, we have a signature h' of the form $[0, a, 0, 0, 0, b]$ or $[a, 0, 0, 0, b, 0]$ of arity exactly 5. Then Pl-Holant(h') is #P-hard by Lemma 8.4.1 since $ab \neq 0$, so Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

3. Suppose f contains no 0 entry. If f has a binary subsignature g such that Pl-Holant$(g \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2, then Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

   Otherwise every binary subsignature $[a, b, c]$ of f satisfies the conditions of some tractable case in Theorem 8.1.2. The three possible tractable cases are degenerate with condition $ac = b^2$ (case 1), affine $\mathscr{A}$ with condition $ac = -b^2 \wedge a = -c$ (case 3), and a Hadamard-transformed product type $\widehat{\mathscr{P}}$ with condition $a = c$ (case 4). If every binary subsignature $[a, b, c]$ of f satisfies $ac = b^2$, then f is degenerate, a contradiction. If every binary subsignature $[a, b, c]$ of f satisfies $ac = -b^2 \wedge a = -c$, then $f = [1, \pm 1, -1, \mp 1, 1, \pm 1, -1, \mp 1, \dots] \in \mathscr{A}$ (up to a scalar) and we are done. If every binary subsignature $[a, b, c]$ of f satisfies $a = c$, then $f \in \widehat{\mathscr{P}}$ and we are done.

Otherwise, there exists two binary subsignatures of f that do not satisfy the same tractable case in Theorem 8.1.2. More specifically, f has arity at least 3 and there exists a ternary subsignature $g = [a, b, c, d]$ such that $h = [a, b, c]$ and $h' = [b, c, d]$ exclusively satisfy the conditions of different tractable cases in Theorem 8.1.2. By symmetry in the statement of the tractable conditions in Theorem 8.1.2, under a holographic transformation by $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$, we can we can switch the order of h and h'. Suppose f contains a binary subsignature that satisfies the condition of the affine case. Let h be that subsignature. Then for either case of h', we have $g = [1, \varepsilon, -1, \varepsilon]$ after normalizing, where $\varepsilon^2 = 1$. Connecting two copies of $[0, 1]$ to g gives $[-1, \varepsilon]$. Connecting this back to g gives $g' = [0, -2\varepsilon, 2]$. Then Pl-Holant$(g' \mid \widehat{\mathcal{EQ}})$ is #P-hard by Theorem 8.1.2, so Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ is also #P-hard.

Otherwise, we may assume that h satisfies the product-type condition (but not the degenerate condition) and h' satisfies the degenerate condition. Then $g = [1, b, 1, b^{-1}]$ after normalizing, where $b^2 \neq 1$. Then Pl-Holant$(g \mid \widehat{\mathcal{EQ}})$ is #P-hard by Lemma 8.4.8, so Pl-Holant$(\{f\} \cup \widehat{\mathcal{EQ}})$ is also #P-hard. □

Now we are ready to prove our main dichotomy theorem.

Theorem 8.5.2. *Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures in Boolean variables. Then* Pl-Holant$(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ *is* #P*-hard unless $\mathcal{F} \subseteq \mathscr{A}$, $\mathcal{F} \subseteq \widehat{\mathscr{P}}$, or $\mathcal{F} \subseteq \mathscr{M}$, in which case the problem is computable in polynomial time.*

Proof. The tractability is given in Theorem 8.1.4. When $\mathcal{F}$ is not a subset of $\mathscr{A}$, $\widehat{\mathscr{P}}$, or $\mathscr{M}$, we prove that Pl-Holant$^c(\mathcal{F} \cup \widehat{\mathcal{EQ}})$ is #P-hard, which is sufficient because of pinning (Theorem 8.4.9).

For any degenerate signature $f \in \mathcal{F}$, we connect some number of $[1, 0]$ to f to get its corresponding unary signature. We replace f by this unary signature, which does not change the complexity.

Thus, assume that the only degenerate signatures in $\mathcal{F}$ are unary signatures.

If $\mathcal{F} \not\subseteq \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$, then the problem is #P-hard by Theorem 8.5.1. Otherwise, $\mathcal{F} \subseteq \mathscr{A} \cup \widehat{\mathscr{P}} \cup \mathscr{M}$ and we are done by Theorem 8.3.6. □

We also have the corresponding theorem for the Pl-#CSP framework in the standard basis.

Theorem 8.5.3. *Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures in Boolean variables. Then* Pl-#CSP($\mathcal{F}$) *is* #P*-hard unless $\mathcal{F} \subseteq \mathscr{A}$, $\mathcal{F} \subseteq \mathscr{P}$, or $\mathcal{F} \subseteq \widehat{\mathscr{M}}$, in which case the problem is computable in polynomial time.*

8.6 Closing Thoughts

Designing the proof One of the anonymous reviewers of [74] made the following comment.

> "It is hard to imagine how hard it is to design the case-by-case proof in [Sections 8.2, 8.3, and 8.4]."

Let me try to explain how this is done. The short answer is that you don't design the case-by-case proof. You design the overall structure of the proof and the case-by-case aspects of the proof design themselves.

One should begin by considering important ideas from previous work that were successful. We had just finished [29], the Holant dichotomy that was proven in Chapter 7. The hardness proof of the dichotomy for any set of signatures (see the end of Section 7.5) was easy to prove given the tools to which we availed ourselves by that point. The primary tool was a dichotomy for a single signature (Theorem 7.5.1). Thus, the plan is to prove a single signature dichotomy for Pl-#CSP, which we eventually did in Theorem 8.5.1.

Nearly all of the previous dichotomy theorems for #CSP (including nearly all of the dichotomy theorems for the special case of counting graph homomorphisms), used the pinning technique. This makes it an easy decision to attempt to pin in our new setting. As previously discussed, the only chance of pinning in Pl-#CSP must occur in the Hadamard basis. Thus, the plan is to pin in Pl-Holant($\mathcal{F} \cup \widehat{\mathcal{EQ}}$).

The other tool used in the Holant dichotomy is "mixing" results, though it is not as clearly delineated as in the current chapter (the whole of Section 8.3). Partial mixing with vanishing signatures was given in Section 7.3. The mixing results for signatures in $\mathscr{A}$ or $\mathscr{P}$ are implicit in the Pl-#CSP^d dichotomy, which was applied in each of the three dichotomies in Subsection 7.4.2. Anyway, after applying a single signature dichotomy, what you are left with is a bunch of individually tractable signatures. Thus, it is natural to prove some mixing results.

Furthermore, it is a good idea to prove this mixing result as soon as possible. The proofs of these dichotomy theorems are typically begun with a (possibly rough) conjecture of the tractable cases. By proving mixing results, you gain a clearer picture of the final dichotomy that you are trying to prove. For example, when we began [29], we knew that a new tractable case would involve vanishing signatures. But even after characterizing them (presented here in Section 4.4), we still didn't know the statement of the dichotomy that we wanted to prove. We had to determine what signatures combine with vanishing ones and remain tractable. This is where (attempting to) prove a dichotomy theorem helps you. You just try to prove as much hardness as possible. When you can prove no more, you stop and try to show instead that the problem is tractable.

Now the question to ask is: how can the mixing results be proved? This is generally done by reduction from small arity cases, which is (partially) why we first proved a planar Holant dichotomy for a single signature of arity 4 (presented here in Chapter 6). It also happened that we used the technique of domain pairing. This technique was also used to prove the Pl-#CSP dichotomy with real weights, so it was natural to see how this technique could help us as well.

Given this plan for the high-level proof structure, you then go and execute. The "design" of our proof at the "case" level (in Sections 8.2, 8.3, and 8.4) is simply a result of optimizing the proof.

Style of presentation I no longer like the

"problem A is #P-hard (or in P) iff problem B is #P-hard (or in P)"

style of presentation. It would not bother me so much if it were just used once in, say, the Pinning Theorem (Theorem 8.4.9). In total though, it is used five times, and the statements of these results are rather wordy because of it.

I think it is possible to reorganize and restate these results as various cases that can be nicely tied together in a newly worded pinning lemma. This style of presentation is similar to that in Chapter 11. Roughly speaking, the style used there can be described as follows. First, cases are shown to be #P-hard whenever possible. Second, if there is some tractable case, then its tractability is briefly mentioned in prose between hardness proofs to motivate why that case is excluded in the next proof of #P-hardness (and often why it is excluded in many future proofs of #P-hardness as well). Finally, everything is tied together in some capstone theorem that considers each case and either shows that it is tractable or shows that it is #P-hard.

Chapter 9

Interlude to Compute Some Gadget Signatures over General Domains

This chapter contains no complexity results. It exists so that complexity results can be proved in the next two chapters, which study the computational complexity of higher domain Holant problems. With higher domains, it becomes increasing difficult to verify that a particular gadget has a particular signature. This work was published in [32, 33].

9.1 Discussion

In the next two chapters, some of the more difficult claims to verify are those when we say that a particular $\mathcal{F}$-gate has a particular signature. This is an essential difficultly that cannot be avoided. We are proving that Pl-Holant$_\kappa(\mathcal{F})$ is #P-hard for various $\mathcal{F}$ (and computing the signature of an $\mathcal{F}$-gate is a generalization of this problem). Thus, one should not expect to be able to compute these signatures significantly faster in general than what the naive algorithm can do.

This has always been an issue for any dichotomy theorem about counting problems, but with larger domain sizes, we seem to be reaching the limit of what can be computed by hand for the signatures of gadget constructions that are presented in our proofs. To counter this, the standard techniques are to utilize the smallest gadgets (that suffice) or an infinite family of related gadgets with a (small) description of finite size, which we certainly employ. Additionally, we point out some tricks, where they exist, to save as much work as possible.

Beyond all this, we also face another problem. We would like to express the signature of a gadget as a function of the domain size. To compute the signature of a gadget for every domain size is no longer a finite computation. However, each entry of the gadget's signature is a polynomial in the domain size of degree at most the number of internal edges in the gadget. To obtain these polynomials, one can interpolate them by computing the signature for small domain sizes. It is easy to write a program to do this.

When computing by hand, there is another possibility that works quite well. One partitions the internal edge assignments into a limited number of parts such that the assignments in each part contribute the same quantity to the Holant sum. This is best explained with some examples.

9.2 Gadget Computations

The contents of this section would typically appear in an appendix (as in [32]). I have included it here so that no proof depends on a forward reference. I recommend that readers initially skip this section and return to it for reference as needed. Those feeling particularly motivated are welcome to try and verify the correctness of these proofs. If so, I highly recommend doing the computation both by hand and by computer. The kinds of mistakes that I tend to make when working by hand are not the the kinds of mistakes that I tend to make when coding (and vise versa). When working solely by hand, it can take me two to three hours to verify the most laborious of these calculations.

The statement of these lemmas and their proofs use the notion of a succinct signature and the succinct signature types defined in Section 10.1. They also use the following expressions:

$$\mathfrak{A} = a - 3b + 2c; \tag{9.2.1}$$

$$\mathfrak{B} = \mathfrak{A} + \kappa(b - c) \qquad = a + (\kappa - 3)b - (\kappa - 2)c; \qquad \text{and} \tag{9.2.2}$$

$$\mathfrak{C} = \mathfrak{B} + \kappa[2b + (\kappa - 2)c] \quad = a + 3(\kappa - 1)b + (\kappa - 1)(\kappa - 2)c. \tag{9.2.3}$$

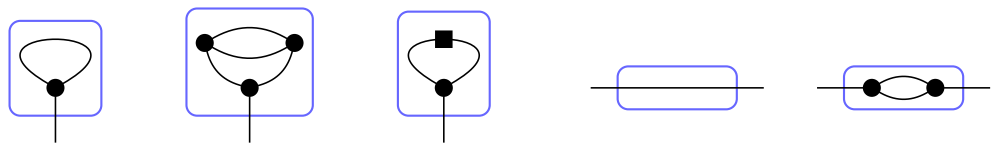


Figure 9.1: Gadgets (a) and (b) are used to construct $\langle 1 \rangle$. They are special cases of (c) and are obtained by replacing the square in (c) with either (d) or (e) respectively. All (circle) vertices are assigned $\langle a, b, c \rangle$.

Lemma 9.2.1. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c, x, y \in \mathbb{C}$. Let $\langle a, b, c \rangle$ be a succinct ternary signature of type τ_3 and let $\langle x, y \rangle$ be a succinct signature of type τ_2. If we assign $\langle a, b, c \rangle$ to the circle vertex and $\langle x, y \rangle$ to the square vertex of the gadget in Figure 9.1c, then the succinct unary signature of type τ_1 of the resulting gadget is $\langle x[a + (\kappa - 1)b] + y(\kappa - 1)[2b + (\kappa - 2)c] \rangle$.*

If the square vertex is replaced by Figure 9.1d, then the resulting signature is $\langle a + (\kappa - 1)b \rangle$. If the square vertex is replaced by Figure 9.1e, and $a + (\kappa - 1)b = 0$, then the resulting signature is

$$\langle -(\kappa - 1)(\kappa - 2)[2b + (\kappa - 2)c][b^2 - 4bc - (\kappa - 3)c^2] \rangle. \tag{9.2.4}$$

Proof. Since $\langle a, b, c \rangle$ and $\langle x, y \rangle$ are domain invariant, the signatures of these gadgets are also domain invariant. Any domain invariant unary signature has a succinct signature of type τ_1.

Let $g \in [\kappa]$ be a possible edge assignment, which we call a color. Suppose the external edge is assigned g and consider all internal edge assignments that assign the same colors to both edges. For such assignments, $\langle x, y \rangle$ contributes a factor of x. Now if this color assigned to both internal edges is also g, then $\langle a, b, c \rangle$ contributes a factor of a. Thus, the Holant sum includes one factor of ax. If the two internal edges are assigned any color different from g, then $\langle a, b, c \rangle$ contributes a factor of b. Since there are $\kappa - 1$ such colors, this adds $(\kappa - 1)bx$ to the Holant sum.

Now consider all internal assignments that assign different colors to the edges. For such assignments, $\langle x, y \rangle$ contributes a factor of y. First, suppose that one of the internal edges is assigned g. There are two ways this could happen and $\langle a, b, c \rangle$ contributes a factor of b. Since there are $\kappa - 1$ choices for the remaining edge assignment, this adds $2(\kappa - 1)by$ to the Holant sum. Lastly, suppose that the two internal edges are not assigned g. Then $\langle a, b, c \rangle$ contributes a factor of c. Since there are $(\kappa - 1)(\kappa - 2)$ such assignments, this adds $(\kappa - 1)(\kappa - 2)cy$ to the Holant sum. Thus, the resulting signature is $\langle x[a + (\kappa - 1)b] + y(\kappa - 1)[2b + (\kappa - 2)c]\rangle$ as claimed.

Replacing the square by Figure 9.1d is equivalent to setting $x = 1$ and $y = 0$, which gives $\langle a + (\kappa - 1)b\rangle$. Replacing the square by Figure 9.1e is equivalent to setting x and y to the values given in Lemma 9.2.2. The resulting signature is indeed (9.2.4). □

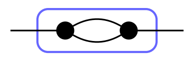


Figure 9.2: A simple binary gadget.

Lemma 9.2.2. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c \in \mathbb{C}$. Let $\langle a, b, c\rangle$ be a succinct ternary signature of type τ_3. If we assign $\langle a, b, c\rangle$ to both vertices of the gadget in Figure 9.2, then the succinct binary signature of type τ_2 of the resulting gadget is $\langle x, y\rangle$, where*

$$x = a^2 + 3(\kappa - 1)b^2 + (\kappa - 1)(\kappa - 2)c^2 \qquad \textit{and}$$
$$y = 2ab + \kappa b^2 + 4(\kappa - 2)bc + (\kappa - 2)(\kappa - 3)c^2.$$

Proof. Since $\langle a, b, c\rangle$ is domain invariant, the signature of this gadget is also domain invariant. Any domain invariant binary signature has a succinct signature of type τ_2.

Let $g, r \in [\kappa]$ be distinct edge assignments. We have two entries to compute. To compute x, suppose that both external edges are assigned g. We begin with the case where both internal edges have the same assignment. If this assignment is g, then a^2 is contributed to the sum. If this assignment is not g, then b^2 is contributed to the sum for a total contribution of $(\kappa - 1)b^2$. Now consider the case that the two internal edges have a different assignment. If one of these assignments is g, then b^2 is contributed to the sum for a total contribution of $2(\kappa - 1)b^2$. If neither assignment is g, then c^2 is contributed to the sum for a total contribution of $(\kappa - 1)(\kappa - 2)c^2$. These total contributions sum to the value for x given in Lemma 9.2.2.

To compute y, suppose one external edge is assigned g and the other is assigned r. We begin with the case where both internal edges have the same assignment. If this assignment is g or r, then ab is contributed to the sum for a total contribution of $2ab$. If this assignment is not g or r, then

b^2 is contributed to the sum for a total contribution of $(\kappa - 2)b^2$. Now consider the case that the two internal edges have a different assignment. If both are assigned g or r, then b^2 is contributed to the sum for a total contribution of $2b^2$. If exactly one is assigned g or r, then bc is contributed to the sum for a total contribution of $4(\kappa - 2)bc$. If neither is assigned g or r, then c^3 is contributed to the sum for a total contribution of $(\kappa - 2)(\kappa - 3)c^3$. These total contributions sum to the value for y given in Lemma 9.2.2. □

When checking these proofs, a concern is that some assignments might not have been counted (or maybe some assignment was counted twice). A sanity check to address this concern is to set $a = b = c = 1$ and inspect the resulting expression. If computed correctly, the result will be κ^m, where m is the number of internal edges, which is the number of internal edge assignments.

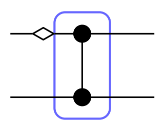


Figure 9.3: A simple quaternary gadget.

Lemma 9.2.3. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c \in \mathbb{C}$. Let $\langle a, b, c \rangle$ be a succinct ternary signature of type τ_3. If we assign $\langle a, b, c \rangle$ to both vertices of the gadget in Figure 9.3, then the succinct quaternary signature of type τ_4 of the resulting gadget is*

$$f = \left\langle f_{\substack{1\,1\\1\,1}}, f_{\substack{1\,2\\1\,1}}, f_{\substack{1\,2\\1\,2}}, f_{\substack{1\,3\\1\,2}}, f_{\substack{1\,2\\2\,1}}, f_{\substack{1\,3\\2\,1}}, f_{\substack{1\,1\\2\,2}}, f_{\substack{1\,1\\2\,3}}, f_{\substack{1\,4\\2\,3}} \right\rangle,$$

where

$$f_{\substack{1\,1\\1\,1}} = a^2 + (\kappa - 1)b^2,$$

$$f_{\substack{1\,2\\1\,1}} = b[a + b + (\kappa - 2)c],$$

$$f_{\substack{1\,2\\1\,2}} = 2b^2 + (\kappa - 2)c^2,$$

$$f_{\substack{1\,3\\1\,2}} = b^2 + 2bc + (\kappa - 3)c^2,$$

$$f_{\substack{1\,2\\2\,1}} = f_{\substack{1\,2\\1\,2}},$$

$$f_{\substack{1\,3\\2\,1}} = f_{\substack{1\,3\\1\,2}},$$

$$f_{\substack{1\,1\\2\,2}} = b[2a + (\kappa - 2)b],$$

$$f_{\substack{1\,1\\2\,3}} = ac + 2b^2 + (\kappa - 3)bc, \text{ and}$$

$$f_{\substack{1\,4\\2\,3}} = c[4b + (\kappa - 4)c].$$

Proof. Since $\langle a, b, c\rangle$ is domain invariant, the signature of this gadget is also domain invariant. The vertical and horizontal symmetry of this gadget implies that its signature has a succinct signature of type τ_4.

Let $w, x, y, z \in [\kappa]$ be distinct edge assignments. We have nine entries to compute. Recall that the edge with the diamond is considered the first input and the rest are ordered counterclockwise.

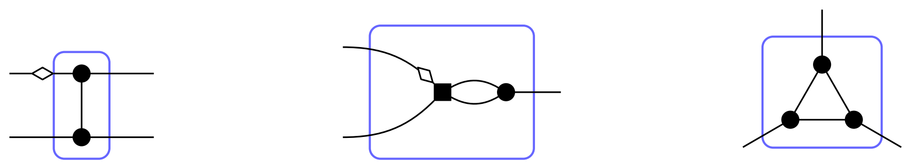
1. To compute $f_{\substack{1\,1\\1\,1}}$, suppose the external assignment is (w, w, w, w). If the internal edge is also assigned w, then a^2 is contributed to the sum. If the internal edge is not assigned w, then b^2 is contributed to the sum for a total contribution of $(\kappa - 1)b^2$.
2. To compute $f_{\substack{1\,2\\1\,1}}$, suppose the external assignment is (w, w, w, x). If the internal edge is assigned w, then ab is contributed to the sum. If the internal edge is assigned x, then b^2 is contributed to the sum. If the internal edge is not assigned w or x, then bc is contributed to the sum for a total contribution of $(\kappa - 2)bc$.
3. To compute $f_{\substack{1\,2\\1\,2}}$, suppose the external assignment is (w, w, x, x). If the internal edge is assigned w, then b^2 is contributed to the sum. If the internal edge is assigned x, then b^2 is contributed to the sum. If the internal edge is not assigned w or x, then c^2 is contributed to

the sum for a total contribution of $(\kappa - 2)c^2$.

4. To compute $f_{\substack{1\,3\\1\,2}}$, suppose the external assignment is (w, w, x, y). If the internal edge is assigned w, then b^2 is contributed to the sum. If the internal edge is assigned x, then bc is contributed to the sum. If the internal edge is assigned y, then bc is contributed to the sum. If the internal edge is not assigned w, x or y, then c^2 is contributed to the sum for a total contribution of $(\kappa - 3)c^2$.
5. To compute $f_{\substack{1\,3\\1\,2}}$, suppose the external assignment is (w, x, w, x). This entry is the same as that for (w, w, x, x). The reason is that the signature is unchanged if the two external edges of the lower vertex are swapped since $\langle a, b, c\rangle$ is symmetric.
6. To compute $f_{\substack{1\,3\\2\,1}}$, suppose the external assignment is (w, x, w, y). This entry is the same as that for (w, w, x, y) for the same reason as the previous entry.
7. To compute $f_{\substack{1\,1\\2\,2}}$, suppose the external assignment is (w, x, x, w). If the internal edge is assigned w, then ab is contributed to the sum. If the internal edge is assigned x, then ab is contributed to the sum. If the internal edge is not assigned w or x, then b^2 is contributed to the sum for a total contribution of $(\kappa - 2)b^2$.
8. To compute $f_{\substack{1\,1\\2\,3}}$, suppose the external assignment is (w, x, y, w). If the internal edge is assigned w, then ac is contributed to the sum. If the internal edge is assigned x, then b^2 is contributed to the sum. If the internal edge is assigned y, then b^2 is contributed to the sum. If the internal edge is not assigned w, x or y, then bc is contributed to the sum for a total contribution of $(\kappa - 3)c^2$.
9. To compute $f_{\substack{1\,4\\2\,3}}$, suppose the external assignment is (w, x, y, z). If the internal edge is assigned w, x, y, or z, then bc is contributed to the sum for a total contribution of $4bc$. If the internal edge is not assigned w, x, y or z, then c^2 is contributed to the sum for a total contribution of $(\kappa - 4)c^2$.

These total contributions each sum to their corresponding entry of f given in the statement of Lemma 9.2.3. □

Although possible, it would be difficult to compute the signature of the gadget in Figure 9.4c through partitioning of the internal edge assignments alone. To simplify matters, we utilize the calculations from Lemma 9.2.3. Since composing the gadget in Figure 9.4a with the one in Figure 9.4b gives a symmetric signature, we refrain from distinguishing the external edges of the gadget in Figure 9.4b.



(a) Inner structure (b) Outer structure (c) Entire binary gadget

Figure 9.4: Decomposition of a ternary gadget. All circle vertices are assigned $\langle a, b, c\rangle$ and the square vertex in (b) is assigned the signature of the gadget in (a).

Lemma 9.2.4. *Suppose $\kappa \ge 3$ is the domain size and $a, b, c \in \mathbb{C}$. Let $\langle a, b, c\rangle$ be a succinct ternary signature of type τ_3. If we assign $\langle a, b, c\rangle$ to all vertices of the gadget in Figure 9.4c, then the succinct ternary signature of type τ_3 of the resulting gadget is $\langle a', b', c'\rangle$, where*

$$\begin{aligned}
a' &= a^3 + 3(\kappa-1)ab^2 + 4(\kappa-1)b^3 + 3(\kappa-1)(\kappa-2)(b^2c+bc^2) + (\kappa-1)(\kappa-2)(\kappa-3)c^3,\\
b' &= a^2b + 4ab^2 + 2(\kappa-2)abc + (\kappa-2)ac^2 + (5\kappa-7)b^3 + (\kappa-2)(\kappa+5)b^2c\\
&\qquad + (\kappa-2)(7\kappa-18)bc^2 + (\kappa-2)(\kappa-3)^2c^3, \qquad \textit{and}\\
c' &= 3ab^2 + 6abc + 3(\kappa-3)ac^2 + (\kappa+5)b^3 + 3(7\kappa-18)b^2c + 9(\kappa-3)^2bc^2\\
&\qquad + (\kappa^3 - 9\kappa^2 + 29\kappa - 32)c^3.
\end{aligned}$$

Furthermore, if $\mathfrak{A} = 0$, then

$$\begin{aligned}
a' &= 3b' - 2c',\\
b' &= (5\kappa+14)b^3 + (\kappa^2+9\kappa-42)b^2c + (7\kappa^2-33\kappa+42)bc^2 + (\kappa-2)(\kappa^2-6\kappa+7)c^3, \qquad \textit{and}\\
c' &= (\kappa+14)b^3 + 21(\kappa-2)b^2c + 3(3\kappa^2-15\kappa+14)bc^2 + (\kappa^3-9\kappa^2+23\kappa-14)c^3.
\end{aligned}$$

Proof. Since $\langle a, b, c\rangle$ is domain invariant, the signature of this gadget is also domain invariant. As a ternary signature, the rotational symmetry of this gadget implies the symmetry of the signature.

Any symmetric domain invariant ternary signature has a succinct signature of type τ_3.

Consider the gadget in Figure 9.4a. We assign $\langle a, b, c\rangle$ to both vertices. Then by Lemma 9.2.3, the succinct quaternary signature of this gadget is the signature f given in Lemma 9.2.3.

Now consider the gadget in Figure 9.4b. We assign $\langle a, b, c\rangle$ to the circle vertex and f to the square vertex. The resulting gadget is the one in Figure 9.4c, which is symmetric. Thus, there is no need to distinguish the external edges. We have three entries to compute.

Let $g, r, y \in [\kappa]$ be distinct edge assignments. To compute a', suppose that all external edges are assigned g. We begin with the case where both internal edges have the same assignment. If this assignment is g, then $af_{\substack{1\,1\\1\,1}}$ is contributed to the sum. If this assignment is not g, then $bf_{\substack{1\,2\\1\,2}}$ is contributed to the sum for a total contribution of $(\kappa-1)bf_{\substack{1\,2\\1\,2}}$. Now consider the case that the two internal edges have a different assignment. If one of these assignments is g, then $bf_{\substack{1\,2\\1\,1}}$ is contributed to the sum for a total contribution of $2(\kappa-1)bf_{\substack{1\,2\\1\,1}}$. If neither assignment is g, then $cf_{\substack{1\,3\\1\,2}}$ is contributed to the sum for a total contribution of $(\kappa-1)(\kappa-2)cf_{\substack{1\,3\\1\,2}}$. After substituting for the entries of f, these total contributions sum to the value for a' given in Lemma 9.2.4.

To compute b', suppose the left external edges are assigned g and the right external edge is assigned r. We begin with the case where both internal edges have the same assignment. If this assignment is g, then $bf_{\substack{1\,1\\1\,1}}$ is contributed to the sum. If this assignment is r, then $af_{\substack{1\,2\\1\,2}}$ is contributed to the sum. If this assignment is not g or r, then $bf_{\substack{1\,2\\1\,2}}$ is contributed to the sum for a total contribution of $(\kappa-2)bf_{\substack{1\,2\\1\,2}}$. Now consider the case that the two internal edges have a different assignments. If both are assigned g or r, then $bf_{\substack{1\,2\\1\,1}}$ is contributed to the sum for a total contribution of $2bf_{\substack{1\,2\\1\,1}}$. If one is assigned g and the other is not assigned r, then $cf_{\substack{1\,2\\1\,1}}$ is contributed to the sum for a total contribution of $2(\kappa-2)cf_{\substack{1\,2\\1\,1}}$. If one is assigned r and the other is not assigned g, then $bf_{\substack{1\,3\\1\,2}}$ is contributed to the sum for a total contribution of $2(\kappa-2)bf_{\substack{1\,3\\1\,2}}$. If neither is assigned g or r, then $cf_{\substack{1\,3\\1\,2}}$ is contributed to the sum for a total contribution of $(\kappa-2)(\kappa-3)cf_{\substack{1\,3\\1\,2}}$. After substituting for the entries of f, these total contributions sum to the value for b' given in Lemma 9.2.4.

To compute c', suppose the upper-left external edge is assigned g, the lower-left external edge is assigned r, and the right external edge is assigned y. We begin with the case where both internal edges have the same assignment. If this assignment is g, then $bf_{\begin{smallmatrix}1&2\\1&1\end{smallmatrix}}$ is contributed to the sum. If this assignment is r, then $bf_{\begin{smallmatrix}1&2\\1&1\end{smallmatrix}}$ is contributed to the sum. If this assignment is y, then $af_{\begin{smallmatrix}1&3\\1&2\end{smallmatrix}}$ is contributed to the sum. If this assignment is not g, r, or y, then $bf_{\begin{smallmatrix}1&3\\1&2\end{smallmatrix}}$ is contributed to the sum for a total contribution of $(\kappa - 3)bf_{\begin{smallmatrix}1&3\\1&2\end{smallmatrix}}$. Now consider the case that the two internal edges have a different assignments. If the top internal edge is assigned g and the bottom one is assigned r, then $cf_{\begin{smallmatrix}1&1\\2&2\end{smallmatrix}}$ is contributed to the sum. If the top internal edge is assigned r and the bottom one is assigned g, then $cf_{\begin{smallmatrix}1&2\\2&1\end{smallmatrix}}$ is contributed to the sum. If the top internal edge is assigned g and the bottom one is assigned y, then $bf_{\begin{smallmatrix}1&1\\2&3\end{smallmatrix}}$ is contributed to the sum. If the top internal edge is assigned y and the bottom one is assigned g, then $bf_{\begin{smallmatrix}1&3\\2&1\end{smallmatrix}}$ is contributed to the sum. If the top internal edge is assigned r and the bottom one is assigned y, then $bf_{\begin{smallmatrix}1&3\\2&1\end{smallmatrix}}$ is contributed to the sum. If the top internal edge is assigned y and the bottom one is assigned r, then $bf_{\begin{smallmatrix}1&1\\2&3\end{smallmatrix}}$ is contributed to the sum. If the top internal edge is assigned g and the bottom one not assigned r or y, then $cf_{\begin{smallmatrix}1&1\\2&3\end{smallmatrix}}$ is contributed to the sum for a total contribution of $(\kappa - 3)cf_{\begin{smallmatrix}1&1\\2&3\end{smallmatrix}}$. If the bottom internal edge is assigned g and the top one not assigned r or y, then $cf_{\begin{smallmatrix}1&3\\2&1\end{smallmatrix}}$ is contributed to the sum for a total contribution of $(\kappa - 3)cf_{\begin{smallmatrix}1&3\\2&1\end{smallmatrix}}$. If the top internal edge is assigned r and the bottom one not assigned g or y, then $cf_{\begin{smallmatrix}1&3\\2&1\end{smallmatrix}}$ is contributed to the sum for a total contribution of $(\kappa - 3)cf_{\begin{smallmatrix}1&3\\2&1\end{smallmatrix}}$. If the bottom internal edge is assigned r and the top one not assigned g or y, then $cf_{\begin{smallmatrix}1&1\\2&3\end{smallmatrix}}$ is contributed to the sum for a total contribution of $(\kappa - 3)cf_{\begin{smallmatrix}1&1\\2&3\end{smallmatrix}}$. If the one internal edge is assigned y and the other is not assigned g or r, then $bf_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}}$ is contributed to the sum for a total contribution of $2(\kappa - 3)bf_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}}$. If neither internal edge is assigned g r, or y, then $cf_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}}$ is contributed to the sum for a total contribution of $(\kappa - 3)(\kappa - 4)cf_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}}$. After substituting for the entries of f, these total contributions sum to the value for c' given in Lemma 9.2.4. □

The signature of the gadget in Figure 9.5 is difficult to compute using gadget compositions and partitioning of internal edge assignments as we have been doing. Instead, we compute this signature using matrix product, trace, and polynomial interpolation. One can use the same approach to compute the signature of the gadget in Figure 9.4c as well.

Lemma 9.2.5. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c, x_1, y_1, x_2, y_2 \in \mathbb{C}$. Let $\langle a, b, c \rangle$ be a succinct ternary signature of type τ_3 and $\langle x_1, y_1 \rangle$ and $\langle x_2, y_2 \rangle$ be succinct binary signatures of type*

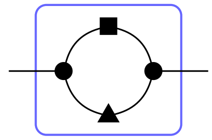


Figure 9.5: A more complicated binary gadget.

τ_2. *If to the gadget in Figure 9.5 we assign* $\langle a, b, c\rangle$ *to the circle vertices,* $\langle x_1, y_1\rangle$ *to the square vertex, and* $\langle x_2, y_2\rangle$ *to the triangle vertex, then the succinct binary signature of type* τ_2 *of the resulting gadget is* $\langle x, y\rangle$*, where*

$$\begin{aligned} x = {} & x_1x_2a^2 + 2(\kappa-1)(x_1y_2 + x_2y_1 + y_1y_2)ab + 2(\kappa-1)(\kappa-2)y_1y_2ac \\ & + (\kappa-1)[3x_1x_2 + \kappa(x_1y_2 + x_2y_1) + (7\kappa-12)y_1y_2]b^2 \\ & + 2(\kappa-1)(\kappa-2)[2(x_1y_2 + x_2y_1) + (3\kappa-7)y_1y_2]bc \\ & + (\kappa-1)(\kappa-2)[x_1x_2 + (\kappa-3)(x_1y_2 + x_2y_1) + (\kappa^2-5\kappa+7)y_1y_2]c^2 \qquad \textit{and} \\ y = {} & y_1y_2a^2 + 2[x_1x_2 + x_1y_2 + x_2y_1 + 3(\kappa-2)y_1y_2]ab + 2(\kappa-2)[x_1y_2 + x_2y_1 + (\kappa-3)y_1y_2]ac \\ & + [\kappa x_1x_2 + (7\kappa-12)(x_1y_2 + x_2y_1) + 3(3\kappa^2-11\kappa+11)y_1y_2]b^2 \\ & + 2(\kappa-2)[2x_1x_2 + (3\kappa-7)(x_1y_2 + x_2y_1) + 3(\kappa^2-4\kappa+5)y_1y_2]bc \\ & + (\kappa-2)[(\kappa-3)x_1x_2 + (\kappa^2-5\kappa+7)(x_1y_2 + x_2y_1) + (\kappa^3-6\kappa^2+14\kappa-13)]c^2. \end{aligned}$$

Furthermore, if $\langle x_1, y_1\rangle = \frac{1}{\kappa}\langle \omega^r + \kappa - 1, \omega^r - 1\rangle$ *and* $\langle x_2, y_2\rangle = \frac{1}{\kappa}\langle \omega^s + \kappa - 1, \omega^s - 1\rangle$*, then*

$$x = \frac{\mathfrak{B}^2}{\kappa^2}\left[\Phi\omega^{r+s} + (\kappa-1)(\omega^r + \omega^s + \Psi + 1)\right] \quad \textit{and} \quad y = \frac{\mathfrak{B}^2}{\kappa^2}\left[\Phi\omega^{r+s} - (\omega^r + \omega^s + \Psi + 1) + \kappa\right],$$

where $\Phi = \frac{\mathfrak{C}^2}{\mathfrak{B}^2}$ *and* $\Psi = \frac{(\kappa-2)\mathfrak{A}^2}{\mathfrak{B}^2}$.

Proof. Since $\langle a, b, c\rangle$, $\langle x_1, y_1\rangle$, and $\langle x_2, y_2\rangle$ are domain invariant, the signature of this gadget is also domain invariant. Any domain invariant binary signature has a succinct signature of type τ_2.

We compute a', b', and c' using the algorithm for $\text{Holant}_\kappa(\mathcal{F})$ when every non-degenerate signature in $\mathcal{F}$ is of arity at most 2, which is to use matrix product and trace. Then we finish with polynomial interpolation. Let $M_\kappa(t)$ be a κ-by-κ matrix such that

$$(M_\kappa(t))_{i,j} = \begin{cases} a & i = j = t \\ b & i = j \neq t \\ b & i \neq j \text{ and } (i = t \text{ or } j = t) \\ c & \text{otherwise.} \end{cases}$$

For example, $M_4(1) = \left[\begin{smallmatrix} a & b & b & b \\ b & b & c & c \\ b & c & b & c \\ b & c & c & b \end{smallmatrix}\right]$. If we fix an input of $\langle a, b, c\rangle$ to $t \in [\kappa]$, then the resulting binary signature (which is no longer domain invariant) has the signature matrix $M_\kappa(t)$.

Consider x and y as polynomials in κ with coefficients in $\mathbb{Z}[a, b, c, x_1, y_1, x_2, y_2]$. Then

$$x(\kappa) = \operatorname{tr}\Big(M_\kappa(1)\big[y_1 J_\kappa + (x_1 - y_1) I_\kappa\big] M_\kappa(1)\big[y_2 J_\kappa + (x_2 - y_2) I_\kappa\big]\Big) \qquad \text{and}$$
$$y(\kappa) = \operatorname{tr}\Big(M_\kappa(1)\big[y_1 J_\kappa + (x_1 - y_1) I_\kappa\big] M_\kappa(2)\big[y_2 J_\kappa + (x_2 - y_2) I_\kappa\big]\Big).$$

Since there are just four internal edges in this gadget, both of $x(\kappa)$ and $y(\kappa)$ are of degree at most 4 in κ. Therefore, we interpolate each of these polynomials using their evaluations at $3 \le \kappa \le 7$ and obtain the expressions for x and y given in Lemma 9.2.5. □

Remark. Lemma 9.2.2 is the special case of Lemma 9.2.5 with $\langle x_1, y_1\rangle = \langle x_2, y_2\rangle = \langle 1, 0\rangle$.

In order to apply a holographic transformation on a particular signature, it is convenient to express the signature as a sum of degenerate signatures. Let $e_{\kappa,i}$ be the standard basis vector of length κ with a 1 at location i and 0 elsewhere. Also let $\mathbf{1}_\kappa$ be the all 1's vector of length κ. Then the succinct ternary signature $\langle a, b, c\rangle$ on domain size κ can be expressed as

$$\langle a,b,c\rangle = c\mathbf{1}_\kappa^{\otimes 3} + (a-c)\sum_{i=1}^{\kappa} e_{\kappa,i}^{\otimes 3} + (b-c)\sum_{\substack{i,j\in[\kappa]\\ i\neq j}} \begin{pmatrix} & e_{\kappa,i}\otimes e_{\kappa,i}\otimes e_{\kappa,j} \\ + & e_{\kappa,i}\otimes e_{\kappa,j}\otimes e_{\kappa,i} \\ + & e_{\kappa,j}\otimes e_{\kappa,i}\otimes e_{\kappa,i} \end{pmatrix} \tag{9.2.5}$$

$$= b\mathbf{1}_\kappa^{\otimes 3} + (a-b)\sum_{i=1}^{\kappa} e_{\kappa,i}^{\otimes 3} + (c-b)\sum_{\substack{\sigma:1,2,3\to[\kappa]\\ \sigma \text{ injective}}} e_{\kappa,\sigma(1)}\otimes e_{\kappa,\sigma(2)}\otimes e_{\kappa,\sigma(3)}. \tag{9.2.6}$$

The expression in (9.2.5) contains $1+\kappa+3\kappa(\kappa-1) = 3\kappa^2-2\kappa+1$ summands. In general, this is smaller than the one in (9.2.6), which contains $1+\kappa+\kappa(\kappa-1)(\kappa-2) = \kappa^3-3\kappa^2+3\kappa+1$ summands. It is advantageous to find an expression that minimizes the number of summands. This leads to less computation in the proof of Lemma 9.2.6. However, determining the fewest number of summands for a given signature is exactly the problem of determining tensor rank, which is a problem well-known to be difficult [76].

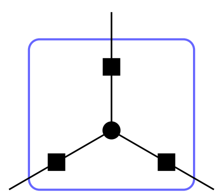


Figure 9.6: Local holographic transformation gadget construction for a ternary signature.

There is a gadget construction that mimics the behavior of a holographic transformation. This construction is called a local holographic transformation [53]. For $x, y \in \mathbb{C}$, let $\langle x, y\rangle$ be a succinct binary signature of type τ_2. Consider the gadget in Figure 9.6. If we assign $\langle a, b, c\rangle$ to the circle vertex and $\langle x, y\rangle$ to the square vertex, then the resulting signature of this gadget is the same as applying a holographic transformation on $\langle a, b, c\rangle$ with basis $T = yJ_\kappa + (x-y)I_\kappa$. We use this fact in the following proof.

Lemma 9.2.6. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c, x, y \in \mathbb{C}$. Let $\langle a, b, c\rangle$ be a succinct* *signature of type* τ_3 *and let* $T = yJ_\kappa + (x-y)I_\kappa$*. Then* $T^{\otimes 3}\langle a,b,c\rangle = \langle a',b',c'\rangle$*, where*

$$\begin{aligned}
a' &= a\left[x^3 + (\kappa-1)y^3\right] \\
&\quad + 3b(\kappa-1)\left[x^2y + xy^2 + (\kappa-2)y^3\right] \\
&\quad + c(\kappa-1)(\kappa-2)\left[3xy^2 + (\kappa-3)y^3\right] \\
b' &= a\left[x^2y + xy^2 + (\kappa-2)y^3\right] \\
&\quad + b\left[x^3 + \kappa x^2y + (7\kappa-12)xy^2 + (3\kappa^2-11\kappa+11)y^3\right] \\
&\quad + c(\kappa-2)\left[2x^2y + (3\kappa-7)xy^2 + (\kappa^2-4\kappa+5)y^3\right], \text{ and} \\
c' &= a\left[3xy^2 + (\kappa-3)y^3\right] \\
&\quad + 3b\left[2x^2y + (3\kappa-7)xy^2 + (\kappa^2-4\kappa+5)y^3\right] \\
&\quad + c\left[x^3 + 3(\kappa-3)x^2y + 3(\kappa^2-5\kappa+7)xy^2 + (\kappa^3-6\kappa^2+14\kappa-13)y^3\right].
\end{aligned}$$

In particular,

$$a' - b' = (x-y)^2[2\mathfrak{D} + \mathfrak{A}(x-y)] \qquad \text{and} \qquad b' - c' = (x-y)^2\mathfrak{D},$$

where $\mathfrak{D} = (b-c)(x-y) + \mathfrak{B}y$*. Furthermore, if* $\mathfrak{A} = 0$*, then*

$$\begin{aligned}
a' &= 3b' - 2c', \\
b' &= [x + (\kappa-1)y]\left\{bx^2 + 2[2b + (\kappa-3)c]xy + [(3\kappa-5)b + (\kappa^2-5\kappa+6)c]y^2\right\} \qquad \text{and} \\
c' &= [x + (\kappa-1)y]\left\{cx^2 + 2[3b + (\kappa-4)c]xy + [(3\kappa-6)b + (\kappa^2-5\kappa+7)c]y^2\right\}.
\end{aligned}$$

If $\kappa = 3$*,* $x = -1$*, and* $y = 2$*, then*

$$a' = -3(5a + 12b - 8c), \qquad b' = -3(2a + 3b + 4c), \qquad \text{and} \qquad c' = 3(4a - 12b - c).$$

Proof. Let $\widehat{f} = T^{\otimes 3}\langle a,b,c\rangle$. Since $\langle a,b,c\rangle$ and $\langle x,y\rangle$ are domain invariant, the signature of the gadget in Figure 9.6, which is the same signature $\widehat{f}$, is also domain invariant. As a ternary signature, the rotational symmetry of this gadget implies the symmetry of the signature. Any symmetric domain invariant ternary signature has a succinct signature of type τ_3.

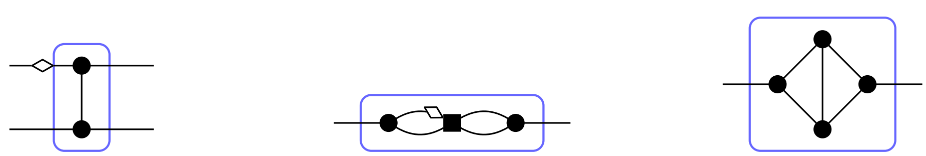
The entries of $\widehat{f}$ are polynomials in κ with coefficients from $\mathbb{Z}[a,b,c,x,y]$. The degree of these polynomials is at most 3 since the arity of $\langle a,b,c\rangle$ is 3. We compute the entries of $\widehat{f} = T^{\otimes 3}\langle a,b,c\rangle$ as elements in $\mathbb{Z}[a,b,c,x,y]$ for domain sizes $3 \le \kappa \le 6$ by replacing $\langle a,b,c\rangle$ with an equivalent expression from either (9.2.5) or (9.2.6). Then we interpolate the entries of $\widehat{f}$ as elements in $(\mathbb{Z}[a,b,c,x,y])[\kappa]$. The resulting expressions for the signature entries are as given in the statement of Lemma 9.2.6.

It is straightforward to verify the expressions for $a' - b'$ and $b' - c'$ given those for a', b', and c'. Recall that $\mathfrak{A} = a - 3b + 2c$. If $\mathfrak{A} = 0$, then it follows that $a' - 3b' + 2c' = 0$ as well since

$$\begin{aligned} a' - 3b' + 2c' &= a' - b' - 2(b' - c') \\ &= (x-y)^2[2\mathfrak{D} + \mathfrak{A}(x-y)] - 2(x-y)^2\mathfrak{D} \\ &= \mathfrak{A}(x-y)^3 = 0. \end{aligned}$$

The expressions for b' and c' when $\mathfrak{A} = 0$ directly follow from their general expressions above. □

By composing smaller gadgets, we can easily compute the signatures of rather large gadgets.



(a) Inner structure (b) Outer structure (c) Entire binary gadget

Figure 9.7: Decomposition of a binary gadget. All circle vertices are assigned $\langle a,b,c\rangle$ and the square vertex in (b) is assigned the signature of the gadget in (a).

Lemma 9.2.7. *Suppose $\kappa \ge 3$ is the domain size and $a,b,c \in \mathbb{C}$. Let $\langle a,b,c\rangle$ be a succinct ternary signature of type τ_3. If $\langle a,b,c\rangle$ is assigned to every vertex of the gadget in Figure 9.7c, then the* *resulting signature is the succinct binary signature $\langle x, y\rangle$ of type τ_2, where*

$$\begin{aligned}
x = {} & a^4 + 6(\kappa-1)a^2b^2 + 16(\kappa-1)ab^3 + 12(\kappa-1)(\kappa-2)ab^2c + 12(\kappa-1)(\kappa-2)abc^2 \\
& + 4(\kappa-1)(\kappa-2)(\kappa-3)ac^3 + 3(\kappa-1)(5\kappa-7)b^4 + 4(\kappa-1)(\kappa-2)(\kappa+5)b^3c \\
& + 6(\kappa-1)(\kappa-2)(7\kappa-18)b^2c^2 + 12(\kappa-3)^2(\kappa-1)(\kappa-2)bc^3 \\
& + (\kappa-1)(\kappa-2)(\kappa^3-9\kappa^2+29\kappa-32)c^4 \qquad \textit{and} \\
y = {} & 2a^3b + (\kappa+4)a^2b^2 + 4(\kappa-2)a^2bc + (\kappa-2)a^2c^2 + 2(9\kappa-11)ab^3 + 2(\kappa-2)(3\kappa+8)ab^2c \\
& + 2(\kappa-2)(12\kappa-31)abc^2 + 2(\kappa-2)(2\kappa^2-11\kappa+16)ac^3 + (7\kappa^2+3\kappa-24)b^4 \\
& + 2(\kappa-2)(\kappa^2+31\kappa-70)b^3c + (\kappa-2)(48\kappa^2-234\kappa+301)b^2c^2 \\
& + 2(\kappa-2)(6\kappa^3-45\kappa^2+121\kappa-116)bc^3 + (\kappa-2)(\kappa-3)(\kappa^3-7\kappa^2+19\kappa-20)c^4.
\end{aligned}$$

Proof. Since $\langle a, b, c\rangle$ is domain invariant, the signature of this gadget is also domain invariant. Any domain invariant binary signature has a succinct signature of type τ_2.

Consider the gadget in Figure 9.7a. We assign $\langle a, b, c\rangle$ to both vertices. By Lemma 9.2.3, this gadget has the succinct quaternary signature f of type τ_4, where f is given in Lemma 9.2.3.

Now consider the gadget in Figure 9.7b. We assign $\langle a, b, c\rangle$ the circle vertices and f to the square vertex. By partitioning the internal edge assignments into parts with the same contribution

Table 9.1: The signature of the gadget in Figure 9.8c is $\langle x, y\rangle$ for the x and y above.

$$\begin{aligned}
x = {} & a^6 + 9(\kappa-1)a^4b^2 + 32(\kappa-1)a^3b^3 + 18(\kappa-1)(\kappa-2)a^3b^2c + 12(\kappa-1)(\kappa-2)a^3bc^2 \\
& + 2(\kappa-1)(\kappa-2)(\kappa-3)a^3c^3 + 3(\kappa-1)(16\kappa-7)a^2b^4 + 6(\kappa-1)(\kappa-2)(\kappa+19)a^2b^3c \\
& + 18(\kappa-1)(\kappa-2)(4\kappa-7)a^2b^2c^2 + 6(\kappa-1)(\kappa-2)(\kappa^2+2\kappa-13)a^2bc^3 \\
& + 3(\kappa-1)(\kappa-2)(3\kappa^2-17\kappa+25)a^2c^4 + 6(\kappa-1)(\kappa^2+27\kappa-42)ab^5 \\
& + 6(\kappa-1)(\kappa-2)(40\kappa-41)ab^4c + 24(\kappa-1)(\kappa-2)(3\kappa^2+8\kappa-36)ab^3c^2 \\
& + 6(\kappa-1)(\kappa-2)(\kappa^3+50\kappa^2-285\kappa+393)ab^2c^3 \\
& + 6(\kappa-1)(\kappa-2)(13\kappa^3-108\kappa^2+311\kappa-307)abc^4 \\
& + 6(\kappa-1)(\kappa-2)(\kappa-3)(\kappa^3-8\kappa^2+24\kappa-26)ac^5 \\
& + (\kappa-1)(\kappa^3+83\kappa^2-189\kappa+81)b^6 + 18(\kappa-1)(\kappa-2)(4\kappa^2+13\kappa-43)b^5c \\
& + 3(\kappa-1)(\kappa-2)(7\kappa^3+222\kappa^2-1156\kappa+1442)b^4c^2 \\
& + 2(\kappa-1)(\kappa-2)(\kappa^4+221\kappa^3-1725\kappa^2+4576\kappa-4153)b^3c^3 \\
& + 3(\kappa-1)(\kappa-2)(43\kappa^4-441\kappa^3+1791\kappa^2-3393\kappa+2505)b^2c^4 \\
& + 6(\kappa-1)(\kappa-2)(\kappa-3)(3\kappa^4-29\kappa^3+116\kappa^2-228\kappa+182)bc^5 \\
& + (\kappa-1)(\kappa-2)(\kappa^6-15\kappa^5+98\kappa^4-361\kappa^3+798\kappa^2-1004\kappa+556)c^6
\end{aligned}$$

and

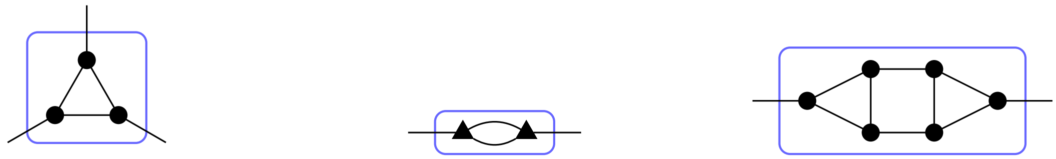
$$\begin{aligned}
y = {} & 2a^5b + (\kappa+8)a^4b^2 + 4(\kappa-2)a^4bc + 2(\kappa-2)a^4c^2 + 4(9\kappa-11)a^3b^3 + 2(\kappa-2)(3\kappa+17)a^3b^2c \\
& + 4(\kappa-2)(7\kappa-18)a^3bc^2 + 2(\kappa-3)^2(\kappa-2)a^3c^3 + (23\kappa^2+49\kappa-114)a^2b^4 \\
& + 2(\kappa-2)(\kappa^2+94\kappa-147)a^2b^3c + 6(\kappa-2)(12\kappa^2-34\kappa+17)a^2b^2c^2 \\
& + 2(\kappa-2)(3\kappa^3+9\kappa^2-97\kappa+149)a^2bc^3 + (\kappa-2)(9\kappa^3-68\kappa^2+181\kappa-171)a^2c^4 \\
& + 2(3\kappa^3+73\kappa^2-183\kappa+99)ab^5 + 2(\kappa-2)(96\kappa^2-43\kappa-255)ab^4c \\
& + 4(\kappa-2)(16\kappa^3+94\kappa^2-655\kappa+855)ab^3c^2 \\
& + 2(\kappa-2)(3\kappa^4+159\kappa^3-1233\kappa^2+3164\kappa-2809)ab^2c^3 \\
& + 2(\kappa-2)(39\kappa^4-375\kappa^3+1425\kappa^2-2555\kappa+1825)abc^4 \\
& + 2(\kappa-2)(3\kappa^5-36\kappa^4+181\kappa^3-482\kappa^2+686\kappa-418)ac^5 \\
& + (\kappa^4+50\kappa^3-17\kappa^2-396\kappa+486)b^6 \\
& + 2(\kappa-2)(28\kappa^3+251\kappa^2-1302\kappa+1467)b^5c \\
& + (\kappa-2)(19\kappa^4+745\kappa^3-5374\kappa^2+12664\kappa-10320)b^4c^2 \\
& + 2(\kappa-2)(\kappa^5+224\kappa^4-2062\kappa^3+7371\kappa^2-12357\kappa+8227)b^3c^3 \\
& + (\kappa-2)(129\kappa^5-1464\kappa^4+6952\kappa^3-17464\kappa^2+23397\kappa-13387)b^2c^4 \\
& + 2(\kappa-2)(9\kappa^6-123\kappa^5+727\kappa^4-2405\kappa^3+4754\kappa^2-5374\kappa+2718)bc^5 \\
& + (\kappa-3)(\kappa-2)(\kappa^6-13\kappa^5+74\kappa^4-239\kappa^3+470\kappa^2-544\kappa+292)c^6.
\end{aligned}$$

to the sum, one can verify that this gadget has the succinct binary signature $\langle x, y\rangle$ of type τ_2, where

$$\begin{aligned}
x = {} & f_{\substack{1\,1\\1\,1}} \left[a^2 + (\kappa-1)b^2\right] \\
& + 4(\kappa-1) f_{\substack{1\,2\\1\,1}} \left[ab + b^2 + (\kappa-2)bc\right] \\
& + (\kappa-1) f_{\substack{1\,2\\1\,2}} \left[2ab + (\kappa-2)b^2\right] \\
& + 2(\kappa^2-3\kappa+2) f_{\substack{1\,3\\1\,2}} \left[ac + 2b^2 + (\kappa-3)bc\right] \\
& + (\kappa-1) f_{\substack{1\,2\\2\,1}} \left[2b^2 + (\kappa-2)c^2\right] \\
& + 2(\kappa^2-3\kappa+2) f_{\substack{1\,3\\2\,1}} \left[b^2 + 2bc + (\kappa-3)c^2\right] \\
& + (\kappa-1) f_{\substack{1\,1\\2\,2}} \left[2b^2 + (\kappa-2)c^2\right] \\
& + 2(\kappa^2-3\kappa+2) f_{\substack{1\,1\\2\,3}} \left[b^2 + 2bc + (\kappa-3)c^2\right] \\
& + (\kappa^3-6\kappa^2+11\kappa-6) f_{\substack{1\,4\\2\,3}} \left[4bc + (\kappa-4)c^2\right] \qquad \text{and} \\
y = {} & f_{\substack{1\,1\\1\,1}} \left[2ab + (\kappa-2)b^2\right] \\
& + 4 f_{\substack{1\,2\\1\,1}} \left[ab + (\kappa-2)ac + (2\kappa-3)b^2 + (\kappa-2)^2 bc\right] \\
& + f_{\substack{1\,2\\1\,2}} \left[a^2 + 2(\kappa-2)ab + (\kappa^2-3\kappa+3)b^2\right] \\
& + 2(\kappa-2) f_{\substack{1\,3\\1\,2}} \left[2ab + (\kappa-3)ac + 2(\kappa-2)b^2 + (\kappa^2-4\kappa+5)bc\right] \\
& + f_{\substack{1\,2\\2\,1}} \left[2b^2 + 4(\kappa-2)bc + (\kappa^2-5\kappa+6)c^2\right] \\
& + 2(\kappa-2) f_{\substack{1\,3\\2\,1}} \left[3b^2 + 2(2\kappa-5)bc + (\kappa^2-5\kappa+7)c^2\right] \\
& + f_{\substack{1\,1\\2\,2}} \left[2b^2 + 4(\kappa-2)bc + (\kappa^2-5\kappa+6)c^2\right] \\
& + 2(\kappa-2) f_{\substack{1\,1\\2\,3}} \left[3b^2 + 2(2\kappa-5)bc + (\kappa^2-5\kappa+7)c^2\right] \\
& + (\kappa^2-5\kappa+6) f_{\substack{1\,4\\2\,3}} \left[4b^2 + 4(\kappa-3)bc + (\kappa^2-5\kappa+8)c^2\right].
\end{aligned}$$

Substituting for the entries of f gives the result stated in Lemma 9.2.7. □

Lemma 9.2.8. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c \in \mathbb{C}$. Let $\langle a, b, c\rangle$ be a succinct ternary signature of type τ_3. If $\langle a, b, c\rangle$ is assigned to every vertex of the gadget in Figure 9.8c, then the resulting signature is the binary succinct signature $\langle x, y\rangle$ of type τ_2, where x and y are given in Table 9.1.*



(a) Inner structure (b) Outer structure (c) Entire binary gadget

Figure 9.8: Decomposition of a binary gadget. All circle vertices are assigned $\langle a, b, c\rangle$ and the triangle vertices in (b) is assigned the signature of the gadget in (a).

Proof. Since $\langle a, b, c\rangle$ is domain invariant, the signature of this gadget is also domain invariant. Any domain invariant binary signature has a succinct signature of type τ_2.

Consider the gadget in Figure 9.8a. We assign $\langle a, b, c\rangle$ to all vertices. By Lemma 9.2.4, this gadget has the succinct ternary signature $f = \langle a_0, b_0, c_0\rangle$ of type τ_4, where a_0, b_0, and c_0 are given in the statement of Lemma 9.2.4 as a', b', and c' respectively.

Now consider the gadget in Figure 9.8b. We assign f to the vertices. By Lemma 9.2.2, the resulting gadget has the binary succinct signature $\langle x, y\rangle$ of type τ_2, where

$$x = a_0^2 + 3(\kappa - 1)b_0^2 + (\kappa - 1)(\kappa - 2)c_0^2 \qquad \text{and}$$
$$y = 2a_0b_0 + \kappa b_0^2 + 4(\kappa - 2)b_0c_0 + (\kappa - 2)(\kappa - 3)c_0^2.$$

Substituting for a_0, b_0, and c_0 gives the result in Table 9.1. □

Beyond the gadgets in this section, there are two 9-by-9 recurrence matrices that appear in our proofs in Chapter 11 (see Table 11.1 and Table 11.3). No entry in those recurrence matrices is any harder to compute than any signature entry appearing in this section. The difficulty with these recurrence matrices is the sheer number of terms that must be computed.

Chapter 10

Dichotomy for Counting Edge Colorings over Planar Regular Graphs

In this chapter, we prove that counting edge colorings using κ colors is #P-hard over planar r-regular graphs for all $\kappa \geq r \geq 3$. The problem is polynomial-time computable in all other parameter settings. We reduce from evaluating the Tutte polynomial of a planar graph. If $\kappa = r$, then we reduce from the point $(\kappa+1, \kappa+1)$. If $\kappa > r$, then we reduce from the point $(1-\kappa, 0)$, which is the problem of counting vertex colorings using κ colors. To express our reductions, we introduce notion to succinctly represent signatures with duplicate entries. This work was published in [32, 33].

10.1 Background

An edge κ-coloring of a graph G is a labeling of the edges of G with colors such that any two adjacent edges have different colors. A fundamental problem in graph theory is to determine how many colors are required to edge color a graph G. The obvious lower bound is $\Delta(G)$, the maximum degree of the graph. By Vizing's Theorem [139], an edge coloring using just $\Delta(G)+1$ colors always exists for simple graphs. Whether $\Delta(G)$ colors suffice depends on the graph G.

Consider the edge coloring problem over 3-regular graphs. It follows from the parity condition (Lemma 10.2.3) that any graph containing a bridge does not have an edge 3-coloring. For bridgeless planar simple graphs, Tait [119] showed that the existence of an edge 3-coloring is equivalent to the Four-Color Theorem. Thus, the answer for the decision problem over planar 3-regular simple graphs is that there is an edge 3-coloring iff the graph is bridgeless.

Without the planarity restriction, determining if a simple 3-regular graph has an edge 3-coloring is NP-complete [79]. This hardness extends to finding an edge κ-coloring over simple κ-regular graphs for all $\kappa \geq 3$ [99]. However, these reductions are not parsimonious, and, in fact, it is claimed that no parsimonious reduction exists unless P = NP [141, p. 118]. The counting complexity of this problem has remained open.

We prove that counting edge colorings over planar regular graphs is #P-hard.[1] This solves a long-standing open problem.

Theorem 10.1.1. *#κ-EdgeColoring is #P-hard over planar r-regular graphs for all $\kappa \geq r \geq 3$.*

When this considers fails to hold, the problem is trivially computable in polynomial time. This theorem is proved in Theorem 10.2.7 for $\kappa = r$ and Theorem 10.3.8 for $\kappa > r$. In both cases, we reduce from evaluating the Tutte polynomial of a planar graph. If $\kappa = r$, then we reduce from the point $(\kappa + 1, \kappa + 1)$. If $\kappa > r$, then we reduce from the point $(1 - \kappa, 0)$, which is the problem of counting vertex colorings using κ colors.

As a Holant problem, counting edge κ-colorings over planar r-regular graphs is expressed as Pl-Holant$_\kappa$(All-Distinct$_r$), where All-Distinct$_r$ is a signature of arity r over a domain of size κ that takes value 1 when all inputs are distinct and 0 otherwise. We also denote this signature by AD$_r$. An arity r signature defined over a domain size κ is fully specified by κ^r values. However, some special cases like AD$_r$ can be defined using far fewer values. In addition to being symmetric, it is also invariant under any permutation of the κ domain elements. We call the second property *domain invariance*. The signature of an $\mathcal{F}$-gate in which all signatures in $\mathcal{F}$ are domain invariant is itself domain invariant.

Definition 10.1.2 (Succinct signature)**.** Let $\tau = (P_1, P_2, \ldots, P_\ell)$ be a partition of $[\kappa]^r$ listed in some order. We say that f is a *succinct signature* of type τ if f is constant on each P_i. A set $\mathcal{F}$ of

[1] Vizing's Theorem is for simple graphs. In Holant problems as well as counting complexity such as graph homomorphism or #CSP, one does not typically require the graphs to be simple. However, our hardness result for counting edge 3-colorings over planar 3-regular graphs also holds for simple graphs (Theorem 10.2.8).

signatures is of type τ if every $f \in \mathcal{F}$ has type τ. We denote a succinct signature f of type τ by $\langle f(P_1), \dots, f(P_\ell)\rangle$, where $f(P) = f(x)$ for any $x \in P$.

Furthermore, we may omit 0 entries. If f is a succinct signature of type τ, we also say f is a *succinct signature* of type τ' with length ℓ', where τ' lists ℓ' parts of the partition τ and we write f as $\langle f_1, f_2, \dots, f_{\ell'}\rangle$, provided all nonzero values $f(P_i)$ are listed. When using this notation, we will make it clear which zero entries have been omitted.

For example, a symmetric signature over the Boolean domain (i.e. $\kappa = 2$) is denoted by $[f_0, f_1, \dots, f_r]$, where f_w is the output on inputs of Hamming weight w. This corresponds to the succinct signature type $(P_0, P_1, \dots, P_r)$, where P_w is the set of inputs of Hamming weight w. A similar succinct signature notation was used for symmetric signatures on domain size 3 [51, p. 1282]. Other uses of this definition have been implicit (cf. the comments just before Lemma 6.2.2 and Lemma 6.3.1).

We use several succinct signature types in our proof of Theorem 10.1.1. For domain invariant unary signatures, there are only two signatures up to a nonzero scalar. Using the trivial partition that contains all inputs, we denote these two succinct unary signatures as $\langle 0\rangle$ and $\langle 1\rangle$ and say that they have succinct type τ_1.

Domain invariant binary signatures are necessarily symmetric. We call their succinct signature type $\tau_2 = (P_1, P_2)$, where $P_i = \{(x, y) \in [\kappa]^2 : |\{x, y\}| = i\}$ for $1 \le i \le 2$.

Domain invariant ternary signatures may not be symmetric, but for those that are, then have a succinct signature of type τ_3 defined as follows. The succinct signature type $\tau_3 = (P_1, P_2, P_3)$ is a partition of $[\kappa]^3$ with $P_i = \{(x, y, z) \in [\kappa]^3 : |\{x, y, z\}| = i\}$ for $1 \le i \le 3$. The notation $\{x, y, z\}$ denotes a multiset and $|\{x, y, z\}|$ denotes the number of distinct elements in it. In particular, the succinct ternary signature for AD_3 is $\langle 0, 0, 1\rangle$.

We note that the number of succinct signature types for signatures of arity r over a domain of size κ that are both symmetric and domain invariant is the number of partitions of r into at most κ parts. This is related to the partition function from number theory, which is not to be confused with the partition function with its origins in statistical mechanics and has been intensively studied in complexity theory of counting problems.

While there are some other succinct signature types that we define later as needed, there is one more important type that we define here. Any quaternary signature f that is domain invariant has a succinct signature of length at most 15. When a signature has both vertical and horizontal symmetry, there is a shorter succinct signature that has only length 9. We say a signature f has vertical symmetry if $f(w,x,y,z) = f(x,w,z,y)$ and horizontal symmetry if $f(w,x,y,z) = f(z,y,x,w)$. For example, the signature of the gadget in Figure 10.6 has both vertical and horizontal symmetry. Accordingly, let $\tau_4 = (P_{\substack{1\,1\\1\,1}}, P_{\substack{1\,2\\1\,1}}, P_{\substack{1\,2\\1\,2}}, P_{\substack{1\,3\\1\,2}}, P_{\substack{1\,2\\2\,1}}, P_{\substack{1\,3\\2\,1}}, P_{\substack{1\,1\\2\,2}}, P_{\substack{1\,1\\2\,3}}, P_{\substack{1\,4\\2\,3}})$ be a type of succinct quaternary signature with partitions

$$
\begin{aligned}
P_{\substack{1\,1\\1\,1}} &= \{(w,x,y,z) \in [\kappa]^4 \mid w = x = y = z\},\\
P_{\substack{1\,2\\1\,1}} &= \left\{ (w,x,y,z) \in [\kappa]^4 \;\middle|\; \begin{array}{c} (w = x = y \neq z) \vee (w = x = z \neq y) \\ \vee\, (w = y = z \neq x) \vee (x = y = z \neq w) \end{array} \right\},\\
P_{\substack{1\,2\\1\,2}} &= \{(w,x,y,z) \in [\kappa]^4 \mid w = x \neq y = z\},\\
P_{\substack{1\,3\\1\,2}} &= \{(w,x,y,z) \in [\kappa]^4 \mid (w = x \neq y \neq z \neq x) \vee (y = z \neq w \neq x \neq z)\},\\
P_{\substack{1\,2\\2\,1}} &= \{(w,x,y,z) \in [\kappa]^4 \mid w = y \neq x = z\},\\
P_{\substack{1\,3\\2\,1}} &= \{(w,x,y,z) \in [\kappa]^4 \mid (w = y \neq x \neq z \neq y) \vee (x = z \neq w \neq y \neq z)\},\\
P_{\substack{1\,1\\2\,2}} &= \{(w,x,y,z) \in [\kappa]^4 \mid w = z \neq x = y\},\\
P_{\substack{1\,1\\2\,3}} &= \{(w,x,y,z) \in [\kappa]^4 \mid (w = z \neq x \neq y \neq z) \vee (x = y \neq w \neq z \neq y)\}, \text{ and}\\
P_{\substack{1\,4\\2\,3}} &= \{(w,x,y,z) \in [\kappa]^4 \mid w,x,y,z \text{ are all distinct}\}.
\end{aligned}
$$

Lemma 6.2.4 is about interpolating unary signatures over the Boolean domain for planar Holant problems, but the same proof applies equally well for higher domains using a binary recursive construction (like that in Figure 10.4) and a succinct signature type with length 2.

Lemma 10.1.3 (Rephrasing of Lemma 6.2.4)**.** *Suppose $\mathcal{F}$ is a set of signatures and τ is a succinct signature type with length 2. If there exists an infinite sequence of planar $\mathcal{F}$-gates defined by an initial succinct signature $s \in \mathbb{C}^{2 \times 1}$ of type τ and recurrence matrix $M \in \mathbb{C}^{2 \times 2}$ satisfying the following conditions,*

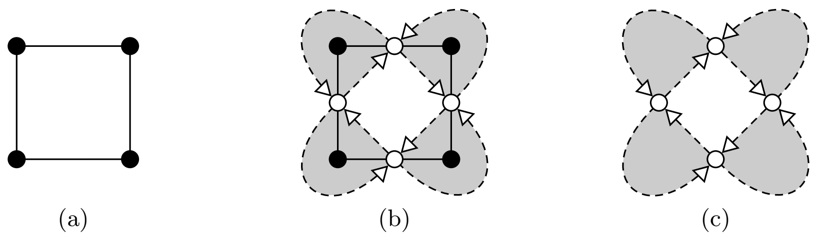


Figure 10.1: Plane graph (a), its directed medial graph (c), and both superimposed (b).

1. $\det(M) \neq 0$*;*

2. $\det([s\ Ms]) \neq 0$*;*

3. M *has infinite order modulo a scalar;*

then

$$\text{Pl-Holant}(\mathcal{F} \cup \{\langle x, y\rangle\}) \leq_T \text{Pl-Holant}(\mathcal{F}),$$

for any $x, y \in \mathbb{C}$*, where* $\langle x, y\rangle$ *is a succinct binary signature of type* τ*.*

10.2 Number of Colors Equals the Regularity Parameter

In this section, we consider $\kappa = r \geq 3$. We reduce from evaluating the Tutte polynomial of a planar graph at the point $(\kappa + 1, \kappa + 1)$. This problem is #P-hard by Theorem 6.1.4.

To express our reduction from these points on the Tutte polynomial, we need to consider Eulerian subgraphs in directed medial graphs. The definition of an undirected medial graph was given in Definition 6.4.2. We extend this definition to directed graphs.

Definition 10.2.1 (cf. Section 4 in [64])**.** The *directed medial graph* $\vec{G}_m$ of G colors the faces of G_m black or white depending on whether they contain or do not contain, respectively, a vertex of G. Then the edges of the medial graph are directed so that the black face is on the left.

Figure 6.4 and Figure 10.1 give examples of a medial graph and a directed medial graph respectively. A (directed) medial graph is always a planar 4-regular graph. We say a graph is an Eulerian (di)graph if every vertex has even degree (resp. in-degree equal to out-degree), but connectedness is not required. Thus, a (directed) medial graph is Eulerian.

Building on previous work [105, 135, 63, 1], Ellis-Monaghan gave the following connection with the diagonal of the Tutte polynomial. A monochromatic vertex is a vertex with all its incident edges having the same color.

Lemma 10.2.2 (Equation (17) in [64])**.** *Suppose G is a connected plane graph and $\vec{G}_m$ is its directed medial graph. For $\kappa \in \mathbb{N}$, let $\mathcal{C}(\vec{G}_m)$ be the set of all edge κ-labelings of $\vec{G}_m$ so that each (possibly empty) set of monochromatic edges forms an Eulerian digraph. Then*

$$\kappa\, \mathrm{T}(G; \kappa+1, \kappa+1) = \sum_{c\, \in\, \mathcal{C}(\vec{G}_m)} 2^{m(c)}, \tag{10.2.1}$$

where $m(c)$ is the number of monochromatic vertices in the coloring c.

The Eulerian partitions in $\mathcal{C}(\vec{G}_m)$ have the property that the subgraphs induced by each partition do not intersect (or crossover) each other due to the orientation of the edges in the medial graph. We call the counting problem defined by the sum on the right-hand side of (10.2.1) counting weighted Eulerian partitions over planar 4-regular graphs. This problem also has an expression as a Holant problem using a succinct signature. To define this succinct signature, it helps to know the following basic result about edge colorings.

When the number of available colors coincides with the regularity parameter of the graph, the cuts in any coloring satisfy a well-known parity condition. This parity condition follows from a more general parity argument (see (1.2) and the Parity Argument on page 95 in [117]). We state this simpler parity condition and provide a short proof for completeness.

Lemma 10.2.3 (Parity Condition)**.** *Let G be a κ-regular graph and consider a cut C in G. For any edge κ-coloring of G,*

$$c_1 \equiv c_2 \equiv \cdots \equiv c_\kappa \pmod 2,$$

where c_i is the number of edges in C colored i for $1 \le i \le \kappa$.

Proof. Consider two distinct colors i and j. Remove from G all edges not colored i or j. The resulting graph is a disjoint union of cycles consisting of alternating colors i and j. Each cycle in this graph must cross the cut C an even number of times. Therefore, $c_i \equiv c_j \pmod 2$. □

Consider all quaternary $\{\mathrm{AD}_\kappa\}$-gates on domain size $\kappa \geq 3$. These gadgets have a succinct signature of type $\tau_{\text{color}} = (P_{\begin{smallmatrix}1&1\\1&1\end{smallmatrix}}, P_{\begin{smallmatrix}1&2\\1&2\end{smallmatrix}}, P_{\begin{smallmatrix}1&2\\2&1\end{smallmatrix}}, P_{\begin{smallmatrix}1&1\\2&2\end{smallmatrix}}, P_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}}, P_0)$, where

$$\begin{aligned}
P_{\begin{smallmatrix}1&1\\1&1\end{smallmatrix}} &= \{(w,x,y,z) \in [\kappa]^4 \mid w = x = y = z\},\\
P_{\begin{smallmatrix}1&2\\1&2\end{smallmatrix}} &= \{(w,x,y,z) \in [\kappa]^4 \mid w = x \neq y = z\},\\
P_{\begin{smallmatrix}1&2\\2&1\end{smallmatrix}} &= \{(w,x,y,z) \in [\kappa]^4 \mid w = y \neq x = z\},\\
P_{\begin{smallmatrix}1&1\\2&2\end{smallmatrix}} &= \{(w,x,y,z) \in [\kappa]^4 \mid w = z \neq x = y\},\\
P_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}} &= \{(w,x,y,z) \in [\kappa]^4 \mid w,x,y,z \text{ are distinct}\}, \text{ and}\\
P_0 &= [\kappa]^4 - P_{\begin{smallmatrix}1&1\\1&1\end{smallmatrix}} - P_{\begin{smallmatrix}1&2\\1&2\end{smallmatrix}} - P_{\begin{smallmatrix}1&2\\2&1\end{smallmatrix}} - P_{\begin{smallmatrix}1&1\\2&2\end{smallmatrix}} - P_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}}.
\end{aligned}$$

Any quaternary signature of an $\{\mathrm{AD}_\kappa\}$-gate is constant on the first five parts of τ_{color} since AD_κ is domain invariant. Using Lemma 10.2.3, we can show that the entry corresponding to P_0 is 0.

Lemma 10.2.4. *Suppose $\kappa \geq 3$ is the domain size and let F be a quaternary $\{\mathrm{AD}_\kappa\}$-gate with succinct signature f of type τ_{color}. Then $f(P_0) = 0$.*

Proof. Let $\sigma_0 \in P_0$ be an edge κ-labeling of the external edges of F. Assume for a contradiction that σ_0 can be extended to an edge κ-coloring of F. We form a graph G from two copies of F (namely, F_1 and F_2) by connecting their corresponding external edges. Then G has a coloring σ that extends σ_0. Consider the cut $C = (F_1, F_2)$ in G whose cut set contains exactly those edges labeled by σ_0. By Lemma 10.2.3, the counts of the colors assigned by σ_0 must satisfy the parity condition. However, this is a contradiction since no edge κ-labeling in P_0 satisfies the parity condition. □

By Lemma 10.2.4, we denote a quaternary signature f of an $\{\mathrm{AD}_\kappa\}$-gate by the succinct signature $\langle f(P_{\begin{smallmatrix}1&1\\1&1\end{smallmatrix}}), f(P_{\begin{smallmatrix}1&2\\1&2\end{smallmatrix}}), f(P_{\begin{smallmatrix}1&2\\2&1\end{smallmatrix}}), f(P_{\begin{smallmatrix}1&1\\2&2\end{smallmatrix}}), f(P_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}})\rangle$ of type τ_{color}, which has the entry for P_0 omitted.[2] When $\kappa = 3$, $P_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}}$ is empty and we define its entry in the succinct signature to be 0.

[2] If $\kappa > 4$, then Lemma 10.2.3 further implies that these signatures are also 0 on $P_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}}$. However, when $\kappa = 4$, this value might be nonzero. The AD_4 signature is an example of this. Instead of using this observation that depends on κ in our proof, we only construct gadgets such that their signatures are 0 on $P_{\begin{smallmatrix}1&4\\2&3\end{smallmatrix}}$ for any value of κ.

Lemma 10.2.5. *Let G be a connected plane graph and let G_m be the medial graph of G. Then*

$$\kappa\,\mathrm{T}(G;\kappa+1,\kappa+1) = \text{Pl-Holant}_\kappa(G_m;\langle 2,1,0,1,0\rangle),$$

where the Holant problem has domain size κ and $\langle 2,1,0,1,0\rangle$ is a succinct signature of type τ_{color}.

Proof. Let $f = \langle 2,1,0,1,0\rangle$. By Lemma 10.2.2, we only need to prove that

$$\sum_{c \,\in\, \mathcal{C}(\vec{G}_m)} 2^{m(c)} = \text{Pl-Holant}_\kappa(G_m; f), \tag{10.2.2}$$

where the notation is from Lemma 10.2.2.

Each $c \in \mathcal{C}(\vec{G}_m)$ is also an edge κ-labeling of G_m. At each vertex $v \in V(\vec{G}_m)$, the four incident edges are assigned at most two distinct colors by c. If all four edges are assigned the same color, then the constraint f on v contributes a factor of 2 to the total weight. This is given by the value in the first entry of f. Otherwise, there are two different colors, say x and y. Because the orientation at v in $\vec{G}_m$ is cyclically "in, out, in, out", the coloring around v can only be of the form $xxyy$ or $xyyx$. These correspond to the second and fourth entries of f. Therefore, every term in the summation on the left-hand side of (10.2.2) appears (with the same nonzero weight) in the summation $\text{Holant}_\kappa(G_m; f)$.

It is also easy to see that every nonzero term in $\text{Holant}_\kappa(G_m; f)$ appears in the sum on the left-hand side of (10.2.2) with the same weight of 2 to the power of the number of monochromatic vertices. In particular, any coloring with a vertex that is cyclically colored $xyxy$ for different colors x and y does not contribute because $f(P_{\substack{1\,2\\2\,1}}) = 0$. □

Remark. This result shows that this planar Holant problem is at least as hard as computing the Tutte polynomial at the point $(\kappa+1, \kappa+1)$ over planar graphs, which implies #P-hardness. Of course they are equally difficult in the sense that both are #P-complete. In fact, they are more directly related since every 4-regular plane graph is the medial graph of some plane graph.

By Theorem 6.1.4 and Lemma 10.2.5, the problem $\text{Pl-Holant}_\kappa(\langle 2,1,0,1,0\rangle)$ is #P-hard. We state this as a corollary.

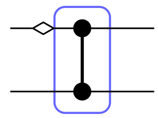


Figure 10.2: Quaternary gadget used in the interpolation construction below. All vertices are assigned AD_κ. The bold edge represents $\kappa - 2$ parallel edges.

Corollary 10.2.6. *Suppose $\kappa \geq 3$ is the domain size. Let $\langle 2, 1, 0, 1, 0 \rangle$ be a succinct quaternary signature of type τ_{color}. Then $\text{Pl-Holant}_\kappa(\langle 2, 1, 0, 1, 0 \rangle)$ is #P-hard.*

With this connection established, we can now show that counting edge colorings is #P-hard over planar regular graphs when the number of colors and the regularity parameter coincide.

Theorem 10.2.7. *#κ-EdgeColoring is #P-hard over planar κ-regular graphs for all $\kappa \geq 3$.*

Proof. Let $\langle 2, 1, 0, 1, 0 \rangle$ be a succinct quaternary signature of type τ_{color}. We reduce from the problem $\text{Pl-Holant}_\kappa(\langle 2, 1, 0, 1, 0 \rangle)$, which is #P-had by Corollary 10.2.6. We reduce to $\text{Pl-Holant}_\kappa(\mathrm{AD}_\kappa)$, which expresses counting edge κ-colorings over planar κ-regular graphs as a Holant problem.

Consider the gadget in Figure 10.2, where the bold edge represents $\kappa - 2$ parallel edges. We assign AD_κ to both vertices. Up to a nonzero factor of $(\kappa - 2)!$, this gadget has the succinct quaternary signature $f = \langle 0, 1, 1, 0, 0 \rangle$ of type τ_{color}. Now consider the recursive construction in Figure 10.3. All vertices are assigned the signature f. Let f_s be the succinct quaternary signature

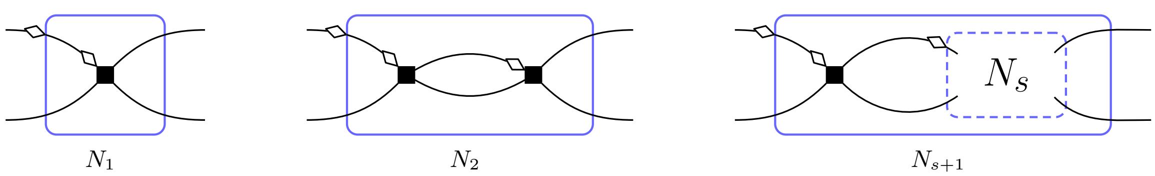


Figure 10.3: Recursive construction to interpolate $\langle 2, 1, 0, 1, 0 \rangle$. The vertices are assigned the signature of the gadget in Figure 10.2.

of type τ_{color} for the sth gadget of the recursive construction. Then $f_1 = f$ and $f_s = M^s f_0$, where

$$M = \begin{bmatrix} 0 & \kappa-1 & 0 & 0 & 0 \\ 1 & \kappa-2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad f_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

The signature f_0 is simply the succinct quaternary signature of type τ_{color} for two parallel edges. We can express M via the Jordan decomposition $M = P \Lambda P^{-1}$, where

$$P = \begin{bmatrix} 1 & 1-\kappa & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and $\Lambda = \text{diag}(\kappa - 1, -1, 1, -1, 1)$. Then for $t = 2s$, we have

$$f_t = P \begin{bmatrix} \kappa-1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}^t P^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = P \begin{bmatrix} x & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} P^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} y+1 \\ y \\ 0 \\ 1 \\ 0 \end{bmatrix},$$

where $x = (\kappa - 1)^t$ and $y = \frac{x-1}{\kappa}$.

Consider an instance Ω of Pl-Holant$_\kappa(\langle 2, 1, 0, 1, 0 \rangle)$. Suppose $\langle 2, 1, 0, 1, 0 \rangle$ appears n times in Ω. We construct from Ω a sequence of instances Ω_t of Pl-Holant$_\kappa(f)$ indexed by t, where $t = 2s$ with $s \geq 0$. We obtain Ω_t from Ω by replacing each occurrence of $\langle 2, 1, 0, 1, 0 \rangle$ with the gadget f_t.

As a polynomial in $x = (\kappa - 1)^t$, Pl-Holant$_\kappa(\Omega_t; f)$ is independent of t and has degree at most n with integer coefficients. Using our oracle for Pl-Holant$_\kappa(f)$, we can evaluate this polynomial at $n+1$ distinct points $x = (\kappa - 1)^{2s}$ for $0 \leq s \leq n$. Then via polynomial interpolation, we can recover the coefficients of this polynomial efficiently. Evaluating this polynomial at $x = \kappa + 1$ (so that $y = 1$) gives the value of Pl-Holant$_\kappa(\Omega; \langle 2, 1, 0, 1, 0 \rangle)$, as desired. □

Remark. For $\kappa = 3$, the interpolation step is actually unnecessary since the succinct signature of f_2 happens to be $\langle 2, 1, 0, 1, 0 \rangle$.

When $\kappa = 3$, it is easy to extend Theorem 10.2.7 by further restricting to simple graphs.

Theorem 10.2.8. *#3-*EdgeColoring *is #*P*-hard over simple planar* 3*-regular graphs.*

Proof. By Theorem 10.2.7, it suffices to efficiently compute the number of edge 3-colorings of a planar 3-regular graph G that might have self-loops and parallel edges. Furthermore, we can assume that G is connected since the number of edge colorings is multiplicative over connected components. If G contains a self-loop, then there are no edge colorings in G, so assume G contains no self-loops. If G also contains no parallel edges, then G is simple and we are done.

Thus, assume that G contains n vertices and parallel edges between some distinct vertices u and v. If u and v are connected by three edges, then this constitutes the whole graph, which has six edge 3-colorings. Otherwise, u and v are connected by two edges. Then there exist vertices u' and v' such that u and u' are connected by a single edge, v and v' are connected by a single edge, and $u' \neq v'$. In any edge 3-coloring of G, it is easy to see that the edges (u, u') and (v, v') must be assigned the same color. By removing u, v, and their incident edges while adding an edge between u' and v', we have a planar 3-regular graph G' on $n - 2$ vertices with half as many edge colorings as G. Then by induction, we can efficiently compute the number of edge 3-colorings in G'. □

Section 11.4, we give an alternative proof of Theorem 10.2.7, which uses the interpolation techniques we develop in Section 11.3. The purpose of Section 11.4 is to show how a recursive construction in an interpolation proof can be used to form a hypothesis about possible invariance properties. One example of an invariance property is that any planar $\{\mathrm{AD}_\kappa\}$-gate with a succinct quaternary signature $\langle a, b, c, d, e \rangle$ of type τ_{color} must satisfy $a + c = b + d$ (Lemma 11.4.1).

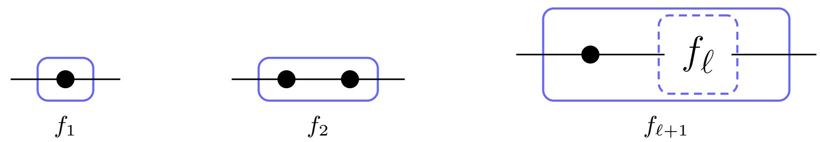


Figure 10.4: Recursive construction to interpolate any succinct binary signature of type τ_2. All vertices are assigned the same succinct binary signature of type τ_2.

10.3 Number of Colors Exceeds the Regularity Parameter

Now we consider $\kappa > r \geq 3$. we reduce from the problem of counting vertex κ-colorings over planar graphs. This problem is also #P-hard by the same dichotomy for the Tutte polynomial (Theorem 6.1.4) since the chromatic polynomial is a specialization the Tutte polynomial.

Proposition 10.3.1 (Proposition 6.3.1 in [12])**.** *Let $G = (V, E)$ be a graph. Then $\chi(G; \lambda)$, the chromatic polynomial of G, is expressed as a specialization of the Tutte polynomial via the relation*

$$\chi(G; \lambda) = (-1)^{|V|-k(G)} \lambda^{k(G)}\, \mathrm{T}(G; 1 - \lambda, 0),$$

where $k(G)$ is the number of connected components of the graph G.

The first step in the proof is to interpolate every possible binary signature that is domain invariant, which are necessarily symmetric. These signatures have the succinct signature type τ_2.

Lemma 10.3.2. *Suppose $\kappa \geq 3$ is the domain size and let $x, y \in \mathbb{C}$. If we assign the succinct binary signature $\langle x, y \rangle$ of type τ_2 to every vertex of the recursive construction in Figure 10.4, then the corresponding recurrence matrix is $\left[\begin{smallmatrix} x & (\kappa-1)y \\ y & x+(\kappa-2)y \end{smallmatrix}\right]$ with eigenvalues $x + (\kappa - 1)y$ and $x - y$.*

Proof. Let f_ℓ be the signature of the ℓth gadget in this construction. To obtain $f_{\ell+1}$ from f_ℓ, we view f_ℓ as a column vector and multiply it by the recurrence matrix $M = \left[\begin{smallmatrix} x & (\kappa-1)y \\ y & x+(\kappa-2)y \end{smallmatrix}\right]$. In general, we have $f_\ell = M^\ell f_0$, where f_0 is the initial signature, which is just a single edge and has the succinct binary signature $\langle 1, 0 \rangle$ of type τ_2. The (column) eigenvectors of M are $\left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$ and $\left[\begin{smallmatrix} 1-\kappa \\ 1 \end{smallmatrix}\right]$ with eigenvalues $x + (\kappa - 1)y$ and $x - y$ respectively, as claimed. □

Consider the recursive construction in Figure 10.4. To every vertex, we assign the succinct binary signature $\langle x, y \rangle$. Since the initial signature is $s = \langle 1, 0 \rangle$, the determinant of the matrix

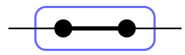


Figure 10.5: Binary gadget used in the interpolation construction of Figure 10.4. Both vertices are assigned AD_r and the bold edge represents $r-1$ parallel edges.

$[s\ Ms]$ is simply y. In order to interpolate all binary succinct signatures of type τ_2, we need to satisfy the second condition of Lemma 10.1.3, which is $y \neq 0$. However when $y = 0$, the recurrence matrix is a scalar multiple of the identity matrix, which implies that the eigenvalues are the same. For two dimensional interpolation using a matrix with a full basis of eigenvectors, as is the case here, the third condition of Lemma 10.1.3 is equivalent to the ratio of the eigenvalues not being a root of unity. In particular, the eigenvalues cannot be the same. Therefore, when using the recursive construction in Figure 10.4, it suffices to satisfy the first and third conditions of Lemma 10.1.3. We state this as a corollary.

Corollary 10.3.3. *Suppose $\mathcal{F}$ is a set of signatures. Let $s = \langle 1, 0\rangle$ of type τ_2 be the initial succinct binary signature and let $M \in \mathbb{C}^{2\times 2}$ be the recurrence matrix for some infinite sequence of planar $\mathcal{F}$-gates defined by the recursive construction in Figure 10.4. If M satisfies the following conditions,*

1. $\det(M) \neq 0$;
2. *M has infinite order modulo a scalar;*

then

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle x, y\rangle\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F}),$$

for any $x, y \in \mathbb{C}$, where $\langle x, y\rangle$ is a succinct binary signature of type τ_2.

Lemma 10.3.4. *Suppose κ is the domain size with $\kappa > r$ for any integer $r \geq 3$, and $x, y \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing AD_r. Then*

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle x, y\rangle\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F}),$$

where $\langle x, y\rangle$ is a succinct binary signature of type τ_2.

Proof. Let $(n)_k = n(n-1)\cdots(n-k+1)$ be the kth falling power of n. Consider the gadget in Figure 10.5. We assign AD_r to both vertices. The succinct binary signature of type τ_2 for this gadget is $f = \langle (\kappa-1)_{r-1}, (\kappa-2)_{r-1} \rangle$. Up to a nonzero factor of $(\kappa-2)_{r-2}$, we have the signature $f' = \frac{1}{(\kappa-2)_{r-2}} f = \langle \kappa - 1, \kappa - r \rangle$.

Consider the recursive construction in Figure 10.4. We assign f' to all vertices. Then by Lemma 10.3.2, the two eigenvalues of the corresponding recurrence matrix are $(r-1) > 0$ and $(\kappa-1)(\kappa-r+1) > 0$. Thus, M is nonsingular. Furthermore, the eigenvalues are not equal since $\kappa \notin \{0, r\}$. Therefore, we are done by Corollary 10.3.3. □

Next we show that Pl-Holant$_\kappa$(AD$_r$) is at least as hard as Pl-Holant$_\kappa$(AD$_3$). To overcome a difficulty when r is even, we use the the concept of a planar pairing, which was introduced in Section 6.3.

Lemma 10.3.5. *Suppose κ is the domain size with $\kappa > r$ for any integer $r \geq 3$. Then*

$$\text{Pl-Holant}_\kappa(\text{AD}_3) \leq_T \text{Pl-Holant}_\kappa(\text{AD}_r).$$

Proof. By Lemma 10.3.4, we can assume that $\langle 1, 1 \rangle$ is available. Take AD$_r$ and first form $t = \lceil \frac{r-4}{2} \rceil$ self-loops. Then add a new vertex on each self-loop and assign $\langle 1, 1 \rangle$ to each of these new vertices. Up to a nonzero factor of $(\kappa-3)_{2t}$, the resulting signature is AD$_3$ if r is odd and AD$_4$ if r is even. To reduce from $r = 3$ to $r = 4$, we use a planar pairing, which can be efficiently computed by Lemma 6.3.4. We add a new vertex on each edge in a planar pairing and assign $\langle 1, 1 \rangle$ to each of these new vertices. Then up to a nonzero factor of $\kappa - 3$, the signature at each vertex of the initial graph is effectively AD$_3$. □

The succinct binary signature $\langle 1-\kappa, 1 \rangle$ of type τ_2 has a special property. Let u be any constant unary signature, which has a succinct signature of type τ_1. If $\langle 1-\kappa, 1 \rangle$ is connected to u, then the resulting unary signature is identically zero.

This observation is the key to what follows. We use it in the next lemma to achieve what would appear to be an impossible task. The requirements, if duly specified, would result in multiple conditions to be satisfied by nine separate polynomials pertaining to some construction in place of the gadget in Figure 10.6. And yet we are able to use just one degree of freedom to cause seven of the polynomials to vanish, satisfying most of these conditions. In addition, the other two

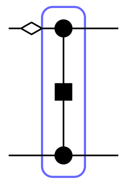


Figure 10.6: The Bobby Fischer gadget, which achieves many objectives using only a single degree of freedom.

polynomials are not forgotten. They are nonzero and their ratio is not a root of unity, which allows interpolation to succeed.

This ability to satisfy a multitude of constraints simultaneously in one magic stroke reminds us of some unfathomably brilliant moves by Bobby Fischer, the chess genius extraordinaire, and so we name this gadget (Figure 10.6) the *Bobby Fischer gadget*.

This gadget is the new idea that allows us to prove Theorem 10.3.8.

Lemma 10.3.6. *Suppose $\kappa \geq 3$ is the domain size and $a, b \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle a, b, b\rangle$ of type τ_3 and the succinct binary signature $\langle 1-\kappa, 1\rangle$ of type τ_2. If $a \neq b$, then*

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{=_4\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F}).$$

Proof. Consider the gadget in Figure 10.6. We assign $\langle a, b, b\rangle$ to the circle vertices and $\langle 1-\kappa, 1\rangle$ to the square vertex. This gadget has a succinct quaternary signature of type τ_4, which has length 9. However, all but two of the entries in this succinct signature must be 0.

To see this, consider an assignment that assigns different values to the two edges incident to the circle vertex on top. Since the assignment to these two edges differ, the signature $\langle a, b, b\rangle$ contributes a factor of b regardless of the value of its third edge, which is connected to the square vertex assigned $\langle 1-\kappa, 1\rangle$. From the perspective of $\langle 1-\kappa, 1\rangle$, this behavior is equivalent to connecting it to the succinct unary signature $b\langle 1\rangle$ of type τ_1. Thus, the sum over the possible assignments to this third edge is 0. The same argument shows that the two edges incident to the circle vertex on the bottom do not contribute anything to the Holant sum when assigned different values.

Thus, any nonzero contribution to the Holant sum comes from assignments where the top two dangling edges are assigned the same value and the bottom two dangling edges are assigned the same value. There are only two entries that satisfy this requirement in the succinct quaternary signature of type τ_4 for this gadget, which are the entries for $P_{\substack{1\,1\\1\,1}}$ and $P_{\substack{1\,1\\2\,2}}$. To compute those two entries, we use the following trick. Since the two external edges of each circle vertex must be assigned the same value, we think of them as just a single edge. Then the effective succinct binary signature of type τ_2 for the circle vertices is just $\langle a, b\rangle$. By connecting the first $\langle a, b\rangle$ with $\langle 1-\kappa, 1\rangle$, the result is $(a-b)\langle 1-\kappa, 1\rangle$ like it is an eigenvector. Connecting the other copy of $\langle a, b\rangle$ to $(a-b)\langle 1-\kappa, 1\rangle$ gives $(a-b)^2\langle 1-\kappa, 1\rangle$. This computation is expressed via the matrix multiplication $[bJ_\kappa + (a-b)I_\kappa][J_\kappa - \kappa I_\kappa][bJ_\kappa + (a-b)I_\kappa] = (a-b)[J_\kappa - \kappa I_\kappa][bJ_\kappa + (a-b)I_\kappa] = (a-b)^2[J_\kappa - \kappa I_\kappa]$. Thus up to a nonzero factor of $(a-b)^2$, the corresponding succinct quaternary signature of type τ_4 for this gadget is $f = \langle 1-\kappa, 0, 0, 0, 0, 0, 1, 0, 0\rangle$.

Consider the recursive construction in Figure 10.3. We assign f to all vertices. Let f_s be the signature of the sth gadget in this construction. The seven entries that are 0 in the succinct signature of type τ_4 for f are also 0 in the succinct signature of type τ_4 for f_s. Thus, we can express f_s via a succinct signature of type τ_4' with length 2, defined as follows. The first two parts in τ_4' are $P_{\substack{1\,1\\1\,1}}$ and $P_{\substack{1\,1\\2\,2}}$ from the succinct signature type τ_4. The last part contains all the remaining assignments. Then the succinct signature for f_s of type τ_4' is $M^s f_0$, where $M = \left[\begin{smallmatrix} 1-\kappa & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and $f_0 = \langle 1, 1\rangle$, which is just the succinct signature of type τ_4' for two parallel edges.

Clearly the conditions in Lemma 10.1.3 hold, so we can interpolate any succinct signature of type τ_4'. In particular, we can interpolate our target signature $=_4$, which is $\langle 1, 0\rangle$ when expressed as a succinct signature of type τ_4'. □

Remark. The nine polynomials mentioned before Lemma 10.3.6 correspond to the nine entries of some quaternary gadget with a succinct signature of type τ_4. In light of Lemma 10.3.4, this gadget might involve many succinct binary signatures $\langle x, y\rangle$ of type τ_2 for various choices of $x, y \in \mathbb{C}$. Each distinct binary signature provides an additional degree of freedom to these polynomials. Our construction in Figure 10.6 only requires one binary signature $\langle x, y\rangle$ and we use our one degree of freedom to set $\frac{x}{y} = 1 - \kappa$.

With the aid of the succinct unary signature $\langle 1\rangle$ of type τ_1 and the succinct binary signature $\langle 0, 1\rangle$ of type τ_2, the assumptions in the previous lemma are sufficient to prove #P-hardness.

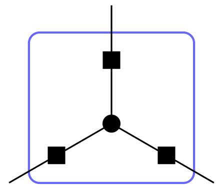


Figure 10.7: Local holographic transformation gadget construction for a ternary signature.

Corollary 10.3.7. *Suppose $\kappa \geq 3$ is the domain size and $a, b \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle a, b, b \rangle$ of type τ_3, the succinct unary signature $\langle 1 \rangle$ of type τ_1, and the succinct binary signatures $\langle 1-\kappa, 1 \rangle$ and $\langle 0, 1 \rangle$ of type τ_2. If $a \neq b$, then* $\operatorname{Pl-Holant}_\kappa(\mathcal{F})$ *is* #P*-hard.*

Proof. By Lemma 10.3.6, we have $=_4$. Connecting $\langle 1 \rangle$ to $=_4$ gives $=_3$. With $=_3$, we can construct the equality signatures of every arity. Along with the binary disequality signature $\neq_2$, which is the succinct binary signature $\langle 0, 1 \rangle$ of type τ_2, we can express the problem of counting the number of vertex κ-colorings over planar graphs. By Proposition 10.3.1, this is, up to a nonzero factor, the problem of evaluating the Tutte polynomial at $(1-\kappa, 0)$, which is #P-hard by Theorem 6.1.4. □

Now we can show that counting edge colorings is #P-hard over planar regular graphs when there are more colors than the regularity parameter.

Theorem 10.3.8. *#κ-*EdgeColoring *is* #P*-hard over planar r-regular graphs for all $\kappa > r \geq 3$.*

Proof. By Lemma 10.3.5, it suffices to consider $r = 3$. By Lemma 10.3.4, we can assume that any succinct binary signature of type τ_2 is available.

Consider the gadget in Figure 10.7. We assign AD_3 to the circle vertex and $\langle 3-\kappa, 1 \rangle$ to the square vertices. By Lemma 9.2.6, the succinct ternary signature of type τ_3 for this gadget is $f = 2(\kappa-2)\langle -(\kappa-3)(\kappa-1), 1, 1 \rangle$.

Now take two edges of AD_3 and connect them to the two edges of $\langle 1, 1 \rangle$. Up to a nonzero factor of $(\kappa-1)(\kappa-2)$, this gadget has the succinct unary signature $\langle 1 \rangle$ of type τ_1. Then we are done by Corollary 10.3.7. □

10.4 Closing Thoughts

Genesis I think the story about how this work began is fascinating. Let me share it with you.

After publishing [73] in December 2012, I was using my free time to work on a draft Wikipedia page about Holant that I had created the previous December. While adding example problems, I came to the section about graph colorings. Edge coloring is the most natural higher domain Holant problem, and I quickly added it. For each example, I would also add a reference that studied the problem and determined its counting complexity. But for this problem, I was at a loss. None of the papers that I knew about or could find had done this. I emailed Jin-Yi Cai on the last day of the year, and we began an informal discussion about the problem.

January of 2013 was a month that I had been looking forward to for some time. From the 13th to the 18th, Heng Guo and I were going to attend Dagstuhl Seminar 13031 about Computational Counting. We arrived a few days early and visited Mingji Xia at the Max Planck Institute for Informatics. We discussed the problem of edge coloring and came up with some ideas. With help from Jin-Yi Cai over email, we had (on January 14th) a sketch of a proof that counting edge κ-colorings was #P-hard over planar κ-regular graphs for $\kappa \geq 3$. While at the seminar, we didn't have the time to check all the details.

In the meantime, no less than three people approached me and asked if I knew the complexity of counting edge colorings. I was stunned. Normally I wouldn't share a result that I wasn't sure about, but I made an exception for this special event. When we returned home, we realized that planarity imposes a restriction that did not arise in our proof sketch.[3] With this restriction, we did not know of any #P-hard problem from which to reduce. Fortunately, we found such a problem. The coincidence is that it came from other conversations at Dagstuhl.

I presented [73] near the beginning of the seminar. One of the highlights of this work (presented here in Section 6.4) is a reduction from evaluating the Tutte polynomial of a planar graph at the point $(3, 3)$, a result by Las Vergnas [136]. One of the people at the seminar was Joanna Ellis-Monaghan. After my talk, she told me that she had generalized the result of Las Vergnas to any positive integer point on the diagonal [64]. We spent many hours together discussing this result

[3] The restriction imposed by planarity is the one given in Lemma 11.4.1.

and many other things, but when I left the seminar, I had still not internalized her generalization.

When I returned home, I began working in parallel. On the one hand, we determined that we could interpolate anything that satisfied a certain equation (cf. Section 11.4). On the other hand, I was emailing Joanna Ellis-Monaghan to help clarify my understanding of her generalization. I realized (on February 7th) that her generalization corresponded to a Holant problem that satisfied the same equation!

I typically think of Holant problems being so expressive and techniques like polynomial interpolation being so strong that there are many problems one can reduce from in order to prove new hardness results. This is *not* one of those cases. I know of no other way to show that counting edge κ-colorings is #P-hard over planar κ-regular graphs for $\kappa \geq 3$ except by using Joanna Ellis-Monaghan's work.

Another remark about the Bobby Fischer gadget The way we presented the Bobby Fischer gadget is not exactly how we found it. We make it look as though we used one degree of freedom to satisfy seven conditions (that is, setting seven signature entries to 0), but I think it is more accurate to say that we use two degrees of freedom to satisfy eight conditions. The eighth condition is implicit in Lemma 10.3.6. It is that we have a succinct ternary signature $\langle a, b, c \rangle$ of type τ_3 with $b = c$. This assumption is satisfied in the proof of Theorem 10.3.8 via a local holographic transformation, which uses one degree of freedom.

The "original Bobby Fischer gadget" was two copies of the same ternary local-holographic-transformation gadget (one degree of freedom) with one edge from each connected together through an arbitrary succinct binary signature of type τ_2 (a second degree of freedom). Then we looked for any interesting choices of those two parameters. At this point, it is rather obvious to consider the parameter choices that we did. What remains is to realize that the resulting signature can be used to interpolate $=_4$ and that $=_4$ quickly leads to a reduction from counting vertex colorings.

The way we presented it is our attempt to be conceptual as possible. We presented it as though one might have had the foresight to select these parameter choices from the beginning.

Chapter 11

Dichotomy for Higher Domain Holant Problems over Planar 3-Regular Graphs

We adapt and extend our proof techniques from the previous chapter to obtain a dichotomy for Pl-Holant$_\kappa(f)$, where f is a succinct ternary signature of type τ_3 on domain size $\kappa \geq 3$ with complex weights. A special case of this result is that counting edge κ-colorings is #P-hard over planar 3-regular graphs for all $\kappa \geq 3$. A key ingredient in the proof includes an effective version of Siegel's Theorem on finiteness of integer solutions for a specific algebraic curve. We use this to show that a specific polynomial $p(x, y)$ has an explicitly listed finite set of integer solutions. We also apply some elementary Galois theory and determine the Galois groups of some specific polynomials. This work was published in [32, 33].

11.1 Background

What do Siegel's Theorem and Galois theory have to do with complexity theory? In this chapter, we show that an effective version of Siegel's Theorem on finiteness of integer solutions for a specific algebraic curve and an application of elementary Galois theory are key ingredients in a chain of steps that lead to a complexity classification of some counting problems. More specifically, we consider a certain class of counting problems that are expressible as Holant problems with an arbitrary domain of size κ over 3-regular graphs, and prove a dichotomy theorem for this class of problems. The hardness, and thus the dichotomy, holds even when restricted to planar graphs. Among other things, the proof of the dichotomy theorem depends on the following: (A) the specific polynomial $p(x, y) = x^5 - 2x^3y - x^2y^2 - x^3 + xy^2 + y^3 - 2x^2 - xy$ has only the integer solutions $(x, y) = (-1, 1), (0, 0), (1, -1), (1, 2), (3, 3)$, and (B) the determination of the Galois groups of some specific polynomials.

The techniques we developed to prove Theorem 10.1.1 naturally extend to Holant problems defined by any succinct ternary signature of type τ_3. We prove a dichotomy theorem for such functions with complex weights.

Theorem 11.1.1. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c \in \mathbb{C}$. Then either* $\text{Holant}_\kappa(\langle a, b, c\rangle)$ *is computable in polynomial time or* $\text{Pl-Holant}_\kappa(\langle a, b, c\rangle)$ *is* #P*-hard. Furthermore, given a, b, c, there is a polynomial-time algorithm that decides which is the case.*

See Theorem 11.8.1 for an explicit listing of the tractable cases. Counting edge κ-colorings over 3-regular graphs is the special case when $\langle a, b, c\rangle = \langle 0, 0, 1\rangle$.

There is only one previous dichotomy theorem for higher domain Holant problems [51]. It is a dichotomy for a single symmetric ternary signature on domain size $\kappa = 3$ in the framework of Holant* problems, which means that all unary signatures are assumed to be freely available. An important difference is that Theorem 11.1.1 is for general domain size $\kappa \geq 3$ while the previous result is for domain size $\kappa = 3$. Dichotomy theorems for an arbitrary domain size are generally difficult to prove. The Feder-Vardi Conjecture for decision Constraint Satisfaction Problems (CSP) is still open [67]. It was a major achievement to prove this conjecture for domain size 3 [15]. The #CSP dichotomy was proved after a long series of work [17, 16, 14, 58, 13, 53, 22, 35, 60, 71, 36, 20].

In Theorem 11.1.2, the notation $f^\frown g$ denotes the signature that results from connecting one edge incident to a vertex assigned the signature f to one edge incident to a vertex assigned the signature g. When f and g are both unary signatures, which are represented by vectors, then the resulting 0-ary signature is just a scalar.

Theorem 11.1.2 (Theorem 3.1 in [51])**.** *Let f be a symmetric ternary signature on domain size* 3*. Then* $\operatorname{Holant}_3^*(f)$ *is* #P*-hard unless f is of one of the following forms, in which case, the problem is computable in polynomial time.*

1. *There exists $\alpha, \beta, \gamma \in \mathbb{C}^3$ that are mutually orthogonal (i.e. $\alpha^\frown\beta = \alpha^\frown\gamma = \beta^\frown\gamma = 0$) and*

$$f = \alpha^{\otimes 3} + \beta^{\otimes 3} + \gamma^{\otimes 3};$$

2. *There exists $\alpha, \beta_1, \beta_2 \in \mathbb{C}^3$ such that $\alpha^\frown\beta_1 = \alpha^\frown\beta_2 = \beta_1^\frown\beta_1 = \beta_2^\frown\beta_2 = 0$ and*

$$f = \alpha^{\otimes 3} + \beta_1^{\otimes 3} + \beta_2^{\otimes 3};$$

3. *There exists $\beta, \gamma \in \mathbb{C}^3$ and $f_\beta \in (\mathbb{C}^3)^{\otimes 3}$ such that $\beta \neq \mathbf{0}$, $\beta^\frown\beta = 0$, $f_\beta^\frown\beta = \mathbf{0}$, and*

$$f = f_\beta + \beta^{\otimes 2} \otimes \gamma + \beta \otimes \gamma \otimes \beta + \gamma \otimes \beta^{\otimes 2}.$$

Some domain invariant signatures are tractable by Theorem 11.1.2.

Corollary 11.1.3. *Suppose the domain size is* 3 *and $a, b, \lambda \in \mathbb{C}$. Let f be a succinct ternary signature of type τ_3. Then* $\operatorname{Holant}(f)$ *is computable in polynomial time when f has one of the following forms:*

1. $f = \lambda\langle 1,0,0\rangle = \lambda\left[(1,0,0)^{\otimes 3} + (0,1,0)^{\otimes 3} + (0,0,1)^{\otimes 3}\right]$;
2. $f = 3\lambda\langle -5,-2,4\rangle = \lambda\left[(1,-2,-2)^{\otimes 3} + (-2,1,-2)^{\otimes 3} + (-2,-2,1)^{\otimes 3}\right]$;
3. $f = \langle a,b,a\rangle = \frac{a+2b}{3}(1,1,1)^{\otimes 3} + \frac{a-b}{3}\left[(1,\omega,\omega^2)^{\otimes 3} + (1,\omega^2,\omega)^{\otimes 3}\right]$,

where ω is a primitive third root of unity.

In Corollary 11.1.3, form 1 is the ternary equality signature $=_3$, which is trivially tractable for any domain size. Then form 2 is just form 1 after a holographic transformation by the matrix $T = \left[\begin{smallmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{smallmatrix}\right]$, which is orthogonal after scaling by $\frac{1}{3}$. This is an example of two problems that must have the same complexity by Lemma 3.2.2.

The tractability of these two problems for higher domain sizes is stated in the following corollary.

Corollary 11.1.4. *Suppose $\kappa \geq 3$ is the domain size and $\lambda \in \mathbb{C}$. Let f be a succinct ternary signature of type τ_3. Then* $\operatorname{Holant}(f)$ *is computable in polynomial time if f has one of the following forms:*

1. $f = \lambda\langle 1, 0, 0\rangle$;
2. $f = \lambda T^{\otimes 3}\langle 1, 0, 0\rangle = \lambda\kappa\langle \kappa^2 - 6\kappa + 4, -2(\kappa - 2), 4\rangle$,

where $T = \kappa I_\kappa - 2J_\kappa$.

Note that $T = \kappa I_\kappa - 2J_\kappa$ is an orthogonal matrix after scaling by $\frac{1}{\kappa}$.

The two cases in Corollary 11.1.4 are respectively product type and transformable to product type. There are also two affine tractable cases that appear in our dichotomy. We already discussed them in Section 4.2. We restate the first using our succinct signature notation.

Corollary 11.1.5 (Restatement of Corollary 4.2.11)**.** *Suppose the domain size is 3 and suppose $a, c \in \mathbb{C}$. Let $T \in \mathbf{O}_3(\mathbb{C})$ and let $\langle a, 0, c\rangle$ be a succinct ternary signature of type τ_3. If $a^3 = c^3$, then* $\operatorname{Holant}_3(T^{\otimes 3}\langle a, 0, c\rangle)$ *is computable in polynomial time.*

The second result requires some additional proof.

Lemma 11.1.6. *Suppose the domain size is 4 and $\lambda, \mu \in \mathbb{C}$. Let $\langle \mu^2, 1, \mu\rangle$ be a succinct ternary signature of type τ_3. If $\mu = -1 + \varepsilon 2i$ with $\varepsilon = \pm 1$, then* $\operatorname{Holant}(\lambda\langle \mu^2, 1, \mu\rangle)$ *is computable in polynomial time.*

Proof. Let $T = \frac{1}{2}\left[\begin{smallmatrix} x & y & y & y \\ y & x & y & y \\ y & y & x & y \\ y & y & y & x \end{smallmatrix}\right]$, where $x = -3 - \varepsilon i$ and $y = 1 - \varepsilon i$. Then up to a factor of λ^n on graphs with n vertices, the output of $\operatorname{Holant}(\lambda\langle \mu^2, 1, \mu\rangle)$ is the same as the output for

$$\begin{aligned} \operatorname{Holant}(\langle \mu^2, 1, \mu\rangle) &= \operatorname{Holant}(\langle -3 - \varepsilon 4i, 1, -1 + \varepsilon 2i\rangle) \\ &\equiv_T \operatorname{Holant}(=_2 \mid T^{\otimes 3}(=_3)) \\ &\equiv \operatorname{Holant}((=_2)T^{\otimes 2} \mid =_3) \\ &\equiv \operatorname{Holant}(2\langle 1, -1\rangle \mid =_3). \end{aligned}$$

After removing the factor of 2 from the signature on the left, we are done by Lemma 4.2.14. $\square$

We restate this lemma as a simple corollary for later convenience.

Corollary 11.1.7. *Suppose the domain size is* 4 *and* $a, b, c \in \mathbb{C}$. *Let* $\langle a, b, c\rangle$ *be a succinct ternary signature of type* τ_3. *If* $a + 5b + 2c = 0$ *and* $5b^2 + 2bc + c^2 = 0$, *then* $\operatorname{Holant}(\langle a, b, c\rangle)$ *is computable in polynomial time.*

Proof. Since $a = -5b - 2c$ and $b = \frac{1}{5}(-1 \pm 2i)c$, after scaling by $\mu = -1 \mp 2i$, we have $\mu\langle a, b, c\rangle = c\langle \mu^2, 1, \mu\rangle$ and are done by Lemma 11.1.6. □

11.2 Proof Outline and Techniques

As usual, the difficult part of a dichotomy theorem is to carve out *exactly* the tractable problems in the class, and prove all the rest #P-hard. A dichotomy theorem for Holant problems has the additional difficulty that some tractable problems are only known to be tractable via a holographic reduction, which can make the appearance of the problem rather unexpected. The problem $\operatorname{Holant}_4(\langle -3-4i, 1, -1+2i\rangle)$ that we just discussed is of this type. In order to understand all problems in a Holant problem class, we must deal with such problems. Dichotomy theorems for graph homomorphisms and for #CSP do not have to deal with as varied a class of such problems, since they implicitly assume all Equality signatures are available and must be preserved. This restricts the possible transformations.

Our #P-hardness results are obtained by reducing from evaluations of the Tutte polynomial over planar graphs. A dichotomy is known for such problems (Theorem 6.1.4).

The chromatic polynomial, a specialization of the Tutte polynomial (Proposition 10.3.1), is concerned with vertex colorings. On domain size κ, one starting point of our hardness proofs is the chromatic polynomial, which contains the problem of counting vertex colorings using at most κ colors. By the planar dichotomy for the Tutte polynomial, this problem is #P-hard for all $\kappa \geq 3$.

Another starting point for our hardness reductions is the evaluation of the Tutte polynomial at an integer diagonal point (x, x), which is #P-hard for all $x \geq 3$ by the same planar Tutte dichotomy. These are new starting places for reductions involving Holant problems. These problems were known to have a so-called state-sum expression (Lemma 10.2.2), which is a sum over weighted Eulerian partitions. This sum is not over the original planar graph but over its directed medial graph, which is always a planar 4-regular graph (Figure 6.4 and Figure 10.1). We show that this state-sum expression is naturally expressed as a Holant problem with a particular quaternary signature (Lemma 10.2.5).

To reduce from these two problems, we execute the following strategy. First, we attempt to construct the succinct unary signature $\langle 1 \rangle$ of type τ_1. (Lemma 11.5.1). Second, we attempt to interpolate all succinct binary signatures assuming that we have $\langle 1 \rangle$ (Section 11.6). Lastly, we attempt to construct a succinct ternary signature $\langle a, b, c \rangle$ of type τ_3 with the special property $b = c$ assuming that all these binary signatures are available (Lemma 11.7.1). At each step, there are some problems specified by $\langle a, b, c \rangle$ for which our attempts fail. In such cases, we directly obtain a dichotomy without the help of additional signatures. See Figure 11.1 for a flow chart of hardness reductions.

Below we highlight some of our proof techniques.

Interpolation within an orthogonal subspace We develop the ability to interpolate when faced with some nontrivial null spaces inherently present in interpolation constructions. In any construction involving an initial signature and a recurrence matrix, it is possible that the initial signature is orthogonal to some row eigenvectors of the recurrence matrix. Previous interpolation results always attempt to find a construction that avoids this. In the present chapter, this avoidance seems impossible. In Section 11.3, we prove an interpolation result that can succeed in this situation to the greatest extent possible. We prove that one can interpolate any signature provided that it is orthogonal to the same set of row eigenvectors, and the relevant eigenvalues satisfy a lattice condition (Lemma 11.3.6).

Satisfy lattice condition via Galois theory A key requirement for this interpolation to succeed is the lattice condition (Definition 11.3.3), which involves the roots of the characteristic polynomial of the recurrence matrix. We use Galois theory to prove that our constructions satisfy this condition. If a polynomial has a large Galois group, such as S_n or A_n, and its roots do not all have the same complex norm, then we show that its roots satisfy the lattice condition (Lemma 11.3.5).

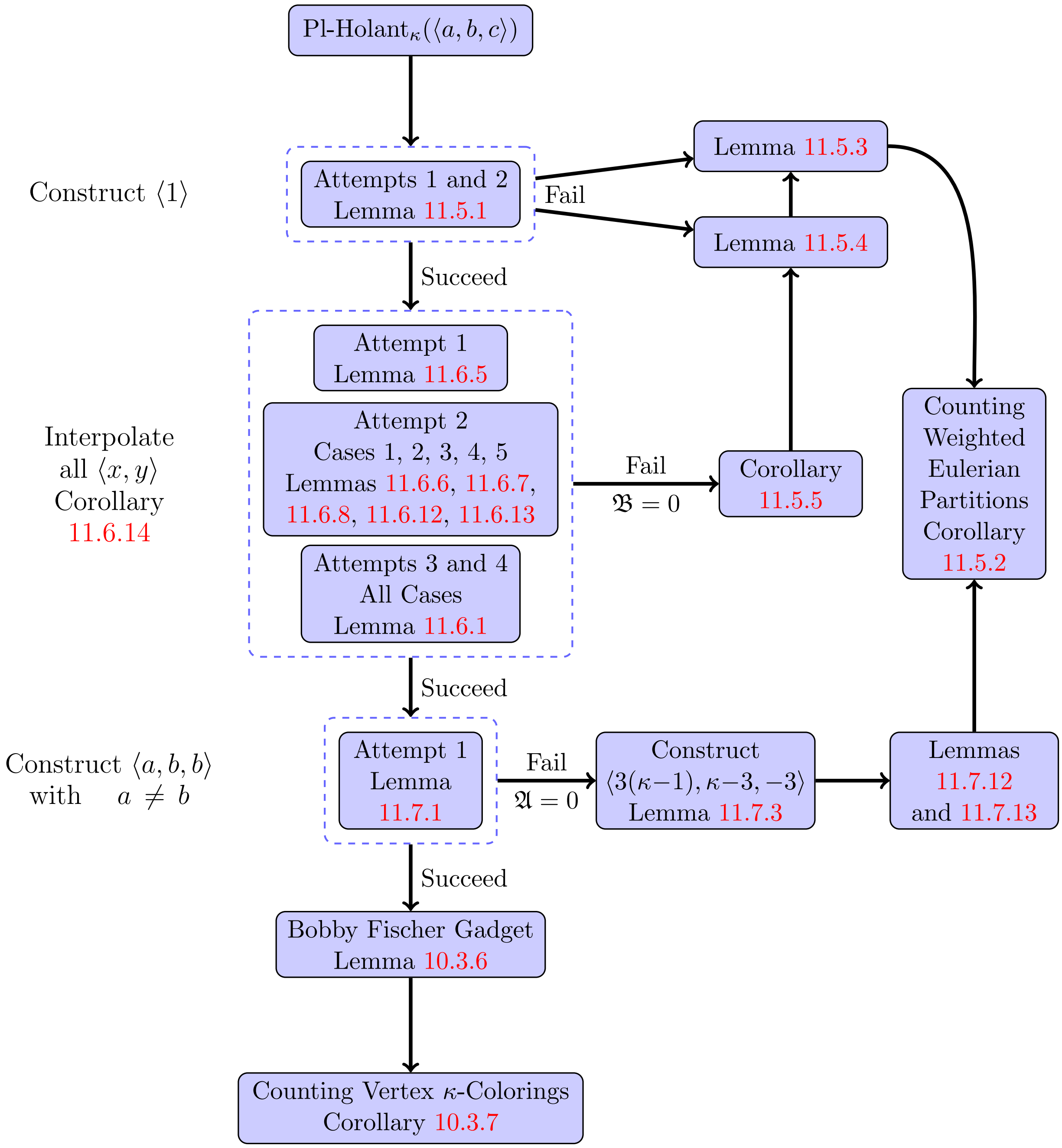


Figure 11.1: Flow chart of hardness reductions in our proof of Theorem 11.1.1 going back to our two starting points of hardness.

Effective Siegel's Theorem via Puiseux series We need to determine the Galois groups for an infinite family of polynomials, one for each domain size. If these polynomials are irreducible, then we can show they all have the full symmetric group as their Galois group, and hence fulfill the lattice condition. We suspect that these polynomials are all irreducible but are unable to prove it.

A necessary condition for irreducibility is the absence of any linear factor. This infinite family of polynomials, as a single bivariate polynomial in (x, κ), defines an algebraic curve, which has genus 3. By a well-known theorem of Siegel [115], there are only a finite number of integer values of κ for which the corresponding polynomial has a linear factor. However this theorem and others like it are not *effective* in general. There are some effective versions of Siegel's Theorem that can be applied to the algebraic curve, but the best general effective bound is over $10^{20,000}$ [140] and hence cannot be checked in practice. Instead, we use Puiseux series to show that this algebraic curve has exactly five explicitly listed integer solutions (Lemma 11.7.6).

Eigenvalue Shifted Triples For a pair of eigenvalues, the lattice condition is equivalent to the statement that the ratio of these eigenvalues is not a root of unity. A sufficient condition is that the eigenvalues have distinct complex norms. We prove three results, each of which is a different way to satisfy this sufficient condition. Chief among them is the technique we call an *Eigenvalue Shifted Triple* (EST). These generalize the technique of Eigenvalue Shifted Pairs from [93]. In an EST, we have three recurrence matrices, each of which differs from the other two by a nonzero additive multiple of the identity matrix. Provided these two multiples are linearly independent over $\mathbb{R}$, we show at least one of these matrices has eigenvalues with distinct complex norms (Lemma 11.6.11). (However determining which one succeeds is a difficult task; but we need not know that).

E Pluribus Unum When the ratio of a pair of eigenvalues is a root of unity, it is a challenge to effectively use this failure condition. Direct application of this cyclotomic condition is often of limited use. We introduce an approach that uses this cyclotomic condition effectively. A direct recursive construction involving these two eigenvalues only creates a finite number of different signatures. We reuse all of these signatures in a multitude of new interpolation constructions (Lemma 11.6.4), one of which we hope will succeed. If the eigenvalues in all of these constructions also satisfy a cyclotomic condition, then we obtain a more useful condition than any of the previous cyclotomic conditions. This idea generalizes the anti-gadget technique from Chapter 5, which only reuses the "last" of these signatures.

Local holographic transformation One reason to obtain all succinct binary signatures is for use in the gadget construction known as a local holographic transformation (Figure 11.8). This construction mimics the effect of a holographic transformation applied on a single signature. In particular, using this construction, we attempt to obtain a succinct ternary signature of the form $\langle a, b, b \rangle$, where $a \neq b$ (Lemma 11.7.1). This signature turns out to have some magical properties in the Bobby Fischer gadget, which we discuss next.

Bobby Fischer gadget Typically, any combinatorial construction for higher domain Holant problems produces very intimidating looking expressions that are nearly impossible to analyze. In our case, it seems necessary to consider a construction that has to satisfy multiple requirements involving at least nine polynomials. However, we are able to combine the signature $\langle a, b, b \rangle$, where $a \neq b$, with a succinct binary signature of our choice in a special construction that we call the *Bobby Fischer gadget* (Figure 10.6). This gadget is able to satisfy seven conditions using just one degree of freedom (Lemma 10.3.6). This ability to satisfy a multitude of constraints simultaneously in one magic stroke reminds us of some unfathomably brilliant moves by Bobby Fischer, the chess genius extraordinaire. We first encountered the Bobby Fischer gadget in the previous chapter in Section 10.3, but it is a crucial ingredient in the current chapter as well.

11.3 An Interpolation Result

The goal of this section is to generalize an interpolation result from [50], which we rephrase using our notion of a succinct signature (cf. Lemma 10.1.3).

Theorem 11.3.1 (Theorem 3.5 in [50]). *Suppose $\mathcal{F}$ is a set of signatures over a domain of size κ and τ is a succinct signature type with length 3. If there exists an infinite sequence of planar $\mathcal{F}$-gates defined by an initial succinct signature $s \in \mathbb{C}^{3\times 1}$ of type τ and a recurrence matrix $M \in \mathbb{C}^{3\times 3}$*

with eigenvalues α, β, and γ satisfying the following conditions:

1. $\det(M) \neq 0$;
2. *s is not orthogonal to any row eigenvector of M;*
3. *for all $(i,j,k) \in \mathbb{Z}^3 - \{(0,0,0)\}$ with $i+j+k=0$, we have $\alpha^i\beta^j\gamma^k \neq 1$;*

then

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{f\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F}),$$

for any succinct ternary signature f of type τ.

Our generalization of this result is designed to relax the second condition so that s can be orthogonal to some row eigenvectors of M. Suppose r is a row eigenvector of M, with eigenvalue λ, that is orthogonal to s (i.e. the dot product $r \cdot s$ is 0). Consider $M^k s$, the kth signature in the infinite sequence defined by M and s. This signature is also orthogonal to r since $r \cdot M^k s = \lambda^k r \cdot s = 0$. We do not know of any way of interpolating a signature using this infinite sequence that is not also orthogonal to r. On the other hand, we would like to interpolate those signatures that do satisfy this orthogonality condition. Our interpolation result gives a sufficient condition to achieve this.

We assume our n-by-n matrix M is diagonalizable, i.e., it has n linearly independent (row and column) eigenvectors. We do not assume that M necessarily has n distinct eigenvalues (although this would be a sufficient condition for it to be diagonalizable). The relaxation of the second condition is that, for some positive integer ℓ, the initial signature s is *not* orthogonal to exactly ℓ of these linearly independent row eigenvectors of M. To satisfy this condition, we use a two-step approach. First, we explicitly exhibit $n - \ell$ linearly independent row eigenvectors of M that are orthogonal to s. Then we use the following lemma to show that the remaining row eigenvectors of M are not orthogonal to s. The justification for this approach is that the eigenvectors orthogonal to s are often simple to express while the eigenvectors not orthogonal to s tend to be more complicated.

Lemma 11.3.2. *For $n \in \mathbb{Z}^+$, let $s \in \mathbb{C}^{n\times 1}$ be a vector and let $M \in \mathbb{C}^{n\times n}$ be a diagonalizable matrix. If $\operatorname{rank}([s \; Ms \; \dots \; M^{n-1}s]) \geq \ell$, then for any set of n linearly independent row eigenvectors, s is not orthogonal to at least ℓ of them.*

Proof. Since M is diagonalizable, it has n linearly independent eigenvectors. Suppose for a contradiction that there exists a set of n linearly independent row eigenvectors of M such that s is orthogonal to $t > n-\ell$ of them. Let $N \in \mathbb{C}^{t\times n}$ be the matrix whose t rows are the row eigenvectors of M that are orthogonal to s. Then $N[s\ Ms\ \ldots\ M^{n-1}s]$ is the zero matrix. From this, it follows that $\operatorname{rank}([s\ Ms\ \ldots\ M^{n-1}s]) < \ell$, a contradiction. □

Remark. The last inequality is known as Sylvester's rank inequality.

The third condition of Theorem 11.3.1 is also known as the lattice condition.

Definition 11.3.3. Fix some $\ell \in \mathbb{N}$. We say that $\lambda_1, \lambda_2, \ldots, \lambda_\ell \in \mathbb{C} - \{0\}$ satisfy the *lattice condition* if for all $x \in \mathbb{Z}^\ell - \{\mathbf{0}\}$ with $\sum_{i=1}^{\ell} x_i = 0$, we have $\prod_{i=1}^{\ell} \lambda_i^{x_i} \neq 1$.

When $\ell \geq 3$, we use Galois theory to show that the lattice condition is satisfied. The idea is that the lattice condition must hold if the Galois group of the polynomial, whose roots are the λ_i's, is large enough. In [50], for the special case $n = \ell = 3$, it was shown that the roots of most cubic polynomials satisfy the lattice condition using this technique.

Lemma 11.3.4 (Lemma 5.2 in [50])**.** *Let $f(x) \in \mathbb{Q}[x]$ be an irreducible cubic polynomial. Then the roots of $f(x)$ satisfy the lattice condition iff $f(x)$ is not of the form $ax^3 + b$ for some $a, b \in \mathbb{Q}$.*

In the following lemma, we show that if the Galois group for a polynomial of degree n is one of the two largest possible groups, S_n or A_n, then its roots satisfy the lattice condition provided these roots do not all have the same complex norm.

Lemma 11.3.5. *Let f be a polynomial of degree $n \geq 2$ with rational coefficients. If the Galois group of f over $\mathbb{Q}$ is S_n or A_n and the roots of f do not all have the same complex norm, then the roots of f satisfy the lattice condition.*

Proof. We consider A_n since the same argument applies to $S_n \supset A_n$. For $1 \leq i \leq n$, let a_i be the roots of f such that $|a_1| \leq \cdots \leq |a_n|$. By assumption, as least one of these inequalities is strict. Suppose for a contradiction that these roots fail to satisfy the lattice condition. This means there exists $x \in \mathbb{Z}^n - \{\mathbf{0}\}$ satisfying $\sum_{i=1}^{n} x_i = 0$ such that

$$a_1^{x_1} \cdots a_n^{x_n} = 1. \tag{11.3.1}$$

Since x is not all 0, it must contain some positive entries and some negative entries. We can rewrite (11.3.1) as $b_1^{y_1} \cdots b_s^{y_s} = c_1^{z_1} \cdots c_t^{z_t}$, where $s, t \geq 1$, $b_1, \ldots, b_s, c_1, \ldots, c_t$ are $s+t$ distinct members from $\{a_1, \ldots, a_n\}$, $y_i > 0$ for $1 \leq i \leq s$, $z_i > 0$ for $1 \leq i \leq t$, and $y_1 + \cdots + y_s = z_1 + \cdots + z_t$. We omit factors in (11.3.1) with exponent 0.

If $n = 2$, then $s = t = 1$ and $|b_1| = |c_1|$. This is a contradiction to the assumption that roots of f do not all have the same complex norm. Otherwise, assume $n \geq 3$. If $s = t = 1$, then $|b_1| = |c_1|$ again. We apply 3-cycles from A_n to conclude that all roots of f have the same complex norm, a contradiction. Otherwise $s + t > 2$. Without loss of generality, suppose $s \geq t$, which implies $s \geq 2$. Pick $j \in \{0, \ldots, n-s-t\}$ such that $|a_{j+1}| \leq \cdots \leq |a_{j+s+t}|$ contains a strict inequality. We permute the roots so that $b_i = a_{j+i}$ for $1 \leq i \leq s$ and $c_i = a_{j+s+i}$ for $1 \leq i \leq t$ (or possibly swapping b_1 and b_2 if necessary to ensure the permutation is in A_n). Then taking the complex norm of both sides gives a contradiction. □

Remark. This result can simplify the interpolation arguments in [50]. Since each of their cubic polynomials is irreducible, the corresponding Galois groups are transitive subgroups of S_3, namely S_3 or A_3 (and in fact by inspection, they are all S_3). Then Lemma 4.5 from [94] shows that the eigenvalues of these polynomials do not all have the same complex norm. Thus, the roots of all polynomials exhibited in [50] satisfy the lattice condition by Lemma 11.3.5.

We apply Lemma 11.3.5 to an infinite family of quintic polynomials that we encounter in Section 11.7. If the polynomials are irreducible, then we are able to apply this lemma. Unfortunately, we are unable to show that all these polynomials are irreducible and thus also have to consider the possible ways in which they could factor. Nevertheless, we are still able to show that all these polynomials satisfy the lattice condition.

To conclude, we state and prove our new interpolation result.

Lemma 11.3.6. *Suppose $\mathcal{F}$ is a set of signatures over a domain of size κ and τ is a succinct signature type with length $n \in \mathbb{Z}^+$. If there exists an infinite sequence of planar $\mathcal{F}$-gates defined by an initial succinct signature $s \in \mathbb{C}^{n \times 1}$ of type τ and a recurrence matrix $M \in \mathbb{C}^{n \times n}$ satisfying the following conditions,*

1. *M is diagonalizable with n linearly independent eigenvectors;*
2. *s is not orthogonal to exactly ℓ of these linearly independent row eigenvectors of M with eigenvalues $\lambda_1, \ldots, \lambda_\ell$;*
3. *$\lambda_1, \ldots, \lambda_\ell$ satisfy the lattice condition;*

then

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{f\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F})$$

for any succinct signature f of type τ that is orthogonal to the $n - \ell$ of these linearly independent eigenvectors of M to which s is also orthogonal.

Proof. Let $\lambda_1, \ldots, \lambda_n$ be the n eigenvalues of M, with possible repetition. Since M is diagonalizable, we can write M as $T\Lambda T^{-1}$, where Λ is the diagonal matrix $\left[\begin{smallmatrix} B_1 & \mathbf{0} \\ \mathbf{0} & B_2 \end{smallmatrix}\right]$ with $B_1 = \operatorname{diag}(\lambda_1, \ldots, \lambda_\ell)$ and $B_2 = \operatorname{diag}(\lambda_{\ell+1}, \ldots, \lambda_n)$. Notice that the columns of T are the column eigenvectors of M and the rows of T^{-1} are the row eigenvectors of M. Let $\mathbf{t}_i$ be the ith column T and let $T^{-1}s = [\alpha_1 \ \ldots \ \alpha_n]^\intercal$. Then $\alpha_i \neq 0$ for $1 \leq i \leq \ell$ and $\alpha_i = 0$ for $\ell < i \leq n$, since s is not orthogonal to exactly the first ℓ row eigenvectors of M.

Now we can write

$$\begin{aligned} M^k s = T \begin{bmatrix} B_1^k & \mathbf{0} \\ \mathbf{0} & B_2^k \end{bmatrix} T^{-1} s = T \begin{bmatrix} B_1^k & \mathbf{0} \\ \mathbf{0} & B_2^k \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_\ell \\ 0 \\ \vdots \\ 0 \end{bmatrix} &= T \operatorname{diag}(\alpha_1 \lambda_1^k, \ldots, \alpha_\ell \lambda_\ell^k, 0, \ldots, 0) \\ = T \operatorname{diag}(\alpha_1, \ldots, \alpha_\ell, 0, \ldots, 0) \begin{bmatrix} \lambda_1^k \\ \vdots \\ \lambda_\ell^k \\ 0 \\ \vdots \\ 0 \end{bmatrix} &= [\alpha_1 \mathbf{t}_1, \ldots, \alpha_\ell \mathbf{t}_\ell, \mathbf{0}, \ldots, \mathbf{0}] \begin{bmatrix} \lambda_1^k \\ \vdots \\ \lambda_\ell^k \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \end{aligned}$$

For $1 \leq i \leq \ell$, let $\mathbf{t}'_i = \alpha_i \mathbf{t}_i$. Both the columns of T and the rows of T^{-1} are linearly independent. From $T^{-1}T = I_m$, we see that $\mathbf{t}_i$ for $1 \leq i \leq \ell$ is orthogonal to the last $n - \ell$ rows of T^{-1}. Thus $\operatorname{span}\{\mathbf{t}_1, \ldots, \mathbf{t}_\ell\} = \operatorname{span}\{\mathbf{t}'_1, \ldots, \mathbf{t}'_\ell\}$ is precisely the space of vectors orthogonal to the last $n - \ell$ rows of T^{-1}.

Consider an instance Ω of $\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{f\})$. Let V_f be the subset of vertices assigned f

with $n_f = |V_f|$. Since f is orthogonal to any row eigenvector of M to which s is also orthogonal, we have $f \in \operatorname{span}\{\mathbf{t}'_1, \ldots, \mathbf{t}'_\ell\}$. Let $f = \beta_1 \mathbf{t}'_1 + \cdots + \beta_\ell \mathbf{t}'_\ell$. Then $\operatorname{Holant}_\kappa(\Omega; \mathcal{F} \cup \{f\})$ is a homogeneous polynomial in the β_i's of total degree n_f. For $y = (y_1, \ldots, y_\ell) \in \mathbb{N}^\ell$, let c_y be the coefficient of $\beta_1^{y_1} \cdots \beta_\ell^{y_\ell}$ in $\operatorname{Holant}_\kappa(\Omega; \mathcal{F} \cup \{f\})$ so that

$$\operatorname{Holant}_\kappa(\Omega; \mathcal{F} \cup \{f\}) = \sum_{y_1 + \cdots + y_\ell = n_f} c_y \beta_1^{y_1} \cdots \beta_\ell^{y_\ell}.$$

We construct from Ω a sequence of instances Ω_k of Pl-Holant$_\kappa(\mathcal{F})$ indexed by $k \in \mathbb{N}$. We obtain Ω_k from Ω by replacing each occurrence of f with $M^k s$, for $k \geq 0$. Then

$$\operatorname{Holant}_\kappa(\Omega_k; \mathcal{F}) = \sum_{y_1 + \cdots + y_\ell = n_f} c_y \left(\lambda_1^{y_1} \cdots \lambda_\ell^{y_\ell}\right)^k.$$

Note that, crucially, the same c_y coefficients appear. We treat this as a linear system with the c_y's as the unknowns. The coefficient matrix is a Vandermonde matrix of order $\binom{n_f + \ell - 1}{\ell - 1}$, which is polynomial in n_f and thus polynomial in the size of Ω. It is nonsingular if every $\lambda_1^{y_1} \cdots \lambda_\ell^{y_\ell}$ is distinct, which is indeed the case since $\lambda_1, \ldots, \lambda_\ell$ satisfy the lattice condition.

Therefore, we can solve for the c_y's in polynomial time and compute $\operatorname{Holant}_\kappa(\Omega; \mathcal{F} \cup \{f\})$. □

Remark. When restricted to $n = \ell = 3$, this proof is simpler than the one given in [50] for Theorem 11.3.1 due to our implicit use of a local holographic transformation (i.e. the writing of f as a linear combination of $\mathbf{t}'_1, \ldots, \mathbf{t}'_\ell$ and expressing $\operatorname{Holant}_\kappa(\Omega; \mathcal{F} \cup \{f\})$ in terms of this).

11.4 Invariance Properties from Row Eigenvectors

Before we launch into our hardness proof of Theorem 11.1.1, we consider a simple example that shows how to apply the interpolation results from the previous section. This section also shows how a recursive construction in an interpolation proof can be used to form a hypothesis about possible invariance properties.

We often find that no matter what constructions one considers, all signatures they produce satisfy certain invariance. Instead of defining this notion formally, we prove the following lemma as an example. After this lemma and its proof, we explain that this invariance can be suggested by certain recursive constructions in an alternative proof of Theorem 10.2.7, that it is #P-hard to count edge κ-coloring over planar κ-regular graphs for all $\kappa \geq 3$. This alternative proof uses the interpolation techniques that we developed in Section 6.2.

Lemma 11.4.1. *Suppose $\kappa \geq 3$ is the domain size. If F is a planar $\{\mathrm{AD}_\kappa\}$-gate with succinct quaternary signature $\langle a, b, c, d, e \rangle$ of type τ_{color}, then $a + c = b + d$.*

Proof. Fix two distinct colors $g, y \in [\kappa]$. We define the *swap* of an edge colored g or y to be the opposite of these two colors. That is, swapping the color of an edge colored g (resp. y) gives the same edge colored y (resp. g). The ith external edge of F is the external edge that corresponds to the ith input of F. Recall that the input edges of F are ordered cyclically.

For $1 \leq i \leq 4$, let S_i (resp. S'_i) be the set of colorings of the edges (both internal and external) of F with an external coloring in the partition P_i of the succinct signature type τ_{color} such that the first external edge of F is colored g (resp. y) and the remaining external edges are either colored g or y (as dictated by P_i). Note that $|S_i| = |S'_i|$ for $1 \leq i \leq 4$. Furthermore, the sizes of these sets do not depend on the choice of $g, y \in [\kappa]$. Thus, it suffices to show that

$$|S_1 \cup S'_1 \cup S_3 \cup S'_3| = |S_2 \cup S'_2 \cup S_4 \cup S'_4|. \tag{11.4.2}$$

Let $\sigma \in S_1 \cup S'_1 \cup S_3 \cup S'_3$ be a coloring of F. Starting at the first external edge of F, there is a unique path π_1 that alternates in edge colors between g and y and terminates at another external edge of F. Suppose for a contradiction that this path terminates at the third external edge of F. Also consider the unique path π_2 that starts at the second external edge of F, alternates in edge colors between g and y, and must terminate at the fourth external edge of F. These two paths must cross somewhere since their ends are crossed. By planarity, they must cross at a vertex, and yet they must be vertex disjoint. This is a contradiction. Therefore, the path π_1 either terminates at the second or fourth external edge of F.

Suppose π_1 terminates at the second external edge of F. If $\sigma \in S_1$ (resp. $\sigma \in S'_1$), then swapping the colors of every edge in π_1 gives a new coloring $\pi'_1 \in S'_2$ (resp. $\pi'_1 \in S_2$). Similarly, if

$\sigma \in S_3$ (resp. $\sigma \in S_3'$), then swapping the colors of every edge in π_1 gives a new coloring $\pi_1' \in S_4'$ (resp. $\pi_1' \in S_4$).

Otherwise, π_1 terminates at the fourth external edge of F. If $\sigma \in S_1$ (resp. $\sigma \in S_1'$), then swapping the colors of every edge in π_1 gives a new coloring $\pi_1' \in S_4'$ (resp. $\pi_1' \in S_4$). Similarly, if $\sigma \in S_3$ (resp. $\sigma \in S_3'$), then swapping the colors of every edge in π_1 gives a new coloring $\pi_1' \in S_2'$ (resp. $\pi_1' \in S_2$).

Furthermore, this mapping from $S_1 \cup S_1' \cup S_3 \cup S_3'$ to $S_2 \cup S_2' \cup S_4 \cup S_4'$ is invertible. Therefore, we have established (11.4.2), as desired. □

Now we give an alternative proof of Theorem 10.2.7. The recursive construction in this proof will suggest the invariance in Lemma 11.4.1.

Let $q(x,\kappa) = x^3 - x^2 + x - (\kappa - 1)$. First we determine the nature of the roots of $q(x,\kappa)$.

Lemma 11.4.2. *For all $\kappa \in \mathbb{Z}$, the polynomial $q(x,\kappa)$ in x has one real root $r \in \mathbb{R}$ and two nonreal complex conjugate roots $\alpha, \overline{\alpha} \in \mathbb{C}$, such that $\alpha + \overline{\alpha} = 1 - r$ and $\alpha\overline{\alpha} = r^2 - r + 1$.*

Furthermore, if $q(x,\kappa)$ is reducible in $\mathbb{Q}[x]$ and $\kappa \geq 3$, then $r \geq 2$ is an integer.

Proof. The discriminant of $q(x,\kappa)$ with respect to x is $\operatorname{disc}_x(q(x,\kappa)) = -27\kappa^2 + 68\kappa - 44 < 0$, so $q(x,\kappa)$ has one real root $r \in \mathbb{R}$ and two nonreal complex conjugate roots $\alpha, \overline{\alpha} \in \mathbb{C}$. We have

$$\begin{aligned} \alpha + \overline{\alpha} + r &= 1 \\ \alpha\overline{\alpha} + (\alpha + \overline{\alpha})r &= 1 \\ \alpha\overline{\alpha}r &= \kappa - 1. \end{aligned}$$

It follows that $\alpha + \overline{\alpha} = 1 - r$, $\alpha\overline{\alpha} = r^2 - r + 1$, and

$$\kappa = r^3 - r^2 + r + 1. \tag{11.4.3}$$

If $q(x,\kappa)$ is reducible in $\mathbb{Q}[x]$ with $\kappa \geq 3$, then $r \in \mathbb{Z}$ by Gauss's Lemma and so $r \geq 2$ by (11.4.3). □

Lemma 11.4.3. *If $\kappa \geq 3$ is an integer, then the roots of $x^3 - x^2 + x - (\kappa - 1)$ satisfy the lattice condition.*

Proof. If $q(x, \kappa)$ is irreducible in $\mathbb{Q}[x]$, then its roots satisfy the lattice condition by Lemma 11.3.4.

Otherwise, $q(x, \kappa)$ is reducible in $\mathbb{Q}[x]$. By Lemma 11.4.2, $q(x, \kappa)$ has one real root $r \in \mathbb{Z}$ satisfying $r \geq 2$ and two nonreal complex conjugate roots $\alpha, \overline{\alpha} \in \mathbb{C}$ satisfying $\alpha + \overline{\alpha} = 1 - r$ and $\alpha\overline{\alpha} = r^2 - r + 1$. Suppose there exist $i, j, k \in \mathbb{Z}$ such that $\alpha^i \overline{\alpha}^j = r^k$ and $i + j = k$. We want to show that $i = j = k = 0$.

There is an element in the Galois group of $q(x, \kappa)$ that fixes $\mathbb{Q}$ pointwise and swaps α and $\overline{\alpha}$. Thus $\alpha^j \overline{\alpha}^i = r^k$. Dividing these two equations gives $(\alpha/\overline{\alpha})^{i-j} = 1$. We claim that $\omega = \alpha/\overline{\alpha}$ cannot be a root of unity and hence $i = j$. For a contradiction, suppose ω is a dth primitive root of unity. Let $f(x) = (x - \alpha)(x - \overline{\alpha}) = x^2 + (r - 1)x + (r^2 - r + 1) \in \mathbb{Z}[x]$. Then ω belongs to the splitting field of f over $\mathbb{Q}$, which is an extension of degree two over $\mathbb{Q}$. This implies that the Euler totient function $\phi(d) \mid 2$. Therefore $d \in \{1, 2, 3, 4, 6\}$. Let $\rho = \frac{\alpha+\overline{\alpha}}{\alpha\overline{\alpha}} = \frac{1+\omega}{\omega\overline{\alpha}} = \frac{1-r}{r^2-r+1} \in \mathbb{Q}$. Since $r \geq 2$, we have $\rho \neq 0$ and hence $d \neq 2$. Moreover, $f(x) = x^2 - (2 + \omega + \omega^{-1})\rho^{-1}x + (2 + \omega + \omega^{-1})\rho^{-2}$. Notice that the quantity $2 + \omega + \omega^{-1}$ is $4, 1, 2, 3$ respectively, when $d = 1, 3, 4, 6$. As $(2 + \omega + \omega^{-1})\rho^{-2} \in \mathbb{Z}$, we get that ρ^{-1} must be an integer when $d = 3, 4, 6$ and half an integer when $d = 1$. However $\rho^{-1} = -r + \frac{1}{r-1}$. The only possibility is $r = 3$ and $d = 1$; yet it is easy to check that $\omega \neq 1$ when this holds. This proves the claim.

From $\alpha\overline{\alpha} = r^2 - r + 1$, we have $(r^2 - r + 1)^i = (\alpha\overline{\alpha})^i = r^k$. Since r and $r^2 - r + 1$ are relatively prime and $r \geq 2$, we must have $i = k = 0$. □

Alternative proof of Theorem 10.2.7. As before, let $\langle 2, 1, 0, 1, 0 \rangle$ be a succinct quaternary signature of type τ_{color}. We reduce from Pl-Holant$_\kappa(\langle 2, 1, 0, 1, 0 \rangle)$ to Pl-Holant$_\kappa(\text{AD}_\kappa)$, which denotes the problem of counting edge κ-colorings in planar κ-regular graphs as a Holant problem. Then by Corollary 10.2.6, we conclude that Pl-Holant$_\kappa(\text{AD}_\kappa)$ is #P-hard.

Consider the gadget in Figure 10.2, where the bold edge represents $\kappa - 2$ parallel edges. We assign AD_κ to both vertices. Up to a nonzero factor of $(\kappa - 2)!$, this gadget has the succinct quaternary signature $f = \langle 0, 1, 1, 0, 0 \rangle$ of type τ_{color}. Now consider the recursive construction in

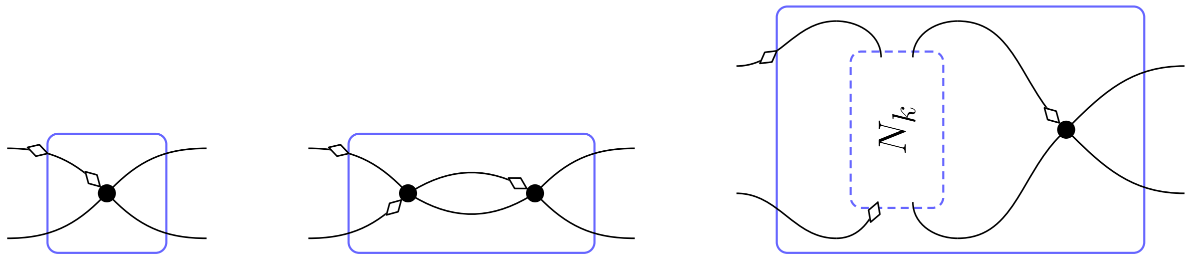


Figure 11.2: Alternate recursive construction to interpolate $\langle 2,1,0,1,0\rangle$. The vertices are assigned the signature of the gadget in Figure 10.2.

Figure 11.2. All vertices are assigned the signature f. Let f_s be the succinct quaternary signature of type τ_{color} for the sth gadget of the recursive construction. Then $f_0 = f$ and $f_s = M^s f_0$, where

$$M = \begin{bmatrix} 0 & 0 & 0 & \kappa-1 & 0 \\ 1 & 0 & 0 & \kappa-2 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The row vectors

$$(1,-1,1,-1,0) \qquad \text{and} \qquad (0,0,0,0,1)$$

are linearly independent row eigenvectors of M, with eigenvalues -1 and 1 respectively, that are orthogonal to the initial signature f_0. Note that our target signature $\langle 2,1,0,1,0\rangle$ is also orthogonal to these two row eigenvectors.

Up to a factor of $(x-1)(x+1)$, the characteristic polynomial of M is $x^3 - x^2 + x - (\kappa - 1)$. The roots of this polynomial satisfy the lattice condition by Lemma 11.4.3. In particular, these three roots are distinct. By Lemma 11.4.2, the only real root is at least 2. Thus, all five eigenvalues of M are distinct, so M is diagonalizable.

The 3-by-3 matrix in the upper-left corner of $[f_0\ Mf_0\ \ldots\ M^4 f_0]$ is $\left[\begin{smallmatrix} 0 & 0 & \kappa-1 \\ 1 & 0 & \kappa-2 \\ 1 & 1 & 0 \end{smallmatrix}\right]$. Its determinant is $\kappa - 1 \neq 0$. Thus, $[f_0\ Mf_0\ \ldots\ M^4 f_0]$ has rank at least 3, so by Lemma 11.3.2, f_0 is not orthogonal to the three remaining row eigenvectors of M.

Therefore, by Lemma 11.3.6, we can interpolate $\langle 2, 1, 0, 1, 0\rangle$, which completes the proof. $\square$

Notice that the row eigenvector $(1, -1, 1, -1, 0)$ suggests that $a - b + c - d = 0$ is an invariance shared by all signatures of symmetric ternary constructions. Some row eigenvectors, like $(0, 0, 0, 0, 1)$, only indicate an invariance present in some recursive constructions. (When $\kappa = 4$, there are recursive constructions for which $(0, 0, 0, 0, 1)$ is not a row eigenvector of the recurrence matrix.) The row eigenvector $(1, -1, 1, -1, 0)$ is more intrinsic; it must appear because of the invariance present in all constructions as shown in Lemma 11.4.1.

This suggests an approach to discover new invariance properties. Given a set $\mathcal{F}$ of signatures, create some recursive construction and inspect the row eigenvectors of the resulting recurrence matrix. For example, consider the set $\mathcal{F}_{\mathfrak{A}} = \{\langle a, b, c\rangle \mid a, b, c \in \mathbb{C} \text{ and } \mathfrak{A} = 0\}$, where $\mathfrak{A} = a - 3b + 2c$. It seems that $\mathcal{F}_{\mathfrak{A}}$ is closed under symmetric ternary constructions, such as those in Section 11.7.1. In particular, $(1, -3, 2)$ is a row eigenvector of the recurrence matrix for every recursive ternary construction with symmetric signatures that we tried. However, we do not know how to prove this closure property.

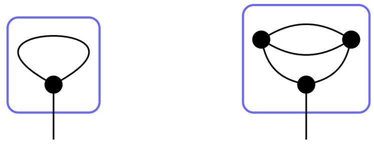
11.5 Constructing a Nonzero Unary Signature

The primary goal of this section is to construct the succinct unary signature $\langle 1\rangle$ of type τ_1. However, this is not always possible. For example, the succinct ternary signature $\langle 0, 0, 1\rangle = \mathrm{AD}_3$ of type τ_3 (over a domain of size 3) cannot construct $\langle 1\rangle$. This follows from the parity condition (Lemma 10.2.3). In such cases, we show that the problem is either computable in polynomial time or #P-hard without the help of additional signatures.

Lemma 11.5.1 handles two easy cases for which it is possible to construct $\langle 1\rangle$.

Lemma 11.5.1. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle a, b, c\rangle$ of type τ_3. If $a + (\kappa - 1)b \neq 0$ or $[2b + (\kappa - 2)c][b^2 - 4bc - (\kappa - 3)c^2] \neq 0$, then*

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle 1\rangle\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F}),$$



(a) (b)

Figure 11.3: Two simple unary gadgets.

where $\langle 1 \rangle$ *is a succinct unary signature of type* τ_1.

Proof. Suppose $a + (\kappa - 1)b \neq 0$. Consider the gadget in Figure 11.3a. We assign $\langle a, b, c\rangle$ to its vertex. By Lemma 9.2.1, this gadget has the succinct unary signature $\langle u \rangle$ of type τ_1, where $u = a + (\kappa - 1)b$. Since $u \neq 0$, this signature is equivalent to $\langle 1 \rangle$.

Otherwise, $a + (\kappa - 1)b = 0$, and $[2b + (\kappa - 2)c][b^2 - 4bc - (\kappa - 3)c^2] \neq 0$. Consider the gadget in Figure 11.3b. We assign $\langle a, b, c\rangle$ to all three vertices. By Lemma 9.2.1, this gadget has the succinct unary signature $\langle u' \rangle$ of type τ_1, where $u' = -(\kappa - 1)(\kappa - 2)[2b + (\kappa - 2)c][b^2 - 4bc - (\kappa - 3)c^2]$. Since $u' \neq 0$, this signature is equivalent to $\langle 1 \rangle$. □

One of the failure conditions of Lemma 11.5.1 is when both $a + (\kappa - 1)b = 0$ and $b^2 - 4bc - (\kappa - 3)c^2 = 0$ hold. In this case, $\langle a, b, c\rangle = c\langle -(\kappa - 1)(2 \pm \sqrt{\kappa + 1}), 2 \pm \sqrt{\kappa + 1}, 1\rangle$. If $c = 0$, then $a = b = c = 0$ and the signature is trivial. Otherwise, $c \neq 0$. Then up to a nonzero factor of c, this signature further simplifies to AD_3 by taking the minus sign when $\kappa = 3$. Just like AD_3, we show (in Lemma 11.5.3) that all of these signatures are #P-hard.

Similar to the proof of Theorem 10.2.7, we prove the hardness in Lemma 11.5.3 by reducing from counting weighted Eulerian partitions. The succinct signature type τ_4 is a refinement of τ_{color}, so any succinct signature of type τ_{color} can also be expressed as a succinct signature of type τ_4. In particular, the succinct signature $\langle 2, 1, 0, 1, 0\rangle$ of type τ_{color} is written $\langle 2, 0, 1, 0, 0, 0, 1, 0, 0\rangle$ of type τ_4. Then the following is a restatement of Corollary 10.2.6.

Corollary 11.5.2. *Suppose* $\kappa \geq 3$ *is the domain size. Let* $\langle 2, 0, 1, 0, 0, 0, 1, 0, 0\rangle$ *be a succinct quaternary signature of type* τ_4. *Then* $\text{Pl-Holant}_3(\langle 2, 0, 1, 0, 0, 0, 1, 0, 0\rangle)$ *is* #P*-hard.*

Table 11.1: The recurrence matrix M, up to a factor of $(\gamma+1)$, for the recursive construction in the proof of Lemma 11.5.3.

$$\begin{bmatrix}
(\kappa-1)(\gamma-3)\gamma^2 & -2(\kappa-2)(\kappa-1)\gamma & (\kappa-1)(3\gamma-1) & 2(\kappa-2)(\kappa-1)\gamma & 0 & 0 & 0 & 0 & 0 \\
0 & \kappa^2(\gamma+1)-4\kappa\gamma+2(\gamma+1) & 0 & (\kappa-2)(3\gamma-1) & -(\kappa-2)\gamma & -(\kappa-4)(\kappa-2)\gamma & -(\kappa-2)\gamma & -(\kappa-4)(\kappa-2)\gamma & 2(\kappa-2)(\gamma-4)\gamma^2 \\
3\gamma-1 & 2(\kappa-2)\gamma & \kappa^2(\gamma+1)+\kappa(3\gamma-5)-7\gamma+5 & -2(\kappa-2)\gamma & 0 & 0 & 0 & 0 & 0 \\
0 & 2(3\gamma-1) & 0 & (\kappa-2)\gamma(\kappa+\gamma+1) & 2\gamma & 2(\kappa-4)\gamma & 2\gamma & 2(\kappa-4)\gamma & -4(\gamma-4)\gamma^2 \\
0 & -2(\kappa-2)\gamma & 0 & 2(\kappa-2)\gamma & -(\gamma-3)\gamma^2 & 4(\kappa-2)\gamma & 3\gamma-1 & 4(\kappa-2)\gamma & (\kappa-2)(\gamma-4)\gamma(\gamma+1) \\
0 & -(\kappa-4)\gamma & 0 & (\kappa-4)\gamma & 2\gamma & 2(\kappa-4)\gamma & 2\gamma & \kappa(3\gamma+1)-4(\gamma+1) & (\gamma-4)\gamma(\gamma\kappa+\kappa-4) \\
0 & -2(\kappa-2)\gamma & 0 & 2(\kappa-2)\gamma & 3\gamma-1 & 4(\kappa-2)\gamma & -(\gamma-3)\gamma^2 & 4(\kappa-2)\gamma & (\kappa-2)(\gamma-4)\gamma(\gamma+1) \\
0 & -(\kappa-4)\gamma & 0 & (\kappa-4)\gamma & 2\gamma & \kappa(3\gamma+1)-4(\gamma+1) & 2\gamma & 2(\kappa-4)\gamma & (\gamma-4)\gamma(\gamma\kappa+\kappa-4) \\
0 & 4\gamma & 0 & -4\gamma & \gamma+1 & 2(\gamma\kappa+\kappa-4) & \gamma+1 & 2(\gamma\kappa+\kappa-4) & \kappa^2(\gamma+1)-2(\gamma+5)-2(5\gamma-11)
\end{bmatrix}$$

Table 11.2: The matrix P whose rows are the row eigenvectors of the matrix in Table 11.1.

$$\begin{bmatrix}
0 & 0 & 0 & 0 & 1 & -2 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & -1 & 0 & 1 & -(\gamma-3)\gamma & (\gamma-3)\gamma & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\
0 & (\kappa-2)\gamma & 0 & -(\kappa-2)\gamma & 0 & (\kappa-2)(\gamma-1) & 0 & (\kappa-2)(\gamma-1) & (\kappa-2)(\gamma-4)(\gamma-1)\gamma \\
0 & -(\kappa-2)\gamma & 0 & (\kappa-2)\gamma & \gamma-1 & (\kappa-2)(\gamma-1) & \gamma-1 & (\kappa-2)(\gamma-1) & 0 \\
0 & 2 & 0 & \kappa-2 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & \kappa-1 & 0 & 0 & 0 & 0 & 0 & 0 \\
(\gamma-3)\gamma & \kappa^2+\kappa(2\gamma-7)-2(\gamma-5) & -(\gamma-3)\gamma & -\kappa^2-\kappa(2\gamma-7)+2(\gamma-5) & -(\gamma-3)\gamma & -(\kappa-4)(\gamma-3)\gamma & -(\gamma-3)\gamma & -(\kappa-4)(\gamma-3)\gamma & 2(\gamma-4)(\gamma-3)\gamma^2
\end{bmatrix}$$

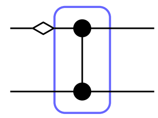


Figure 11.4: Quaternary gadget used in the interpolation construction below. All vertices are assigned $\langle -(\kappa-1)\gamma, \gamma, 1\rangle$.

Lemma 11.5.3. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c \in \mathbb{C}$. Let $\langle a, b, c\rangle$ be a succinct ternary signature of type τ_3. If $a + (\kappa-1)b = 0$ and $b^2 - 4bc - (\kappa-3)c^2 = 0$, then*

$$\langle a, b, c\rangle = c\langle -(\kappa-1)(2+\varepsilon\sqrt{\kappa+1}), 2+\varepsilon\sqrt{\kappa+1}, 1\rangle,$$

where $\varepsilon = \pm 1$, and Pl-Holant$_\kappa(\langle a, b, c\rangle)$ *is #*P*-hard unless $c = 0$, in which case, the problem is computable in polynomial time.*

Proof. If $c = 0$, then $a = b = c = 0$ so the output is always 0. Otherwise, $c \neq 0$. Up to a nonzero factor of c, $\langle a, b, c\rangle$ can be written as $\langle -(\kappa-1)(2+\varepsilon\sqrt{\kappa+1}), 2+\varepsilon\sqrt{\kappa+1}, 1\rangle$ under the given assumptions, where $\varepsilon = \pm 1$.

Suppose $\kappa = 3$. If $\varepsilon = -1$, then we have $\langle 0, 0, 1\rangle = \text{AD}_3$ and we are done by Theorem 10.2.7. Otherwise, $\varepsilon = 1$ and we have $\langle 8, -4, -1\rangle$. Let $T = \frac{1}{3}\left[\begin{smallmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{smallmatrix}\right]$, which is an orthogonal matrix. It follows from Theorem 3.2.2 and Lemma 9.2.6 that

$$\text{Pl-Holant}_3(\langle 8, -4, -1\rangle) \equiv \text{Pl-Holant}_3(T^{\otimes 3}\langle 8, -4, -1\rangle) \equiv \text{Pl-Holant}_3(\langle 0, 0, 1\rangle),$$

so again we are done by Theorem 10.2.7.

Now we suppose $\kappa \geq 4$. Let $g = \langle 2, 0, 1, 0, 0, 0, 1, 0, 0\rangle$ be a succinct quaternary signature of type τ_4. We reduce from Pl-Holant$_\kappa(g)$ to Pl-Holant$_\kappa(\langle a, b, c\rangle)$. Then by Corollary 11.5.2, Pl-Holant$_\kappa(\langle a, b, c\rangle)$ is #P-hard. We write this signature as $\langle -(\kappa-1)\gamma, \gamma, 1\rangle$, where $\gamma = 2+\varepsilon\sqrt{\kappa+1}$.

Consider the gadget in Figure 11.4. We assign $\langle -(\kappa-1)\gamma, \gamma, 1\rangle$ to both vertices. By Lemma 9.2.3, up to a nonzero factor of $\gamma - 1$, this gadget has the succinct quaternary signature f of type τ_4, where

$$f = \big\langle (\kappa-1)(\gamma-3)\gamma^2, \quad -(\kappa-2)\gamma, \quad 3\gamma-1, \quad 2\gamma, \quad 3\gamma-1, \quad 2\gamma, \quad -(\gamma-3)\gamma^2, \quad 2\gamma, \quad \gamma+1\big\rangle.$$

Now consider the recursive construction in Figure 10.3. We assign f to all vertices. Let f_s be the succinct signature of type τ_4 for the sth gadget in this recursive construction. The initial signature, which is just two parallel edges, has the succinct signature $f_0 = \langle 1, 0, 0, 0, 0, 0, 1, 0, 0 \rangle$ of type τ_4. We can express f_s as $M^s f_0$, where M is the matrix in Table 11.1.

Consider an instance Ω of Pl-Holant$_\kappa(g)$. Suppose g appears n times in Ω. We construct from Ω a sequence of instances Ω_s of Pl-Holant$_\kappa(f)$ indexed by $s \ge 0$. We obtain Ω_s from Ω by replacing each occurrence of g with the gadget f_s.

We can express M as $(\gamma - 1)^3 P^{-1} \Lambda P$, where P is the matrix in Table 11.2,

$$\Lambda = \operatorname{diag}(-1, -1, -1, -1, \kappa - 2, \kappa - 2, \kappa - 1, \kappa - 1, \lambda),$$

and $\lambda = \frac{(\kappa-2)(\kappa+2\gamma-4)}{(\gamma-1)^2}$. The rows of P are linearly independent since

$$\det(P) = (\kappa - 1)(\kappa - 2)^2 (\gamma - 1)^6 (\gamma - 3)^3 \gamma \neq 0.$$

For $1 \le i \le 9$, let r_i be the ith row of P. Notice that the initial signature f_0 and the target signature g are orthogonal to the same set of row eigenvectors of M, namely $\{r_1, r_2, r_3, r_5, r_7, r_9\}$. Up to a common factor of $(\gamma - 1)^3$, the eigenvalues for M corresponding to r_4, r_6, and r_8 (the three row eigenvectors of M *not* orthogonal to f_0) are -1, $\kappa - 2$, and $\kappa - 1$ respectively. Since $\kappa \ge 4$, $\kappa - 2$ and $\kappa - 1$ are relatively prime and greater than 1, so these three eigenvalues satisfy the lattice condition. Thus by Lemma 11.3.6, we can interpolate g as desired. □

Remark. Although the matrices in Table 11.1 and Table 11.2 seem large, they are probably the smallest possible to succeed in this recursive quaternary construction. In fact, for quaternary signatures one would normally expect these matrices to be even larger since there are typically fifteen different entries in a domain invariant signature of arity 4.

The other failure condition of Lemma 11.5.1 is when both $a + (\kappa - 1)b = 0$ and $2b + (\kappa - 2)c = 0$ hold. In this case, $\langle a, b, c \rangle = c\langle (\kappa - 1)(\kappa - 2), -(\kappa - 2), 2 \rangle$. If this signature is connected to $\langle 1 \rangle$, then the first entry of the resulting succinct binary signature of type τ_2 is $(\kappa-1)(\kappa-2) \cdot 1 - (\kappa-2) \cdot (\kappa-1) =$

0 while the second entry is $-(\kappa-2)\cdot 2+2\cdot(\kappa-2)=0$. That is, the resulting binary signature is identically zero. This suggests we apply a holographic transformation such that the support of the resulting signature is only on $\kappa-1$ of the domain elements.

If $c=0$, then $a=b=c=0$ and the signature is trivial. Otherwise, $c\neq 0$. If $\kappa=3$, then up to a nonzero factor of c, this signature further simplifies to $\langle 2,-1,2\rangle$, which is tractable by case 3 of Corollary 11.1.3. Otherwise $\kappa\geq 4$, and we show the problem is #P-hard.

Lemma 11.5.4. *Suppose $\kappa\geq 4$ is the domain size. Let $f=\langle(\kappa-1)(\kappa-2),-(\kappa-2),2\rangle$ be a succinct ternary signature of type τ_3. Then* Pl-Holant$_\kappa(f)$ *is* #P*-hard.*

Proof. Consider the matrix $T=\left[\begin{smallmatrix}1 & \mathbf{1}\\ \mathbf{1} & T'\end{smallmatrix}\right]\in\mathbf{GL}_\kappa(\mathbb{C})$, where $T'=yJ_{\kappa-1}+(x-y)I_{\kappa-1}$ with $x=-\frac{\kappa+\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$ and $y=\frac{1}{\sqrt{\kappa}+1}$. After scaling by $\frac{1}{\sqrt{\kappa}}$, we claim that T is an orthogonal matrix.

Let r_i be the ith row of $\frac{1}{\sqrt{\kappa}}T$. First we compute the diagonal entries of $\frac{1}{\kappa}TT^{\intercal}$. Clearly $r_1r_1^{\intercal}=1$. For $2\leq i\leq\kappa$, we have

$$r_ir_i^{\intercal}=\frac{1}{\kappa}\left[1+x^2+(\kappa-2)y^2\right]=\frac{1}{\kappa}\left[1+\frac{(\kappa+\sqrt{\kappa}-1)^2}{(\sqrt{\kappa}+1)^2}+\frac{\kappa-2}{(\sqrt{k}+1)^2}\right]=1.$$

Now we compute the off-diagonal entries. For $2\leq i\leq\kappa$, we have

$$r_1r_i^{\intercal}=\frac{1}{\kappa}\left[1+x+(\kappa-2)y\right]=\frac{1}{\kappa}\left[1-\frac{\kappa+\sqrt{\kappa}-1}{\sqrt{\kappa}+1}+\frac{\kappa-2}{\sqrt{\kappa}+1}\right]=0.$$

For $2\leq i<j\leq\kappa$, we have

$$r_ir_j^{\intercal}=\frac{1}{\kappa}\left[1+2xy+(\kappa-3)y^2\right]=\frac{1}{\kappa}\left[1-\frac{2(\kappa+\sqrt{\kappa}-1)}{(\sqrt{\kappa}+1)^2}+\frac{\kappa-3}{(\sqrt{k}+1)^2}\right]=0.$$

This proves the claim.

We apply a holographic transformation by T to the signature f to obtain $\widehat{f}=T^{\otimes 3}f$, which does not change the complexity of the problem by Lemma 3.2.2. Since the first row of T is a row of all 1's, the output of $\widehat{f}$ on any input containing the first domain element is 0. When restricted to the remaining $\kappa-1$ domain elements, $\widehat{f}$ is domain invariant and symmetric, so it can be expressed as a succinct ternary signature of type τ_3.

Up to a nonzero factor of $\frac{\kappa^3}{(\sqrt{\kappa}+1)^2}$, it can be verified that $\widehat{f} = \langle -(\kappa-2)(2+\sqrt{\kappa}), 2+\sqrt{\kappa}, 1\rangle$. One way to do this is as follows. We write $f = \langle a, b, 2\rangle$ and $T = \left[\begin{smallmatrix} 1 & \mathbf{1} \\ \mathbf{1} & T' \end{smallmatrix}\right] \in \mathbf{GL}_\kappa(\mathbb{C})$, where $T' = yJ_{\kappa-1} + (x-y)I_{\kappa-1}$. The entries of $\widehat{f}$ are polynomials in κ with coefficients from $\mathbb{Z}[a,b,x,y]$. The degree of these polynomials is at most 3 since the arity of f is 3. After computing the entries of $\widehat{f}$ for domain sizes $3 \le \kappa \le 6$ as elements in $\mathbb{Z}[a,b,x,y]$, we interpolate the entries of $\widehat{f}$ as elements in $(\mathbb{Z}[a,b,x,y])[\kappa]$. Then replacing a,b,x,y with their actual values gives the claimed expression for the signature.

Since $\kappa \ge 4$, $\widehat{f}$ is #P-hard by Lemma 11.5.3, which completes the proof. □

At this point, we have achieved the broader goal of this section. For any $a,b,c \in \mathbb{C}$ and domain size $\kappa \ge 3$, either $\operatorname{Pl-Holant}_\kappa(\langle a,b,c\rangle)$ is computable in polynomial time, or $\operatorname{Pl-Holant}_\kappa(\langle a,b,c\rangle)$ is #P-hard, or we can use $\langle a,b,c\rangle$ to construct $\langle 1\rangle$ (i.e. the reduction $\operatorname{Pl-Holant}_\kappa(\{\langle a,b,c,\rangle, \langle 1\rangle\} \le_T \operatorname{Pl-Holant}_\kappa(\langle a,b,c\rangle)$ holds). However, Lemma 11.5.4 is easily generalized and this generalization turns out to be necessary to obtain our dichotomy.

Recall that connecting $f = \langle (\kappa-1)(\kappa-2), -(\kappa-2), 2\rangle$ to $\langle 1\rangle$ results in an identically zero signature. This suggests that we consider the more general signature $\widetilde{f} = \alpha\langle 1\rangle^{\otimes 3} + \beta f$ for any $\alpha \in \mathbb{C}$ and any nonzero $\beta \in \mathbb{C}$ since this does not change the complexity (as we argue in Corollary 11.5.5). For any $a,b,c \in \mathbb{C}$ satisfying $\mathfrak{B} = 0$ (cf. (9.2.2)), if $\alpha = \frac{2b+(\kappa-2)c}{\kappa}$ and $\beta = \frac{-b+c}{\kappa}$, then $\widetilde{f} = \langle a,b,c\rangle$. We note that the condition $\mathfrak{B} = 0$ can also be written as $(\kappa-2)(b-c) = b-a$. We now prove a dichotomy for the signature $\widetilde{f}$.

Corollary 11.5.5. *Suppose $\kappa \ge 3$ is the domain size and $a,b,c \in \mathbb{C}$. Let $\langle a,b,c\rangle$ be a succinct ternary signature of type τ_3. If $\mathfrak{B} = 0$, then* $\operatorname{Pl-Holant}_\kappa(\langle a,b,c\rangle)$ *is #P-hard unless $b = c$ or $\kappa = 3$, in which case, the problem is computable in polynomial time.*

Proof. If $b = c$, then by $\mathfrak{B} = 0$, we have $a = b = c$, which means the signature is degenerate and the problem is trivially tractable. If $\kappa = 3$, then $a = c$ and the problem is tractable by case 3 of Corollary 11.1.3. Otherwise $b \ne c$ and $\kappa \ge 4$.

Since $\mathfrak{B} = 0$, it can be verified that

$$\langle a,b,c\rangle = \frac{2b+(\kappa-2)c}{\kappa}\langle 1\rangle^{\otimes 3} + \frac{-b+c}{\kappa} f, \qquad \text{where} \qquad f = \langle (\kappa-1)(\kappa-2), -(\kappa-2), 2\rangle.$$

We show that Pl-Holant$_\kappa(\langle a, b, c\rangle)$ is #P-hard iff Pl-Holant$_\kappa(f)$ is. Since Pl-Holant$_\kappa(f)$ is #P-hard by Lemma 11.5.4, this proves the result.

Let $G = (V, E)$ be a connected planar 3-regular graph with $n = |V|$ and $m = |E|$. We can view $\operatorname{Holant}_\kappa(G; \langle a, b, c\rangle)$ as a sum of 2^n Holant computations using the signatures $\alpha\langle 1\rangle^{\otimes 3}$ and βf. Each of these Holant computations considers a different assignment of either $\alpha\langle 1\rangle^{\otimes 3}$ or βf to each vertex. Since connecting f to $\langle 1\rangle$ gives an identically zero signature, if any connected signature grid contains both $\alpha\langle 1\rangle^{\otimes 3}$ and βf, then that particular Holant computation is 0. This is because a vertex of degree three assigned $\langle 1\rangle^{\otimes 3}$ is equivalent to three vertices of degree one connected to the same three neighboring vertices and each assigned $\langle 1\rangle$. There are only two possible assignments that could be nonzero. If all vertices are assigned $\alpha\langle 1\rangle^{\otimes 3}$, then the Holant is $\alpha^n \kappa^m$. Otherwise, all vertices are assigned βf and the Holant is $\beta^n \operatorname{Holant}_\kappa(G; f)$. Thus, $\operatorname{Holant}_\kappa(G; \alpha\langle 1\rangle^{\otimes 3} + \beta f) = \alpha^n \kappa^m + \beta^n \operatorname{Holant}_\kappa(G; f)$. Since $\beta \neq 0$, one can solve for either Holant value given the other. □

11.6 Interpolating All Binary Signatures of Type τ_2

In this section, we show how to interpolate all binary succinct signatures of type τ_2 in most settings. We use two general techniques to achieve this goal. In Subsection 11.6.2, we use a generalization of the anti-gadget technique that creates a multitude of gadgets. They are so numerous that one is most likely to succeed. In Subsection 11.6.3, we introduce a new technique called *Eigenvalue Shifted Triples* (EST). These generalize the technique of Eigenvalue Shifted Pairs from [93], and we use EST to interpolate binary succinct signatures in cases where the anti-gadget technique cannot handle. There are a few isolated problems for which neither technique works. However, these problems are easily handled separately in Lemma 11.6.1 in Subsection 11.6.1.

From Section 11.5, every problem fits into one of three cases: either (1) the problem is tractable, (2) the problem is #P-hard, or (3) we can construct the succinct unary signature $\langle 1\rangle$ of type τ_1. Thus, many results in this section assume that $\langle 1\rangle$ is available.

11.6.1 Specific Cases

For some settings of $a, b, c \in \mathbb{C}$, Lemma 11.6.4 and Lemma 11.6.12 do not apply. However, these settings are easily handled on a case-by-case basis.

Lemma 11.6.1. *Suppose $\kappa \geq 3$ is the domain size. Let $\mathcal{F}$ be a signature set containing the succinct unary signature $\langle 1 \rangle$ of type τ_1 and any of the following succinct ternary signatures of type τ_3:*

1. $\langle \kappa - 2 \pm i\kappa\sqrt{2(\kappa-2)}, \kappa - 2, -2 \rangle$;
2. $\langle (\kappa-2)^2 \pm i\kappa\sqrt{\kappa^2-4}, -2(\kappa-2), 4 \rangle$;
3. $\langle -(2\kappa-3)\left[2(\kappa-2) \pm i\kappa\sqrt{2(\kappa-2)}\right], -2(\kappa-3)(\kappa-2) \pm i\kappa\sqrt{2(\kappa-2)}, 4(2\kappa-3) \rangle$ *with* $\kappa \neq 4$;
4. $\langle -\kappa^2 + 2, 2, 2 \rangle$;
5. $\langle \kappa^2 - 6\kappa + 6, -2(\kappa-3), 6 \rangle$;
6. $\langle (\kappa-3)(\kappa-2)^2 \pm i\kappa(2\kappa-3)\sqrt{\kappa^2-4}, -3(\kappa-2)^2 \mp i\kappa\sqrt{\kappa^2-4}, 2(5\kappa-6) \rangle$;
7. $\langle -(\kappa-1)\left[5(\kappa-2) \pm 3i\kappa\sqrt{2(\kappa-2)}\right], -(\kappa-2)(3\kappa-5) \pm i\kappa\sqrt{2(\kappa-2)}, 9\kappa-10 \rangle$;
8. $\langle (\kappa-1)\left[(\kappa-2)(2\kappa+3) \pm 3\kappa\sqrt{\kappa^2-5\kappa+6}\right], (\kappa-3)(\kappa-2) \mp \kappa\sqrt{\kappa^2-5\kappa+6}, -5\kappa+6 \rangle$;
9. $\langle (\kappa-1)\left[(\kappa-2)(2\kappa-7) \pm 3i\kappa\sqrt{\kappa^2-\kappa-2}\right], -(\kappa-2)(5\kappa-7) \mp i\kappa\sqrt{\kappa^2-\kappa-2}, 13\kappa-14 \rangle$;
10. $\langle 1, 0, -2 \rangle$ *with* $\kappa = 3$;
11. $\langle \pm i\sqrt{2}, 0, 1 \rangle$ *with* $\kappa = 3$;
12. $\langle -1 \pm i\sqrt{2}, 0, 1 \rangle$ *with* $\kappa = 3$;
13. $\langle -1 \pm 3i\sqrt{3}, 0, 2 \rangle$ *with* $\kappa = 3$;

Then

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle x, y \rangle\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F})$$

for any $x, y \in \mathbb{C}$, where $\langle x, y \rangle$ is a succinct binary signature of type τ_2.

Proof. In each case, we use the recursive construction in Figure 10.4. We simply state which gadget we use, the signature of that gadget, and the eigenvalues of its associated recurrence matrix (cf. Lemma 10.3.2). Then the result easily follows from Corollary 10.3.3 as the eigenvalues have distinct complex norms.

We use three possible gadgets, which are in Figure 11.6a, Figure 9.7c, and Figure 9.8c. The signatures for the last two gadgets are given by Lemma 9.2.7 and Lemma 9.2.8 respectively.

1. For $\langle \kappa - 2 \pm i\kappa\sqrt{2(\kappa-2)}, \kappa - 2, -2\rangle$, we first use the gadget in Figure 9.7c. Let $\gamma = \pm i\sqrt{2(\kappa-2)}$. Up to a nonzero factor of $\frac{(\gamma-2)^7\gamma^2(\gamma+2)^3}{64}$, the signature of the gadget is $\langle -1, 1\rangle$, which means the eigenvalues are $\kappa - 2$ and -2. If $\kappa \neq 4$, then these eigenvalues have distinct complex norms. Otherwise, $\kappa = 4$ and we use the gadget in Figure 9.8c. Up to a factor of $\pm 65536i$, the signature of this gadget is $\langle 1, -3\rangle$, which means the eigenvalues are -8 and 4.
2. For $\langle (\kappa-2)^2 \pm i\kappa\sqrt{\kappa^2-4}, -2(\kappa-2), 4\rangle$, we first use the gadget in Figure 9.7c. Let $\gamma = \pm i\sqrt{\kappa^2-4}$. Up to a nonzero factor of $-4(\kappa-2)\kappa^3(\kappa^2-4\gamma-8)$, the signature of this gadget is $\langle \kappa^2-6\kappa+4, -2(\kappa-4)\rangle$, which means the eigenvalues are $-(\kappa-2)^2$ and κ^2-4k-4. If $\kappa \geq 5$, then these eigenvalues have opposite signs but cannot be the negative of each other. Thus, they have distinct complex norms. The same conclusion holds for $\kappa = 3$ by direct inspection. Otherwise, $\kappa = 4$ and we use the gadget in Figure 9.8c. Up to a factor of 2097152, the signature of this gadget is $\langle 5, 1\rangle$, which means the eigenvalues are 8 and 4.
3. For $\langle -(2\kappa-3)\left[2(\kappa-2) \pm i\kappa\sqrt{2(\kappa-2)}\right], -2(\kappa-3)(\kappa-2) \pm i\kappa\sqrt{2(\kappa-2)}, 4(2\kappa-3)\rangle$, we have $\kappa \neq 4$. We use the gadget in Figure 9.7c. Let $\gamma = \pm i\sqrt{2(\kappa-2)}$. Up to a nonzero factor of $-4(\kappa-2)\kappa^6(3\kappa-4)(4\kappa^2-28\kappa+41-4\gamma(2\kappa-5))$, the signature of the gadget is $\frac{1}{\kappa}\langle 3\kappa-4, \kappa-4\rangle$, which means the eigenvalues are $\kappa - 2$ and 2.
4. For $\langle -\kappa^2+2, 2, 2\rangle$, we use the gadget in Figure 9.7c. Up to a nonzero factor of $(\kappa-2)\kappa^5$, the signature for this gadget is $\langle \kappa^2+2\kappa-4, -4\rangle$, which means the eigenvalues are $(\kappa-2)\kappa$ and $\kappa(\kappa+2)$.
5. For $\langle \kappa^2-6\kappa+6, -2(\kappa-3), 6\rangle$, we use the gadget in Figure 9.7c. Up to a nonzero factor of $(\kappa-2)\kappa^5$, the signature for this gadget is $\langle \kappa^2+2\kappa-4, -4\rangle$, which means the eigenvalues are $(\kappa-2)\kappa$ and $\kappa(\kappa+2)$.
6. For $\langle (\kappa-3)(\kappa-2)^2 \pm i\kappa(2\kappa-3)\sqrt{\kappa^2-4}, -3(\kappa-2)^2 \mp i\kappa\sqrt{\kappa^2-4}, 2(5\kappa-6)\rangle$, we use the gadget in Figure 9.7c. Let $\gamma = \pm i\sqrt{\kappa^2-4}$. Up to a nonzero factor of $(\gamma-2)^2(\gamma+2)^2(\kappa-2)\kappa[7\kappa^2+60\kappa-164+8\gamma(3\kappa-10)]$, the signature of the gadget is $\langle -\kappa^4+6\kappa^3+4\kappa^2-24\kappa+16, 2(\kappa^3-2\kappa^2-8\kappa+8)\rangle$, which means the eigenvalues are $\lambda_1 = (\kappa-2)\kappa(\kappa^2+2\kappa-4)$ and $\lambda_2 = -\kappa(\kappa+2)(\kappa^2-6\kappa+4)$. For $3 \leq \kappa \leq 5$, one can directly check that these eigenvalues have distinct complex norms. For $\kappa \geq 6$, we have $\lambda_2 < 0$, so these eigenvalues have the same complex norm preciously when

$\lambda_1 = -\lambda_2$. However, $\lambda_1 + \lambda_2 = 4\kappa^3 \neq 0$, so the eigenvalues have distinct complex norms.

7. For $\langle -(\kappa-1)\left[5(\kappa-2) \pm 3i\kappa\sqrt{2(\kappa-2)}\right], -(\kappa-2)(3\kappa-5) \pm i\kappa\sqrt{2(\kappa-2)}, 9\kappa-10\rangle$, we first use the gadget in Figure 9.7c. Let $\gamma = \pm i\sqrt{2(\kappa-2)}$. Up to a nonzero factor of $-(\kappa-2)(\kappa-1)\kappa^5[81\kappa^2 - 756\kappa + 1252 - 24(9\kappa-26)\gamma]$, the signature of this gadget is $\langle 5\kappa-6, \kappa-6\rangle$, which means the eigenvalues are $\kappa-2$ and 4. If $\kappa \neq 6$, then these eigenvalues have distinct complex norms. Otherwise, $\kappa = 6$ and we use the gadget in Figure 9.8c. Up to a factor of $-17199267840(1169 \pm 450i\sqrt{2})$, the signature of this gadget is $\langle 7, 13\rangle$, which means the eigenvalues are 72 and -6.
8. For $\langle (\kappa-1)\left[(\kappa-2)(2\kappa+3) \pm 3\kappa\sqrt{\kappa^2-5\kappa+6}\right], (\kappa-3)(\kappa-2) \mp \kappa\sqrt{\kappa^2-5\kappa+6}, -5\kappa+6\rangle$, we first use the gadget in Figure 9.7c. Let $\gamma = \pm\sqrt{\kappa^2-5\kappa+6}$. Up to a factor of $(\kappa-2)(\kappa-1)\kappa^5[313\kappa^2 - 1500\kappa + 1764 - 24(13\kappa-30)\gamma]$, the signature of this gadget is $\langle \kappa^3 - 3\kappa^2 + 3, -\kappa+3\rangle$, which means the eigenvalues are $\lambda_1 = (\kappa-2)^2\kappa$ and $\lambda_2 = \kappa(\kappa^2 - 3\kappa + 1)$. If $\kappa \geq 4$, these eigenvalues are positive, so they have the same complex norm preciously when $\lambda_1 = \lambda_2$. However, $\lambda_1 - \lambda_2 = -(\kappa-3)\kappa \neq 0$, so the eigenvalues have distinct complex norms. Otherwise, $\kappa = 3$ and we use the gadget in Figure 9.8c. Up to a factor of 9565938, the signature of this gadget is $\langle 5, 2\rangle$, which means the eigenvalues are 9 and 3.
9. For $\langle (\kappa-1)\left[(\kappa-2)(2\kappa-7) \pm 3i\kappa\sqrt{\kappa^2-\kappa-2}\right], -(\kappa-2)(5\kappa-7) \mp i\kappa\sqrt{\kappa^2-\kappa-2}, 13\kappa-14\rangle$, we use the gadget in Figure 9.7c. Let $\gamma = \pm i\sqrt{\kappa^2-\kappa-2}$. Up to a nonzero factor of $(\kappa-2)(\kappa-1)\kappa^5[119\kappa^2 + 76\kappa - 772 + 24(5\kappa-22)\gamma]$, the signature of this gadget is $\langle -\kappa^3 + 7\kappa^2 - 4\kappa - 3, 2\kappa^2 - 7\kappa - 3\rangle$, which means the eigenvalues are $\lambda_1 = (\kappa-2)\kappa^2$ and $\lambda_2 = -\kappa(k^2 - 5\kappa - 3)$. For $3 \leq \kappa \leq 5$, one can directly check that these eigenvalues have distinct complex norms. For $\kappa \geq 6$, we have $\lambda_2 < 0$, so these eigenvalues have the same complex norm preciously when $\lambda_1 = -\lambda_2$. However, $\lambda_1 + \lambda_2 = 3\kappa(\kappa+1) \neq 0$, so the eigenvalues have distinct complex norms.
10. For $\langle 1, 0, -2\rangle$ with $\kappa = 3$, we use the gadget in Figure 9.7c. Up to a factor of 3, the signature of this gadget is $\langle 11, -4\rangle$, which means the eigenvalues are 3 and 15.
11. For $\langle \pm i\sqrt{2}, 0, 1\rangle$ with $\kappa = 3$, we use the gadget in Figure 11.6a. The signature of this gadget is $\langle \pm i\sqrt{2}, 1\rangle$, which means the eigenvalues are $2 \pm i\sqrt{2}$ and $-1 \pm i\sqrt{2}$.
12. For $\langle -1 \pm i\sqrt{2}, 0, 1\rangle$ with $\kappa = 3$, we use the gadget in Figure 11.6a. The signature of this

gadget is $\langle -1 \pm i\sqrt{2}, 1\rangle$, which means the eigenvalues are $1 \pm i\sqrt{2}$ and $-2 \pm i\sqrt{2}$.

13. For $\langle -1 \pm 3i\sqrt{3}, 0, 2\rangle$ with $\kappa = 3$, we use the gadget in Figure 9.7c. Up to a factor of 72, the signature of this gadget is $\frac{1}{3}\langle 25 \pm 13\sqrt{3}, -5 \pm i\sqrt{3}\rangle$, which means the eigenvalues are $5(1 \pm i\sqrt{3})$ and $2(5 \pm 2\sqrt{3})$. □

11.6.2 E Pluribus Unum

In this subsection, we use Lemma 10.1.3 to prove our interpolation results. The main technical difficulty is to satisfy the third condition of Lemma 10.1.3, which is to prove that some recurrence matrix (that defines a sequence of gadgets) has infinite order up to a scalar. When the matrix has a finite order up to a scalar, we can utilize this failure condition to our advantage by constructing an anti-gadget (cf. Chapter 5), which is the "last" gadget with a distinct signature (up to a scalar) in the infinite sequence of gadgets. To make sure that we construct a multitude of nontrivial gadgets without cancellation, we put the anti-gadget inside another gadget (contrast the gadget in Figure 11.5 with the gadget in Figure 11.6b). From among this plethora of gadgets, at least one must succeed under the right conditions.

Although this idea works quite well in that some gadget among those constructed does succeed, we still must prove that one such gadget succeeds in every setting. We aim to exhibit a recurrence matrix whose ratio of eigenvalues is not a root of unity. We consider three related recurrence matrices at once. The next two lemmas consider two similar situations involving the eigenvalues of three such matrices. When applied, these lemmas show that some recurrence matrix must have eigenvalues with distinct complex norms, even though exactly which one among them succeeds may depend on the parameters in a complicated way.

Lemma 11.6.2. *Let $d_0, d_1, d_2, \Psi \in \mathbb{C}$. If d_0, d_1, and d_2 have the same argument but are distinct, then for all $\rho \in \mathbb{R}$, there exists $i \in \{0, 1, 2\}$ such that $|\Psi + d_i| \neq \rho$.*

Proof. Assume to the contrary that there exists $\rho \in \mathbb{R}$ such that $|\Psi + d_i| = \rho$ for every $i \in \{0, 1, 2\}$. In the complex plane, consider the circle centered at the origin of radius ρ. Each $\Psi + d_i$ is a distinct point on this circle as well as a distinct point on a common line through Ψ. However, the line intersects the circle in at most two points, a contradiction. □

Lemma 11.6.3. *Let $d_0, d_1, d_2, \Psi \in \mathbb{C}$. If d_0, d_1, and d_2 have the same complex norm but are distinct and $\Psi \neq 0$, then for all $\rho \in \mathbb{R}$, there exists $i \in \{0, 1, 2\}$ such that $|\Psi + d_i| \neq \rho$.*

Proof. Let $\ell = |d_0|$. Assume to the contrary that there exists $\rho \in \mathbb{R}$ such that $|\Psi + d_i| = \rho$ for every $i \in \{0, 1, 2\}$. In the complex plane, consider the circle centered at the origin of radius ρ and the circle centered at Ψ of radius ℓ. Since $\Psi \neq 0$, these circles are distinct. Each $\Psi + d_i$ is a distinct point on both circles. However, these circles intersect in at most two points, a contradiction. □

Now we use Lemma 11.6.2 and Lemma 11.6.3 as well as our generalization of the anti-gadget technique to establish a crucial lemma.

Lemma 11.6.4. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c, \omega \in \mathbb{C}$. Let $\mathcal{F}$ be a set of signatures containing the succinct binary signature $\langle \omega + \kappa - 1, \omega - 1 \rangle$ of type τ_2 and the succinct ternary signature $\langle a, b, c \rangle$ of type τ_3. If the following three conditions are satisfied:*

1. *$\omega \notin \{0, \pm 1\}$,*
2. *$\mathfrak{B} \neq 0$, and*
3. *at least one of the following holds:*
   - *(i) $\mathfrak{C} = 0$ or*
   - *(ii) $\mathfrak{C}^2 = \omega^{2\ell} \mathfrak{B}^2$ for some $\ell \in \{0, 1\}$ but either $\mathfrak{C}^2 \neq \mathfrak{A}^2$ or $\kappa \neq 3$,*

then

$$\operatorname{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle x, y \rangle\}) \leq_T \operatorname{Pl-Holant}_\kappa(\mathcal{F})$$

for any $x, y \in \mathbb{C}$, where $\langle x, y \rangle$ is a succinct binary signature of type τ_2.

We use this lemma to establish that various 2-by-2 recurrence matrices have infinite order modulo scalars. When applied, ω will be the ratio of two eigenvalues, one of which is a multiple of $\mathfrak{B}$ or $\mathfrak{B}^2$ by a nonzero function of κ.

Proof of Lemma 11.6.4. Let $\Phi = \frac{\mathfrak{C}^2}{\mathfrak{B}^2}$ and $\Psi = \frac{(\kappa-2)\mathfrak{A}^2}{\mathfrak{B}^2}$. Consider the recursive construction in Figure 10.4. After scaling by a nonzero factor of κ, we assign $f = \frac{1}{\kappa}\langle \omega + \kappa - 1, \omega - 1 \rangle$ to every vertex. Let f_s be the succinct binary signature of type τ_2 for the sth gadget in this construction. We can express f_s as $M^s \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$, where $M = \frac{1}{\kappa} \begin{bmatrix} \omega+\kappa-1 & (\kappa-1)(\omega-1) \\ \omega-1 & (\kappa-1)\omega+1 \end{bmatrix} = \left[\begin{smallmatrix} 1 & 1-\kappa \\ 1 & 1 \end{smallmatrix}\right] \left[\begin{smallmatrix} \omega & 0 \\ 0 & 1 \end{smallmatrix}\right] \left[\begin{smallmatrix} 1 & 1-\kappa \\ 1 & 1 \end{smallmatrix}\right]^{-1}$ by Lemma 10.3.2.

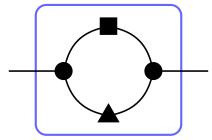


Figure 11.5: Binary gadget that generalizes the anti-gadget technique. The circle vertices are assigned $\langle a, b, c\rangle$ while the square and triangle vertices are each assigned the signature of some gadget.

Then $f_s = \frac{1}{\kappa}\langle \omega^s + \kappa - 1, \omega^s - 1\rangle$. The eigenvalues of M are 1 and ω, so the determinant of M is $\omega \neq 0$. If ω is not a root of unity, then we are done by Corollary 10.3.3.

Otherwise, suppose ω is a primitive root of unity of order n. Since $\omega \neq \pm 1$ by assumption, $n \geq 3$. Now consider the gadget in Figure 11.5. We assign $\langle a, b, c\rangle$ to the circle vertices, $f_r = \frac{1}{\kappa}\langle \omega^r + \kappa - 1, \omega^r - 1\rangle$ to the square vertex, and $f_s = \frac{1}{\kappa}\langle \omega^s + \kappa - 1, \omega^s - 1\rangle$ to the triangle vertex, where $r, s \geq 0$ are parameters of our choice. By Lemma 9.2.5, up to a nonzero factor of $\frac{\mathfrak{B}^2}{\kappa}$, this gadget has the succinct binary signature

$$f(r, s) = \tfrac{1}{\kappa}\langle \Phi\omega^{r+s} + (\kappa - 1)(\omega^r + \omega^s + \Psi + 1),\ \Phi\omega^{r+s} - (\omega^r + \omega^s + \Psi + 1) + \kappa\rangle$$

of type τ_2. Consider using this gadget in the recursive construction of Figure 10.4. Let $f_t(r, s)$ be the succinct binary signature of type τ_2 for the tth gadget in this recursive construction. Then $f_1(r, s) = f(r, s)$ and $f_t(r, s) = (M(r, s))^t \left[\begin{smallmatrix}1\\0\end{smallmatrix}\right]$, where the eigenvalues of $M(r, s)$ are $\Phi\omega^{r+s} + \kappa - 1$ and $\omega^r + \omega^s + \Psi$ by Lemma 10.3.2. Thus, the determinant of $M(r, s)$ is $(\Phi\omega^{r+s} + \kappa - 1)(\omega^r + \omega^s + \Psi)$. Since Φ is either 0 or a power of ω by condition 3, the first factor is nonzero for any choice of r and s. However, for some r and s, it might be that $g(r, s) = \omega^r + \omega^s + \Psi = 0$.

Suppose $\Psi = 0$. We consider the two possible cases of Φ in order to finish the proof under this assumption.

1. Suppose $\Phi = 0$. Consider the gadget $M(0, 1)$. The determinant of $M(0, 1)$ is nonzero since $g(0, 1) \neq 0$ and the ratio of its eigenvalues is not a root of unity because they have distinct complex norms. Thus, we are done by Corollary 10.3.3.
2. Suppose $\Phi = \omega^{2\ell}$ for some $\ell \in \{0, 1\}$. Consider the gadget $M(n-\ell, n-\ell)$. The determinant of $M(n-\ell, n-\ell)$ is nonzero since $g(n-\ell, n-\ell) \neq 0$ and the ratio of its eigenvalues is not a root

of unity because they have distinct complex norms. Thus, we are done by Corollary 10.3.3.

Otherwise, $\Psi \neq 0$. We claim that $g(r,s) = 0$ can hold for at most one choice of $r, s \in \mathbb{Z}_n$ (modulo the swapping of r and s). To see this, consider r_1, s_1, r_2, s_2 such that $g(r_1, s_1) = 0 = g(r_2, s_2)$. Then $\omega^{r_1} + \omega^{s_1} = -\Psi = \omega^{r_2} + \omega^{s_2}$. By taking complex norms and applying the law of cosines, we have $\cos\theta_1 = \cos\theta_2$, where $\theta_j = \arg(\omega^{s_j - r_j})$ is the angle from ω^{r_j} to ω^{s_j} for $j \in \{1, 2\}$. Thus, $\theta_1 = \pm\theta_2$. Since $\Psi \neq 0$, we have $\theta_1 \neq \pm\pi$. If $\theta_1 = \theta_2$, then $\omega^{r_1}(1 + e^{i\theta_1}) = \omega^{r_2}(1 + e^{i\theta_1})$. Since $\theta_1 \neq \pm\pi$, the factor $1 + e^{i\theta_1}$ is nonzero. After dividing by this factor, we conclude that $r_1 = r_2$ and thus $s_1 = s_2$. Otherwise, $\theta_1 = -\theta_2$. Then $\omega^{r_1}(1 + e^{i\theta_1}) = \omega^{s_2}(1 + e^{i\theta_1})$ and we conclude that $r_1 = s_2$ and $s_1 = r_2$. This proves the claim.

Suppose $n \geq 4$ and let $S_0 = \{(0,0), (1, n-1), (2, n-2)\}$ and $S_1 = \{(1,1), (2,0), (3, n-1)\}$. Then $g(r,s) = 0$ holds for at most one $(r,s) \in S_0 \cup S_1$. In particular, $g(r,s)$ is either nonzero for all $(r,s) \in S_0$ or nonzero for all $(r,s) \in S_1$. Pick $j \in \{0,1\}$ such that $g(r,s)$ is nonzero for all $(r,s) \in S_j$. By Lemma 11.6.2 with $d_i = (\omega^i + \omega^{-i})\omega^j$ and $\rho = |\Phi\omega^{2j} + \kappa - 1|$, there exists some $(r,s) \in S_j$ such that the eigenvalues of $M(r,s)$ have distinct complex norms, so we are done by Corollary 10.3.3.

Otherwise, $n = 3$. We consider the two possible cases of Φ in order to finish the proof.

1. Suppose $\Phi = 0$. Let $S_j = \{(0,j), (1, j+1), (2, j+2)\}$. Then $g(r,s) = 0$ holds for at most one $(r,s) \in S_0 \cup S_1$. In particular, $g(r,s)$ is either nonzero for all $(r,s) \in S_0$ or nonzero for all $(r,s) \in S_1$. Pick $j \in \{0,1\}$ such that $g(r,s)$ is nonzero for all $(r,s) \in S_j$. By Lemma 11.6.3 with $d_i = (1 + \omega^j)\omega^i$ and $\rho = \kappa - 1$, there exists some $(r,s) \in S_j$ such that the eigenvalues of $M(r,s)$ have distinct complex norms, so we are done by Corollary 10.3.3.

2. Suppose $\Phi = \omega^{2\ell}$ for some $\ell \in \{0,1\}$ but either $\mathfrak{C}^2 \neq \mathfrak{A}^2$ or $\kappa \neq 3$. Note that this is equivalent to $\Phi \neq \Psi$ or $\kappa \neq 3$. Consider the set $S = \{(0,0), (0,1), (0,2), (1,1), (1,2), (2,2)\}$. If there exists some $(r,s) \in S$ such that $g(r,s) \neq 0$ and the eigenvalues of $M(r,s)$ have distinct complex norms, then we are done by Corollary 10.3.3.

   Otherwise, for every $(r,s) \in S$, either $g(r,s) = 0$ or the eigenvalues of $M(r,s)$ have the same complex norm. If the latter condition were to always hold, then we would have

$$\begin{aligned} |2+\Psi| &= \left|\omega^{2\ell}+\kappa-1\right| &&= |-1+\Psi|\,, \\ \left|2\omega^2+\Psi\right| &= \left|\omega^{2\ell+1}+\kappa-1\right| &&= \left|-\omega^2+\Psi\right|\,\text{, and} \\ |2\omega+\Psi| &= \left|\omega^{2\ell+2}+\kappa-1\right| &&= |-\omega+\Psi|\,, \end{aligned}$$

where each equality corresponds to one of the six $M(r,s)$ having eigenvalues of equal complex norm for $(r,s)\in S$. Of the six equalities, at most one may not hold since $g(r,s)=0$ for at most one $(r,s)\in S$. Since $n=3$, two of the three terms of the form $|\omega^{2\ell+m}+\kappa-1|$ must be equal, so we can write the stronger condition

$$\begin{array}{c} \left|2\omega^2+\Psi\omega^\ell\right| = |\omega+\kappa-1| = \left|-\omega^2+\Psi\omega^\ell\right| \\ \| \\ \left|2\omega+\Psi\omega^\ell\right| = \left|\omega^2+\kappa-1\right| = \left|-\omega+\Psi\omega^\ell\right|. \end{array} \tag{11.6.4}$$

As it is, one of the horizontal equalities in (11.6.4) may not hold. However, even without one of these equalities, we can still reach a contradiction.

We show that $\Psi\omega^\ell\in\mathbb{R}$ even if one of the equalities in (11.6.4) does not hold. In fact, either the left or the right half of the equalities in (11.6.4) hold. In the first case, $|2\omega^2+\Psi\omega^\ell|=|2\omega+\Psi\omega^\ell|$ holds and we get $\Psi\omega^\ell\in\mathbb{R}$. Similarly in the second case, $|-\omega^2+\Psi\omega^\ell|=|-\omega+\Psi\omega^\ell|$ holds and we get $\Psi\omega^\ell\in\mathbb{R}$ as well. Next, we use real and imaginary parts to calculate the complex norms even if one of the equalities in (11.6.4) does not hold. Either the top half of the equalities hold and thus $|2\omega^2+\Psi\omega^\ell|=|-\omega^2+\Psi\omega^\ell|$, or the bottom half of the equalities hold and thus $|2\omega+\Psi\omega^\ell|=|-\omega+\Psi\omega^\ell|$. In any case, it readily follows that $\Psi\omega^\ell=1$. This implies $\Psi=\omega^{2\ell}$,

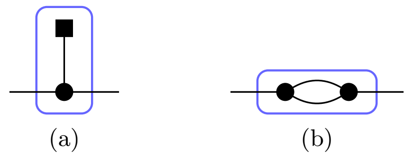


Figure 11.6: Binary gadgets used to interpolate any succinct binary signature of type τ_2. The circle vertices are assigned $\langle a, b, c\rangle$ and the square vertex is assigned $\langle 1\rangle$.

so we can rewrite (11.6.4) as

$$\begin{array}{c} \sqrt{3} = |\omega + \kappa - 1| = \sqrt{3} \\ \| \\ \sqrt{3} = \left|\omega^2 + \kappa - 1\right| = \sqrt{3}, \end{array}$$

where at most one equation may not hold. This forces $\kappa = 3$. However, $\Phi = \omega^{2\ell} = \Psi$ and $\kappa = 3$ is a contradiction. □

The previous lemma is strong enough to handle the typical case.

Lemma 11.6.5. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle a, b, c\rangle$ of type τ_3 and the succinct unary signature $\langle 1\rangle$ of type τ_1. If*

1. *$\mathfrak{B} \neq 0$,*
2. *$\mathfrak{C} \neq 0$,*
3. *$\mathfrak{C}^2 \neq \mathfrak{B}^2$, and*
4. *either $\mathfrak{C}^2 \neq \mathfrak{A}^2$ or $\kappa \neq 3$,*

then

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle x, y\rangle\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F})$$

for any $x, y \in \mathbb{C}$, where $\langle x, y\rangle$ is a succinct binary signature of type τ_2.

Proof. Let $\omega = \frac{\mathfrak{C}}{\mathfrak{B}}$, which is well-defined. Consider the gadget in Figure 11.6a. We assign $\langle a, b, c\rangle$ to the circle vertex and $\langle 1\rangle$ to the square vertex. Up to a nonzero factor of $\frac{\mathfrak{B}}{\kappa}$, this gadget has the succinct binary signature

$$\frac{\kappa}{\mathfrak{B}}\langle a + (\kappa - 1)b, 2b + (\kappa - 2)c\rangle = \langle \omega + \kappa - 1, \omega - 1\rangle$$

of type τ_2. Then we are done by Lemma 11.6.4 with $\ell = 1$ in case (ii) of condition 3. □

If $\mathfrak{B} = 0$, then we already know the complexity by Corollary 11.5.5. The other failure conditions from the previous lemma are:

$$\mathfrak{C} - \mathfrak{B} = \kappa[2b + (\kappa - 2)c] = 0; \tag{11.6.5}$$

$$\mathfrak{C} + \mathfrak{B} = 2a + 2(2\kappa - 3)b + (\kappa - 2)^2 c = 0; \tag{11.6.6}$$

$$\mathfrak{C} = 0; \tag{11.6.7}$$

$$\kappa = 3 \text{ and } \mathfrak{C} - \mathfrak{A} = 0, \qquad \text{or equivalently} \qquad \kappa = 3 \text{ and } b = 0; \tag{11.6.8}$$

$$\kappa = 3 \text{ and } \mathfrak{C} + \mathfrak{A} = 0, \qquad \text{or equivalently} \qquad \kappa = 3 \text{ and } 2a + 3b + 4c = 0. \tag{11.6.9}$$

Notice that these five failure conditions are *linear* in a, b, c.

By starting the proof with a different gadget, Lemma 11.6.4 can handle the first three failure conditions. The last two failure conditions require a new idea, Eigenvalue Shifted Triples, which we introduce in Subsection 11.6.3. In fact, these two cases are equivalent under an orthogonal holographic transformation.

The next lemma considers the failure condition in (11.6.5). Note that $\mathfrak{C} = \mathfrak{B}$ iff the signature can be written as $\langle 2a, -(\kappa - 2)c, 2c\rangle$ up to a factor of 2. The first excluded case in Lemma 11.6.6 is handled by Corollary 11.5.5 and the last two excluded cases are tractable by Corollary 11.1.4.

Lemma 11.6.6. *Suppose $\kappa \geq 3$ is the domain size and $a, c \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle 2a, -(\kappa - 2)c, 2c\rangle$ of type τ_3 and the succinct unary signature $\langle 1\rangle$ of type τ_1. If*

1. $2a \neq (\kappa - 1)(\kappa - 2)c$,
2. $4a \neq (\kappa^2 - 6\kappa + 4)c$, *and*
3. $c \neq 0$,

then

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle x, y\rangle\}) \le_T \text{Pl-Holant}_\kappa(\mathcal{F})$$

for any $x, y \in \mathbb{C}$, where $\langle x, y\rangle$ is a succinct binary signature of type τ_2.

Proof. Note that when $2b = -(\kappa-2)c$, we have $\mathfrak{B} = \mathfrak{C} = 2a - (\kappa-1)(\kappa-2)c$ by (11.6.5), which is nonzero by condition 1 of the lemma. Let $\omega_0 = 4a^2 + (\kappa-2)[4ac + (2\kappa^2+\kappa-2)c^2]$ and assume $\omega_0 \neq 0$. Then let $\omega = \frac{\mathfrak{B}^2}{\omega_0} \neq 0$. By conditions 2 and 3, it follows that $\omega \neq 1$. Also we note that when $2b = -(\kappa-2)c$, we have $2\mathfrak{A} = 2a + (3\kappa-2)c$ and $2\mathfrak{C} = 2a - (\kappa-1)(\kappa-2)c$. By the same conditions, 2 and 3, we have $\mathfrak{C}^2 \neq \mathfrak{A}^2$. We further assume that $\omega \neq -1$, which is equivalent to $8a^2 - 4(\kappa-2)^2ac + (\kappa-2)(\kappa^3 - 2\kappa^2 + 6\kappa - 4)c^2 \neq 0$.

Consider the gadget in Figure 11.6b. We assign $\langle 2a, -(\kappa-2)c, 2c\rangle$ to the vertices. Up to a nonzero factor of $\frac{\omega_0}{\kappa}$, this gadget has the succinct binary signature

$$\frac{\kappa}{\omega_0}\langle 4a^2 + (\kappa-1)(\kappa-2)(3\kappa-2)c^2, \quad -(\kappa-2)[4ac - (\kappa^2-6\kappa+4)c^2]\rangle = \langle \omega + \kappa - 1, \quad \omega - 1\rangle$$

of type τ_2. Then we are done by Lemma 11.6.4 with $\ell = 0$ in case (ii) of condition 3.

Now we deal with the following exceptional cases.

1. If $\omega_0 = 0$, then $2a = -\left[\kappa - 2 \pm i\kappa\sqrt{2(\kappa-2)}\right]c$. Up to a nonzero factor of $-c$, we have $-\frac{1}{c}\langle 2a, -(\kappa-2)c, 2c\rangle = \langle \kappa-2\pm i\kappa\sqrt{2(\kappa-2)}, \kappa-2, -2\rangle$ and are done by case 1 of Lemma 11.6.1.
2. If $8a^2 - 4(\kappa-2)^2ac + (\kappa-2)(\kappa^3 - 2\kappa^2 + 6\kappa - 4)c^2 = 0$, then $4a = \left[(\kappa-2)^2 \pm i\kappa\sqrt{\kappa^2-4}\right]c$. Up to a nonzero factor of $\frac{c}{2}$, we have

$$\frac{2}{c}\langle 2a, -(\kappa-2)c, 2c\rangle = \langle (\kappa-2)^2 \pm i\kappa\sqrt{\kappa^2-4}, -2(\kappa-2), 4\rangle$$

and are done by case 2 of Lemma 11.6.1. □

The next lemma considers the failure condition in (11.6.6). Note that $\mathfrak{C} = -\mathfrak{B}$ iff the signature can be written as $\langle -2(2\kappa-3)b - (\kappa-2)^2c, 2b, 2c\rangle$ up to a factor of 2. The first excluded case in Lemma 11.6.7 is handled by Corollary 11.5.5 and the last excluded case is tractable by Corollary 11.1.7.

Lemma 11.6.7. *Suppose $\kappa \ge 3$ is the domain size and $a, b \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle -2(2\kappa-3)b-(\kappa-2)^2c, 2b, 2c\rangle$ of type τ_3 and the succinct unary signature $\langle 1\rangle$ of type τ_1. If*

1. $2b \neq -(\kappa-2)c$ *and*
2. $\kappa \neq 4$ *or* $5b^2+2bc+c^2 \neq 0$,

then

$$\text{Pl-Holant}_\kappa(\mathcal{F}\cup\{\langle x,y\rangle\}) \le_T \text{Pl-Holant}_\kappa(\mathcal{F})$$

for any $x, y \in \mathbb{C}$, where $\langle x,y\rangle$ is a succinct binary signature of type τ_2.

Proof. Note that when $2a = -2(2\kappa-3)b-(\kappa-2)^2c$, we have $\mathfrak{B} = -\mathfrak{C}$ by (11.6.6) and $2\mathfrak{B} = -\kappa[2b+(\kappa-2)c]$, which is nonzero by condition 1 of the lemma. Let $\omega_0 = 8(2\kappa-3)b^2+(\kappa-2)\left[8(\kappa-3)bc+(\kappa^2-6\kappa+12)c^2\right]$ and assume $\omega_0 \neq 0$. Then let $\omega = \frac{\kappa[2b+(\kappa-2)c]^2}{\omega_0}$. By condition 1, $\omega \neq 0$. It can be shown that $\kappa[2b+(\kappa-2)c]^2 = \omega_0$ is equivalent to $(b-c)[3b+(\kappa-3)c] = 0$. Thus, assume $b \neq c$ and $3b \neq -(\kappa-3)c$. Then $\omega \neq 1$. Also we note that when $2a = -2(2\kappa-3)b-(\kappa-2)^2c$, we have $2\mathfrak{A} = -\kappa[4b+(\kappa-4)c]$ and $2\mathfrak{C} = \kappa[2b+(\kappa-2)c]$. By the same assumptions, $b \neq c$ and $3b \neq -(\kappa-3)c$, we have $\mathfrak{C}^2 \neq \mathfrak{A}^2$. Further assume that $\omega \neq -1$, which is equivalent to $2(5\kappa-6)b^2+(\kappa-2)[6(\kappa-2)bc+(\kappa^2-4\kappa+6)c^2] \neq 0$.

Consider the gadget in Figure 11.6b. We assign $\langle -2(2\kappa-3)b-(\kappa-2)^2c, 2b, 2c\rangle$ to the vertices. Up to a nonzero factor of $\frac{\omega_0}{4}$, this gadget has the succinct binary signature $\frac{1}{\omega_0}\langle x,y\rangle = \langle \omega+\kappa-1, \omega-1\rangle$ of type τ_2, where

$$x = 4(4\kappa^2-9\kappa+6)b^2+(\kappa-2)\left[4(\kappa-2)(2\kappa-3)bc+(\kappa^3-6\kappa^2+16\kappa-12)c^2\right] \quad \text{and}$$
$$y = -4(\kappa-2)\left[3b^3+(\kappa-6)bc-(\kappa-3)c^2\right].$$

Then we are done by Lemma 11.6.4 with $\ell = 0$ in case (ii) of condition 3.

Now we deal with the following exceptional cases.

1. If $\omega_0 = 0$, then we have $-4(2\kappa-3)b = \left[2(\kappa-3)(\kappa-2) \pm i\kappa\sqrt{2(\kappa-2)}\right]c$ but $\kappa \neq 4$ by

condition 2 since otherwise $\omega_0 = 8(5b^2 + 2bc + c^2) \neq 0$. Up to a nonzero factor of $\frac{c}{2(2\kappa-3)}$,

$$\begin{aligned}&\frac{2(2\kappa-3)}{c}\langle -2(2\kappa-3)b-(\kappa-2)^2c, 2b, 2c\rangle \\ &= \langle -(2\kappa-3)\big[2(\kappa-2) \mp i\kappa\sqrt{2(\kappa-2)}\big], \quad -2(\kappa-3)(\kappa-2) \mp i\kappa\sqrt{2(\kappa-2)}, \quad 4(2\kappa-3)\rangle\end{aligned}$$

and are done by case 3 of Lemma 11.6.1.

2. If $b = c$, then up to a nonzero factor of c, we have $\frac{1}{c}\langle -2(2\kappa-3)b-(\kappa-2)^2c, 2b, 2c\rangle = \langle -\kappa^2+2, 2, 2\rangle$ and are done by case 4 Lemma 11.6.1.
3. If $3b = -(\kappa-3)c$, then up to a nonzero factor of $\frac{c}{3}$, we have $\frac{3}{c}\langle -2(2\kappa-3)b-(\kappa-2)^2c, 2b, 2c\rangle = \langle \kappa^2-6\kappa+6, -2(\kappa-3), 6\rangle$ and are done by case 5 of Lemma 11.6.1.
4. If $2(5\kappa-6)b^2 + (\kappa-2)[6(\kappa-2)bc + (\kappa^2-4\kappa+6)c^2] = 0$, then $-2(5\kappa-6)b = \big[3(\kappa-2)^2 \pm i\kappa\sqrt{\kappa^2-4}\big]c$. Up to a nonzero factor of $\frac{c}{5\kappa-6}$,

$$\begin{aligned}&\frac{5\kappa-6}{c}\langle -2(2\kappa-3)b-(\kappa-2)^2c, 2b, 2c\rangle \\ &\quad= \langle (\kappa-3)(\kappa-2)^2 \pm i\kappa(2\kappa-3)\sqrt{\kappa^2-4}, \quad -3(\kappa-2)^2 \mp i\kappa\sqrt{\kappa^2-4}, \quad 2(5\kappa-6)\rangle\end{aligned}$$

and are done by case 6 of Lemma 11.6.1. □

The next lemma considers the failure condition in (11.6.7). Note that $\mathfrak{C} = 0$ iff the signature can be written as $\langle -3(\kappa-1)b-(\kappa-1)(\kappa-2)c, b, c\rangle$. The excluded case in Lemma 11.6.8 is handled by Corollary 11.5.5.

Lemma 11.6.8. *Suppose $\kappa \geq 3$ is the domain size and $b, c \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle -3(\kappa-1)b-(\kappa-1)(\kappa-2)c, b, c\rangle$ of type τ_3 and the succinct unary signature $\langle 1\rangle$ of type τ_1. If $2b \neq -(\kappa-2)c$, then*

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle x, y\rangle\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F})$$

for any $x, y \in \mathbb{C}$, where $\langle x, y\rangle$ is a succinct binary signature of type τ_2.

Proof. Note that when $a = -3(\kappa-1)b-(\kappa-2)(\kappa-1)c$, we have $\mathfrak{C} = 0$ and $2\mathfrak{B} = -\kappa[2b+(\kappa-2)c]$, which is nonzero by assumption. Let $\omega_0 = (9\kappa-10)b^2 + (\kappa-2)[2(3\kappa-5)bc + (\kappa^2-4\kappa+5)c^2]$ and assume $\omega_0 \neq 0$. Then let $\omega = \frac{(\kappa-1)[2b+(\kappa-2)c]^2}{\omega_0}$. By assumption, $\omega \neq 0$. Assume $\omega \neq 1$, which is equivalent to $-(5\kappa-6)b^2 - (\kappa-3)(\kappa-2)(2b-c)c \neq 0$. Further assume $\omega \neq -1$, which is equivalent to $(13\kappa-14)b^2 + (\kappa-2)[2(5\kappa-7)bc + (2\kappa^2-7\kappa+7)c^2] \neq 0$.

Consider the gadget in Figure 11.6b. We assign $\langle -3(\kappa-1)b-(\kappa-1)(\kappa-2)c, b, c\rangle$ to the vertices. Up to a nonzero factor of ω_0, this gadget has the succinct binary signature $\frac{1}{\omega_0}\langle x, y\rangle = \langle \omega+\kappa-1, \omega-1\rangle$ of type τ_2, where

$$x = (\kappa-1)\left\{3(3\kappa-2)b^2 + (\kappa-2)\left[6bc + (\kappa^2-3\kappa+3)c^2\right]\right\} \qquad \text{and}$$
$$y = -(5\kappa-6)b^2 - (\kappa-3)(\kappa-2)(2b-c)c.$$

Then we are done by Lemma 11.6.4 via case (i) of condition 3.

Now we deal with the following exceptional cases.

1. If $\omega_0 = 0$, then $-(9\kappa-10)b = [(\kappa-2)(3\kappa-5) \pm i\kappa\sqrt{2(\kappa-2)}]c$. Up to a nonzero factor of $\frac{c}{9\kappa-10}$, we have

$$\begin{aligned}&\frac{9\kappa-10}{c}\langle -3(\kappa-1)b-(\kappa-1)(\kappa-2)c, b, c\rangle\\ &\quad= \left\langle -(\kappa-1)\left[5(\kappa-2) \mp 3i\kappa\sqrt{2(\kappa-2)}\right],\quad -(\kappa-2)(3\kappa-5) \mp i\kappa\sqrt{2(\kappa-2)},\quad 9\kappa-10\right\rangle\end{aligned}$$

   and we are done by case 7 of Lemma 11.6.1.

2. If $-(5\kappa-6)b^2-(\kappa-3)(\kappa-2)(2b-c)c = 0$, then $-(5\kappa-6)b = \left[(\kappa-3)(\kappa-2)\pm\kappa\sqrt{\kappa^2-5\kappa+6}\right]c$. Up to a nonzero factor of $-\frac{c}{5\kappa-6}$, we have

$$\begin{aligned}&-\frac{5\kappa-6}{c}\langle -3(\kappa-1)b-(\kappa-1)(\kappa-2)c, b, c\rangle\\ &= \left\langle (\kappa-1)\left[(\kappa-2)(2\kappa+3) \mp 3\kappa\sqrt{\kappa^2-5\kappa+6}\right],\quad (\kappa-3)(\kappa-2)\pm\kappa\sqrt{\kappa^2-5\kappa+6},\quad -5\kappa+6\right\rangle\end{aligned}$$

   and are done by case 8 Lemma 11.6.1.

3. If $(13\kappa - 14)b^2 + (\kappa - 2)[2(5\kappa - 7)bc + (2\kappa^2 - 7\kappa + 7)c^2] = 0$, then

$$-(13\kappa - 14)b = \left[(\kappa - 2)(5\kappa - 7) \pm i\kappa\sqrt{\kappa^2 - \kappa - 2}\right]c.$$

   Up to a nonzero factor of $\frac{c}{13\kappa - 14}$, we have

$$\begin{aligned} &\frac{13\kappa - 14}{c}\langle -3(\kappa - 1)b - (\kappa - 1)(\kappa - 2)c, b, c\rangle \\ &= \langle (\kappa - 1)\left[(\kappa - 2)(2\kappa - 7) \pm 3i\kappa\sqrt{\kappa^2 - \kappa - 2}\right], \quad -(\kappa - 2)(5\kappa - 7) \mp i\kappa\sqrt{\kappa^2 - \kappa - 2}, \quad 13\kappa - 14\rangle \end{aligned}$$

   and are done by case 9 of Lemma 11.6.1. □

11.6.3 Eigenvalue Shifted Triples

To handle failure conditions (11.6.8) and (11.6.9) from Lemma 11.6.5, we need another technique. We introduce an Eigenvalue Shifted Triple, which extends the concept of an Eigenvalue Shifted Pair.

Definition 11.6.9 (Definition 4.6 in [93]). A pair of nonsingular matrices $M, M' \in \mathbb{C}^{2\times 2}$ is called an *Eigenvalue Shifted Pair* if $M' = M + \delta I$ for some nonzero $\delta \in \mathbb{C}$, and M has distinct eigenvalues.

Eigenvalue shifted pairs were used in [93] to show that interpolation succeeds in most cases since these matrices correspond to some recursive gadget constructions and at least one of them usually has eigenvalues with distinct complex norms. In [93], it is shown that the interpolation succeeds unless the variables in question take real values. Then other techniques were developed to handle the real case. We use Eigenvalue Shifted Pairs in a stronger way. We exhibit three matrices such that any two form an Eigenvalue Shifted Pair. Provided these shifts are linearly independent over $\mathbb{R}$, this is enough to show that interpolation succeeds for both real and complex settings of the variables. We call this an Eigenvalue Shifted Triple.

Definition 11.6.10. A trio of nonsingular matrices $M_0, M_1, M_2 \in \mathbb{C}^{2\times 2}$ is called an *Eigenvalue Shifted Triple* (EST) if M_0 has distinct eigenvalues and there exist nonzero $\delta_1, \delta_2 \in \mathbb{C}$ satisfying $\frac{\delta_1}{\delta_2} \notin \mathbb{R}$ such that $M_1 = M_0 + \delta_1 I$, and $M_2 = M_0 + \delta_2 I$.

If M_0, M_1, and M_2 form an Eigenvalue Shifted Triple, then any permutation of the matrices is also an Eigenvalue Shifted Triple.

The proof of the next lemma is similar to the proof of Lemma 4.7 in [94].

Lemma 11.6.11. *Suppose $\alpha, \beta, \delta_1, \delta_2 \in \mathbb{C}$. If $\alpha \neq \beta$, $\delta_1, \delta_2 \neq 0$, and $\frac{\delta_1}{\delta_2} \notin \mathbb{R}$, then $|\alpha| \neq |\beta|$ or $|\alpha + \delta_1| \neq |\beta + \delta_1|$ or $|\alpha + \delta_2| \neq |\beta + \delta_2|$.*

Proof. Assume for a contradiction that $|\alpha| = |\beta|$, $|\alpha + \delta_1| = |\beta + \delta_1|$, and $|\alpha + \delta_2| = |\beta + \delta_2|$. After a rotation in the complex plane, we can assume that $\alpha = \overline{\beta}$. Note that all of our assumptions are unchanged by this rotation. For $i \in \{1, 2\}$, we have

$$\begin{aligned}(\alpha + \delta_i)\overline{(\alpha + \delta_i)} &= |\alpha + \delta_i|^2 \\ &= |\beta + \delta_i|^2 \\ &= (\beta + \delta_i)\overline{(\beta + \delta_i)} = (\overline{\alpha} + \delta_i)(\alpha + \overline{\delta_i}).\end{aligned}$$

This implies $(\overline{\alpha} - \alpha)(\overline{\delta_i} - \delta_i) = 0$. Since $\alpha \neq \beta = \overline{\alpha}$, we have $\delta_i \in \mathbb{R}$. Then $\frac{\delta_1}{\delta_2} \in \mathbb{R}$, a contradiction. □

The next lemma considers the failure condition in (11.6.8), which is $\kappa = 3$ and $b = 0$, so the signature has the form $\langle a, 0, c\rangle$. If $a = 0$, then the problem is already #P-hard by Theorem 10.2.7. If $c = 0$, then the problem is tractable by case 1 of Corollary 11.1.3. If $a^3 = c^3$, then the problem is tractable by Corollary 11.1.5.

Lemma 11.6.12. *Suppose the domain size is 3 and $a, c \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle a, 0, c\rangle$ of type τ_3 and the succinct unary signature $\langle 1\rangle$ of type τ_1. If $ac \neq 0$ and $a^3 \neq c^3$, then*

$$\text{Pl-Holant}_3(\mathcal{F} \cup \{\langle x, y\rangle\}) \leq_T \text{Pl-Holant}_3(\mathcal{F})$$

for any $x, y \in \mathbb{C}$, where $\langle x, y\rangle$ is a succinct binary signature of type τ_2.

Proof. Assume $2a + c \neq 0$ and let $\omega = \frac{a^2+2c^2}{c(2a+c)}$. Assume $a^2 + 2c^2$ so that $\omega \neq 0$. Further assume $a^2 + 2ac + 3c^2 \neq 0$ so that $\omega^2 \neq 1$ as well as $a^2 + ac + 7c^2 \neq 0$ so that $\omega^3 \neq 1$. Note that these

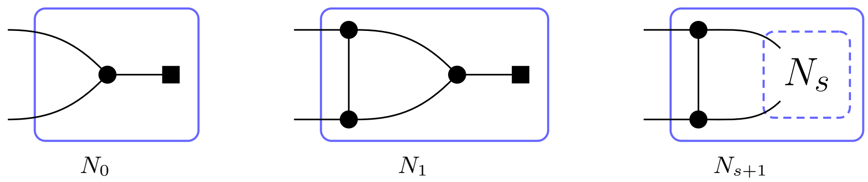


Figure 11.7: Alternative recursive construction to interpolate a binary signature (cf. Figure 10.4). Circle vertices are assigned $\langle a, b, c\rangle$ and the square vertex is assigned $\langle 1\rangle$.

conclusions also require $a \neq c$ and $a^3 \neq c^3$ respectively.

Consider using the recursive construction in Figure 11.7. The circle vertices are assigned $\langle a, 0, c\rangle$ and the square vertex is assigned $\langle 1\rangle$. Let $z = \frac{c}{a}$, which is well-defined by assumption. The succinct signature of type τ_2 for the initial gadget N_0 in this construction is $\langle a, c\rangle$. Up to a nonzero factor of a, this signature is $f_0 = \frac{1}{a}\langle a, c\rangle = \langle 1, z\rangle$. Then up to a nonzero factor of $c(2a + c)$, the succinct signature of type τ_2 for the sth gadget in this construction is $f_s = \langle \omega^k, z\rangle = M^s f_0$, where

$$M = \frac{1}{c(2a+c)} \begin{bmatrix} a^2 + 2c^2 & 0 \\ 0 & c(2a+c) \end{bmatrix} = \begin{bmatrix} \omega & 0 \\ 0 & 1 \end{bmatrix}.$$

Clearly M is nonsingular. The determinant of $[f_0\ Mf_0] = \left[\begin{smallmatrix} a & a\omega \\ c & c \end{smallmatrix}\right]$ is $z(1-\omega) \neq 0$. If ω is not a root of unity, then we are done by Lemma 10.1.3.

Otherwise, suppose ω is a primitive root of unity of order n. By assumption, $n \geq 4$. Now consider the recursive construction in Figure 10.4. We assign f_s to every vertex, where $s \geq 0$ is a parameter of our choice. Let $g_t(s)$ be the signature of the tth gadget in this recursive construction when using f_s. Then $g_1(s) = f_s$ and $g_t(s) = (N(s))^t \left[\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right]$, where $N(s) = \left[\begin{smallmatrix} \omega^s & 2z \\ z & \omega^s + z \end{smallmatrix}\right]$.

By Lemma 10.3.2, the eigenvalues of $N(s)$ are $\omega^s + 2z$ and $\omega^s - z$, which means the determinant of $N(s)$ is $(\omega^s + 2z)(\omega^s - z)$. Each eigenvalue can vanish for at most one value of $s \in \mathbb{Z}_n$ since both eigenvalues are linear polynomials in ω^s that are not identically 0. Furthermore, at least one of the eigenvalues never vanishes for all $s \in \mathbb{Z}_n$ since otherwise $1 = |z| = \frac{1}{2}$.

Thus, at most one matrix among $N(0)$, $N(1)$, $N(2)$, and $N(3)$ can be singular. Pick distinct $j, k, \ell \in \{0, 1, 2, 3\}$ such that $N(j)$, $N(k)$, and $N(\ell)$ are nonsingular. To finish the proof, we show that $N(j)$, $N(k)$, and $N(\ell)$ form an Eigenvalue Shifted Triple. Then by Lemma 11.6.11, at least one of the matrices has eigenvalues with distinct complex norms, so we are done by Corollary 10.3.3.

The eigenvalue shift from $N(j)$ to $N(k)$ is $\delta_{j,k} = \omega^j(\omega^{k-j} - 1)$, which is nonzero since j and k are distinct in $\mathbb{Z}_n$. Assume for a contradiction that $\frac{\delta_{j,k}}{\delta_{j,\ell}} \in \mathbb{R}$, which is equivalent to $\arg(\delta_{j,k}) = \arg(\pm\delta_{j,\ell})$. Then we have

$$\arg\left(\omega^{k-j} - 1\right) = \arg\left(\pm(\omega^{\ell-j} - 1)\right). \tag{11.6.10}$$

In the complex plane, any nonzero $x - 1 \in \mathbb{C}$ with $|x| = 1$ lies on the circle of radius 1 centered at $(-1, 0)$. Such x satisfy $\frac{\pi}{2} < \arg(x-1) < \frac{3\pi}{2}$. Thus, the argument of $x - 1$ is unique, even up to a sign, contradicting (11.6.10). Therefore, M_j, M_k, and M_ℓ form an Eigenvalue Shifted Triple as claimed.

Now we deal with the following exceptional cases.

1. If $2a + c = 0$, then up to a nonzero factor of a, we have $\frac{1}{a}\langle a, 0, c\rangle = \langle 1, 0, -2\rangle$ and are done by case 10 of Lemma 11.6.1.
2. If $a^2 + 2c^2 = 0$, then $a = \pm i\sqrt{2}c$. Up to a nonzero factor of c, we have $\frac{1}{c}\langle a, 0, c\rangle = \langle \pm i\sqrt{2}, 0, 1\rangle$ and are done by case 11 of Lemma 11.6.1.
3. If $a^2 + 2ac + 3c^2 = 0$, then $a = c(-1 \pm i\sqrt{2})$. Up to a nonzero factor of c, we have $\frac{1}{c}\langle a, 0, c\rangle = \langle -1 \pm i\sqrt{2}, 0, 1\rangle$ and are done by case 12 of Lemma 11.6.1.
4. If $a^2 + ac + 7c^2 = 0$, then $2a = c(-1 \pm 3i\sqrt{3})$. Up to a nonzero factor of $\frac{c}{2}$, we have $\frac{2}{c}\langle a, 0, c\rangle = \langle -1 \pm 3i\sqrt{3}, 0, 2\rangle$ and are done by case 13 of Lemma 11.6.1. □

The next lemma considers the failure condition in (11.6.9). Since this failure condition is just a holographic transformation of the failure condition in (11.6.8), the excluded cases in this lemma are handled exactly as those preceding Lemma 11.6.12.

Lemma 11.6.13. *Suppose the domain size is* 3 *and* $b, c \in \mathbb{C}$. *Let* $\mathcal{F}$ *be a signature set containing the succinct ternary signature* $\langle -3b - 4c, 2b, 2c\rangle$ *of type* τ_3 *and the succinct unary signature* $\langle 1\rangle$ *of type* τ_1. *Assume* $T^{\otimes 3}\langle -3b - 4c, 2b, 2c\rangle = \langle \hat{a}, \hat{b}, \hat{c}\rangle$, *where* $T = \left[\begin{smallmatrix} 1 & -2 & -2 \\ -2 & 1 & -2 \\ -2 & -2 & 1 \end{smallmatrix}\right]$. *If* $\hat{a}\hat{c} \neq 0$ *and* $\hat{a}^3 \neq \hat{c}^3$, *then*

$$\text{Pl-Holant}_3(\mathcal{F} \cup \{\langle x, y\rangle\}) \leq_T \text{Pl-Holant}_3(\mathcal{F})$$

for any $x, y \in \mathbb{C}$, where $\langle x, y\rangle$ is a succinct binary signature of type τ_2.

Proof. By Lemma 9.2.6 with $x = 1$ and $y = -2$, we have $\hat{b} = 0$. Thus after a holographic transformation by T, we are in the case covered by Lemma 11.6.12. Since T is orthogonal after scaling by $\frac{1}{3}$, the complexity of these problems are unchanged by Lemma 3.2.2. □

We summarize this section with the following lemma.

Corollary 11.6.14. *Suppose the domain size is $\kappa \geq 3$ and $a, b, c \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle a, b, c\rangle$ of type τ_3 and the succinct unary signature $\langle 1\rangle$ of type τ_1. Then*

$$\operatorname{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle x, y\rangle\}) \leq_T \operatorname{Pl-Holant}_\kappa(\mathcal{F})$$

for any $x, y \in \mathbb{C}$, where $\langle x, y\rangle$ is a succinct binary signature of type τ_2, unless

- *$\mathfrak{B} = 0$ or*
- *there exist $\lambda \in \mathbb{C}$ and $T \in \{I_\kappa, \kappa I_\kappa - 2J_\kappa\}$ such that*

$$\langle a, b, c\rangle = \begin{cases} T^{\otimes 3}\lambda\langle 1, 0, 0\rangle, & \textit{or} \\ T^{\otimes 3}\lambda\langle 0, 0, 1\rangle \textit{ and } \kappa = 3, & \textit{or} \\ T^{\otimes 3}\lambda\langle 1, 0, \omega\rangle \textit{ and } \kappa = 3 \textit{ where } \omega^3 = 1, & \textit{or} \\ T^{\otimes 3}\lambda\langle \mu^2, 1, \mu\rangle \textit{ and } \kappa = 4 \textit{ where } \mu = -1 \pm 2i. \end{cases}$$

Proof. If failure condition (11.6.5), (11.6.6), (11.6.7), (11.6.8), or (11.6.9) holds, then we are done by Lemma 11.6.6, Lemma 11.6.7, Lemma 11.6.8, Lemma 11.6.12, or Lemma 11.6.13 respectively, with the various excluded cases listed. If none of (11.6.5), (11.6.6), (11.6.7), (11.6.8), and (11.6.9) hold, then we are done by Lemma 11.6.5. □

11.7 Puiseux series, Siegel's Theorem, and Galois theory

This section covers the last stage of our hardness proof, which assumes that all succinct binary signatures of type τ_2 are available. Among the ways we utilize this assumption is to build the

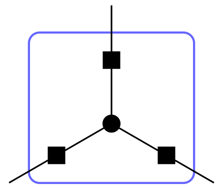


Figure 11.8: Local holographic transformation gadget construction for a ternary signature.

gadget known as a local holographic transformation (see Figure 11.8), which is the focus of Subsection 11.7.1. Then in Subsection 11.7.2, our efforts are largely spent proving that a certain interpolation succeeds. To that end, we employ Galois theory aided by an effective version of Siegel's theorem for a specific algebraic curve, which is made possible by analyzing Puiseux series expansions.

11.7.1 Constructing a Special Ternary Signature

We construct one of two special ternary signatures. Either we construct a signature of the form $\langle a, b, b\rangle$ with $a \neq b$ and can finish the proof with Corollary 10.3.7 or we construct $\langle 3(\kappa-1), \kappa-3, -3\rangle$. With this latter signature, we can interpolate the weighted Eulerian partition signature.

A key step in our dichotomy theorem occurred back in Section 10.3 through Lemma 10.3.6 with the Bobby Fischer gadget. To apply this lemma, we need to construct a gadget with a succinct ternary signature of type τ_3 such that the last two entries are equal and different from the first. This is the goal of the next lemma, which assumes $\mathfrak{B} \neq 0$. We will determine the complexity of the case $\mathfrak{B} = 0$ in Corollary 11.5.5 without using the results from this section.

Lemma 11.7.1. *Suppose $\kappa \geq 3$ is the domain size and $a, b, c \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle a, b, c\rangle$ of type τ_3 and the succinct binary signature $\langle x, y\rangle$ of type τ_2 for all $x, y \in \mathbb{C}$. If $\mathfrak{A}\mathfrak{B} \neq 0$, then there exist $a', b' \in \mathbb{C}$ satisfying $a' \neq b'$ such that*

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle a', b', b'\rangle\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F})$$

where $\langle a', b', b'\rangle$ is a succinct ternary signature of type τ_3.

Proof. Consider the gadget in Figure 11.8. We assign $\langle a, b, c\rangle$ to the circle vertex and $\langle x, y\rangle$ to the square vertices for some $x, y \in \mathbb{C}$ of our choice, to be determined shortly. By Lemma 9.2.6, the succinct ternary signature of type τ_3 for the resulting gadget is $\langle a', b', c'\rangle$, where

$$a' - b' = (x - y)^2[2\mathfrak{D} + \mathfrak{A}(x - y)] \qquad \text{and} \qquad b' - c' = (x - y)^2\mathfrak{D}$$

with $\mathfrak{D} = (b - c)(x - y) + \mathfrak{B}y$. We pick $x = \mathfrak{B} + y$ and $y = -(b - c)$ so that $\mathfrak{D} = 0$ and thus $b' - c' = 0$. Then the first difference simplifies to $a' - b' = \mathfrak{A}\mathfrak{B}^3 \neq 0$. This signature has the desired properties, so we are done. □

The previous proof fails when $\mathfrak{A} = 0$ because such signatures are invariant set-wise under this type of local holographic transformation. With the exception of a single point, we can use this same gadget construction to reduce between any two of these points.

Lemma 11.7.2. *Suppose $\kappa \geq 3$ is the domain size and $b, c, s, t \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle 3b - 2c, b, c\rangle$ of type τ_3 and the succinct binary signature $\langle x, y\rangle$ of type τ_2 for all $x, y \in \mathbb{C}$. If $b \neq c$, $3b + (\kappa - 3)c \neq 0$, and $3s + (\kappa - 3)t \neq 0$, then*

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle 3s - 2t, s, t\rangle\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F}),$$

where $\langle 3s - 2t, s, t\rangle$ is a succinct ternary signature of type τ_3.

Proof. Consider the gadget in Figure 11.8. We assign $\langle 3b - 2c, b, c\rangle$ to the circle vertex and $\langle x, y\rangle$ to the square vertices for some $x, y \in \mathbb{C}$ of our choice, to be determined shortly. By Lemma 9.2.6, the signature of this gadget is $f = [x + (\kappa - 1)y]\langle 3\hat{b} - 2\hat{c}, \hat{b}, \hat{c}\rangle$, where

$$\hat{b} = bx^2 + 2[2b + (\kappa - 3)c]xy + [(3\kappa - 5)b + (\kappa^2 - 5\kappa + 6)c]y^2 \qquad \text{and}$$
$$\hat{c} = cx^2 + 2[3b + (\kappa - 4)c]xy + [(3\kappa - 6)b + (\kappa^2 - 5\kappa + 7)c]y^2.$$

We note that the difference $\hat{b} - \hat{c}$ nicely factors as

$$\hat{b} - \hat{c} = (b - c)(x - y)^2.$$

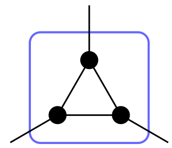


Figure 11.9: Triangle gadget used to construct $\langle 3(\kappa-1), \kappa-3, -3\rangle$.

We pick $x = y + \sqrt{s-t}$ so that $\hat{b} - \hat{c} = (b-c)(s-t)$ is the desired difference $s - t$ up to a nonzero factor of $b-c$. Then we want to set $\hat{c}$ to be $(b-c)t$. With $x = y + \sqrt{s-t}$, we can simplify $(b-c)t - \hat{c}$ to

$$(b-c)t - \hat{c} = -\kappa[3b + (\kappa-3)c]y^2 - 2\sqrt{s-t}[3b + (\kappa-3)c]y + bt - cs. \tag{11.7.11}$$

Since $\kappa[3b + (\kappa-3)c] \neq 0$, (11.7.11) is a nontrivial quadratic polynomial in y, so we can set y such that this expression vanishes. Then the signature is $f = (b-c)[x + (\kappa-1)y]\langle 3s-2t, s, t\rangle$. It remains to check that $x + (\kappa-1)y \neq 0$.

If $x + (\kappa-1)y = 0$, then $y = -\frac{\sqrt{s-t}}{\kappa}$. However, plugging this into (11.7.11) gives

$$\frac{(b-c)[3s + (\kappa-3)t]}{k} \neq 0,$$

so $x + (\kappa-1)y$ is indeed nonzero. □

If $\mathfrak{A} = 0$ and $3b + (\kappa-3)c = 0$, then $-3\langle a, b, c\rangle$ simplifies to $c\langle 3(\kappa-1), \kappa-3, -3\rangle$, which is a failure condition of the previous lemma. The reason is that this signature is pointwise invariant under such local holographic transformations. However, a different ternary construction can reach this point.

Lemma 11.7.3. *Suppose $\kappa \geq 3$ is the domain size and $b, c \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle 3b-2c, b, c\rangle$ of type τ_3 and the succinct binary signature $\langle x, y\rangle$ of type τ_2 for every $x, y \in \mathbb{C}$. If $b \neq c$, then*

$$\text{Pl-Holant}_\kappa(\mathcal{F} \cup \{\langle 3(\kappa-1), \kappa-3, -3\rangle\}) \leq_T \text{Pl-Holant}_\kappa(\mathcal{F}),$$

where $\langle 3(\kappa-1), \kappa-3, -3\rangle$ is a succinct ternary signature of type τ_3.

Proof. If $3b + (\kappa - 3)c = 0$, then up to a nonzero factor of $\frac{-c}{3}$, $\langle 3b - 2c, b, c \rangle$ is already the desired signature. Otherwise, $3b + (\kappa - 3)c \neq 0$. By Lemma 11.7.2, we have $\langle 3s - 2t, s, t \rangle$ for any $s, t \in \mathbb{C}$ satisfying $3s + (\kappa - 3)t \neq 0$.

Consider the gadget in Figure 11.9. We assign $\langle 3s - 2t, s, t \rangle$ to vertices for some $s, t \in \mathbb{C}$ satisfying $3s + (\kappa - 3)t \neq 0$ of our choice, to be determined shortly. By Lemma 9.2.4, the signature of this gadget is $\langle 3s' - 2t', s', t' \rangle$, where

$$s' = (5\kappa + 14)s^3 + (\kappa^2 + 9\kappa - 42)s^2 t + (7\kappa^2 - 33\kappa + 42)st^2 + (\kappa - 2)(\kappa^2 - 6\kappa + 7)t^3, \qquad \text{and}$$
$$t' = (\kappa + 14)s^3 + 21(\kappa - 2)s^2 t + 3(3\kappa^2 - 15\kappa + 14)st^2 + (\kappa^3 - 9\kappa^2 + 23\kappa - 14)t^3.$$

It suffices to pick s and t satisfying $3s + (\kappa - 3)t \neq 0$ such that $s' = \kappa - 3$ and $t' = -3$ up to a common nonzero factor.

We note that the difference $s' - t'$ factors as

$$s' - t' = \kappa(s - t)^2 [4s + (\kappa - 4)t].$$

We pick $s = \frac{-(\kappa-4)t+1}{4}$ so that $s' - t' = \kappa(s - t)^2$ is the desired difference κ up to a factor of $(s - t)^2$. Then we want to set t' to be $-3(s - t)^2$. With $s = \frac{-(\kappa-4)t+1}{4}$, we can simplify $-3(s - t)^2 - t'$ to

$$-3(s - t)^2 - t' = \frac{1}{64} \left[\kappa^3(\kappa - 2)t^3 - 3\kappa^2(\kappa + 2)t^2 + 3\kappa(\kappa - 10)t - (\kappa + 26) \right]. \tag{11.7.12}$$

Since $\kappa \geq 3$, (11.7.12) is a nontrivial cubic polynomial in t, so we can set t such that this expression vanishes. Then $\langle 3s' - 2t', s', t' \rangle = (s - t)^2 \langle 3(\kappa - 1), \kappa - 3, -3 \rangle$. It remains to check that $s \neq t$ and $3s + (\kappa - 3)t \neq 0$.

If $s = t$, then $t = \frac{1}{\kappa}$. Plugging this into (11.7.12) gives -1, so $s \neq t$. If $3s + (\kappa - 3)t = 0$, then $t = -\frac{3}{\kappa}$. Plugging this into (11.7.12) gives $1 - \kappa \neq 0$, so $3s + (\kappa - 3)t \neq 0$. ☐

Remark. We originally proved Theorem 11.1.1 for $\kappa = 3$. For this lemma with $\kappa = 3$, we had picked a specific value for t. When generalizing it to any $\kappa \geq 3$, I knew that we had to something different. For $\kappa = 3$, it was already very complicated to express the value of t. The loose idea that

I had was to make the proof nonconstructive so that we didn't have to explicitly state the value of t. We noticed that the difference of the succinct signature entries gave a nice expression. This allows us to be explicit in our first choice (i.e the choice for s) and more implicit in our second choice (i.e the choice for t). It may seem like a simple matter here, but I think that significant progress on higher domain Holant problems will require nonconstructive ideas like this one.

11.7.2 Dose of an effective Siegel's Theorem and Galois theory

To finish the last part of our hardness proof, it suffices to show that $\langle 3(\kappa-1), \kappa-3, -3\rangle$ is #P-hard for all $\kappa \geq 3$. The general strategy is to use interpolation. However, proving that this interpolation succeeds presents a significant challenge.

Remark. Given a succinct ternary signature of type τ_3 satisfying $\mathfrak{A} = 0$, the results in the previous subsection essentially show that we can construct any succinct ternary signature of type τ_3 satisfying $\mathfrak{A} = 0$. We picked $\langle 3(\kappa-1), \kappa-3, -3\rangle$ by intuition based on the fact that it is the sole invariant point of the local holographic transformation gadget construction (see Figure 11.8). It is possible that another signature might have been easier to prove #P-hard.

Consider the polynomial $p(x, y) \in \mathbb{Z}[x, y]$ defined by

$$\begin{aligned} p(x, y) &= x^5 - 2x^3y - x^2y^2 - x^3 + xy^2 + y^3 - 2x^2 - xy \\ &= x^5 - (2y+1)x^3 - (y^2+2)x^2 + y(y-1)x + y^3. \end{aligned}$$

We consider y as an integer parameter $y \geq 4$, and treat $p(x, y)$ as an infinite family of quintic polynomials in x with integer coefficients. We want to show that the roots of all these quintic polynomials satisfy the lattice condition. First, we determine the number of real and nonreal roots.

Lemma 11.7.4. *For any integer $y \geq 1$, the polynomial $p(x, y)$ in x has three distinct real roots and two nonreal complex conjugate roots.*

Proof. Up to a factor of $-4y^2$, the discriminant of $p(x, y)$ (with respect to x) is

$$27y^{11} - 4y^{10} + 726y^9 - 493y^8 + 2712y^7 - 400y^6 - 2503y^5 + 475y^4 + 956y^3 - 904y^2 + 460y + 104.$$

By replacing y with $z+1$, we get

$$27z^{11}+293z^{10}+2171z^9+10316z^8+33334z^7+77398z^6+127383z^5$$
$$+141916z^4+102097z^3+44373z^2+10336z+1156,$$

which is positive for any $z \geq 0$. Thus the discriminant is negative.

Therefore, $p(x,y)$ has distinct roots in x for all $y \geq 1$. Furthermore, with a negative discriminant, $p(x,y)$ has $2s$ nonreal complex conjugate roots for some odd integer s. Since $p(x,y)$ is a quintic polynomial (in x), the only possibility is $s=1$. □

We suspect that for any integer $y \geq 4$, $p(x,y)$ is in fact irreducible over $\mathbb{Q}$ as a polynomial in x. When considering y as an indeterminate, the bivariate polynomial $p(x,y)$ is irreducible over $\mathbb{Q}$ and the algebraic curve it defines has genus 3, so by Theorem 1.2 in [108], $p(x,y)$ is reducible over $\mathbb{Q}$ for at most a finite number of $y \in \mathbb{Z}$. For any integer $y \geq 4$, if $p(x,y)$ is irreducible over $\mathbb{Q}$ as a polynomial in x, then its Galois group is S_5 and its roots satisfy the lattice condition.

Lemma 11.7.5. *For any integer $y \geq 4$, if $p(x,y)$ is irreducible in $\mathbb{Q}[x]$, then the roots of $p(x,y)$ satisfy the lattice condition.*

Proof. By Lemma 11.7.4, $p(x,y)$ has three distinct real roots and two nonreal complex conjugate roots. With three distinct real roots, we know that not all the roots have the same complex norm. It is well-known that an irreducible polynomial of prime degree n with exactly two nonreal roots has S_n as a Galois group over $\mathbb{Q}$ (for example Theorem 10.15 in [116]). Then we are done by Lemma 11.3.5. □

We know of just five values of $y \in \mathbb{Z}$ for which $p(x,y)$ is reducible as a polynomial in x:

$$p(x,y) = \begin{cases} (x-1)(x^4+x^3+2x^2-x+1) & y=-1 \\ x^2(x^3-x-2) & y=0 \\ (x+1)(x^4-x^3-2x^2-x+1) & y=1 \\ (x-1)(x^2-x-4)(x^2+2x+2) & y=2 \\ (x-3)(x^4+3x^3+2x^2-5x-9) & y=3. \end{cases}$$

These five factorizations also give five integer solutions to $p(x,y)=0$. It is a well-known theorem of Siegel [115] that an algebraic curve of genus at least 1 has only a finite number of integral points. For this curve of genus 3, Faltings' Theorem [66] says that there can be only a finite number of rational points. However these theorems are not *effective* in general. There are some effective versions of Siegel's Theorem that can be applied to our polynomial, but the best effective bound that we can find is over $10^{20{,}000}$ [140] and hence cannot be checked in practice.

However, it is shown in the next lemma that in fact these five are the only integer solutions. In particular, for any integer $y \geq 4$, $p(x,y)$ does not have a linear factor in $\mathbb{Z}[x]$, and hence by Gauss's Lemma, also no linear factor in $\mathbb{Q}[x]$. The following proof is essentially due to Aaron Levin [100]. We thank Aaron for suggesting the key auxiliary function $g_2(x,y) = \frac{y^2}{x} + y - x^2 + 1$, as well as for his permission to include the proof here. We also thank Bjorn Poonen [111] who suggested a similar proof. After the proof, we will explain certain complications in the proof.

Lemma 11.7.6. *The only integer solutions to $p(x,y)=0$ are $(1,-1)$, $(0,0)$, $(-1,1)$, $(1,2)$, and $(3,3)$.*

Proof. Clearly these five points are solutions to $p(x,y)=0$. For $a \in \mathbb{Z}$ with $-3 < a < 17$, one can directly check that $p(a,y)=0$ has no other integer solutions in y.

Let $(a,b) \in \mathbb{Z}^2$ be a solution to $p(x,y)=0$ with $a \neq 0$. We claim $a \mid b^2$. By definition of $p(x,y)$, clearly $a \mid b^3$. If p is a prime that divides a, then let $\mathrm{ord}_p(a)=e$ and $\mathrm{ord}_p(b)=f$ be the exact orders with which p divides a and b respectively. Then $f \geq 1$ since $3f \geq e$ and our claim is that

$2f \geq e$. Suppose for a contradiction that $2f < e$. From $p(a,b) = 0$, we have

$$a^2(a^3 - 2ab - a - b^2 - 2) = -b^3 - ab(b-1).$$

The order with respect to p of the left-hand side is

$$\operatorname{ord}_p\left(a^2(a^3 - 2ab - a - b^2 - 2)\right) \geq \operatorname{ord}_p\left(a^2\right) = 2e.$$

Since p is relatively prime to $b-1$, $\operatorname{ord}_p\left(ab(b-1)\right) = e + f > 3f$, and therefore the order of the right-hand side with respect to p is

$$\operatorname{ord}_p\left(-b^3 - ab(b-1)\right) = \operatorname{ord}_p(b^3) = 3f.$$

However, $2e > 3f$, a contradiction. This proves the claim.

Now consider the functions $g_1(x,y) = y - x^2$ and $g_2(x,y) = \frac{y^2}{x} + y - x^2 + 1$. Whenever $(a,b) \in \mathbb{Z}^2$ is a solution to $p(x,y) = 0$ with $a \neq 0$, $g_1(a,b)$ and $g_2(a,b)$ are integers. However, we show that if $a \leq -3$ or $a \geq 17$, then either $g_1(a,b)$ or $g_2(a,b)$ is not an integer.

Let $c_2 = -(x-1)x$, $c_1 = -x(2x^2+1)$, and $c_0 = x^2(x^3 - x - 2)$ so that $p(x,y) = y^3 + c_2y^2 + c_1y + c_0$. Then the discriminant of $p(x,y)$ with respect to y is

$$\begin{aligned}\operatorname{disc}_y(p(x,y)) &= c_2^2c_1^2 - 4c_1^3 - 4c_2^3c_0 - 27c_0^2 + 18c_2c_1c_0 \\ &= (x-1)x^3(4x^7 + 5x^6 + x^5 + 45x^4 + 151x^3 + 163x^2 + 67x - 4). \qquad (11.7.13)\end{aligned}$$

Suppose $x \leq -3$. Replacing x with $-z-1$ in (11.7.13) gives

$$-(z+1)^3(z+2)(4z^7 + 23z^6 + 55z^5 + 25z^4 + 21z^3 + 39z^2 + 17z + 14).$$

This is clearly negative (for $z \geq 0$), so (11.7.13) is negative. Thus $p(x,y)$ only has one real root as a polynomial in y. Let $y_1(x)$ be that root and consider $y_1^-(x) = x^2 + 2x^{-1}$ and $y_1^+(x) = x^2 + 2x^{-1} + 2x^{-2}$. We have $p(x, y_1^-(x)) = -2x^2 + 6 + 4x^{-1} + 8x^{-3} < 0$. Also $p(x, y_1^+(x)) =$

$6+18x^{-1}+16x^{-2}+12x^{-3}+24x^{-4}+24x^{-5}+8x^{-6}>0$. Hence $y_1^-(x)<y_1(x)<y_1^+(x)$, and all three are positive since $y_1^-(x)$ is positive. Then for $x\le -3$,

$$-1<2x^{-1}=g_1(x,y_1^-(x))<g_1(x,y_1(x))<g_1(x,y_1^+(x))=2x^{-1}+2x^{-2}<0,$$

so $g_1(x,y_1(x))$ is not an integer. Therefore, $y_1(x)$, the only real root for any integer $x\le -3$, is not an integer.

Now suppose $x\ge 17$. Then (11.7.13) is positive and there are three distinct real roots. Similar to the previous argument, we have $p(x,y_1^-(x))<0$ and $p(x,y_1^+(x))>0$. Hence there is some root $y_1(x)$ in the open interval $(y_1^-(x),y_1^+(x))$. All three terms $y_1^-(x)<y_1(x)<y_1^+(x)$ are positive because $y_1^-(x)>0$. Then

$$0<2x^{-1}=g_1(x,y_1^-(x))<g_1(x,y_1(x))<g_1(x,y_1^+(x))=2x^{-1}+2x^{-2}<1,$$

so $g_1(x,y_1(x))$ is not an integer.

There are two more real roots. Consider

$$\begin{aligned} y_2^-(x)&=x^{3/2}-\frac{1}{2}x+\frac{1}{8}x^{1/2}-\frac{65}{128}x^{-1/2}-2x^{-1} \qquad \text{and}\\ y_2^+(x)&=x^{3/2}-\frac{1}{2}x+\frac{1}{8}x^{1/2}-\frac{65}{128}x^{-1/2}. \end{aligned}$$

Replacing x with $(z+2)^2$ in

$$\begin{aligned} p(x,y_2^-(x))=2x^{5/2}-\frac{2495}{512}x^2+\frac{1087}{512}x^{3/2}-\frac{19569}{16384}x-\frac{8579}{16384}x^{1/2}+\frac{126847}{32768}+\frac{1452419}{131072}x^{-1/2}\\ -\frac{317}{256}x^{-1}+\frac{2871103}{2097152}x^{-3/2}-\frac{12675}{8192}x^{-2}-\frac{195}{32}x^{-5/2}-8x^{-3} \end{aligned}$$

gives

$$\frac{1}{2097152(z+2)^6}\begin{pmatrix} 4194304z^{11}+82055168z^{10}+722808832z^9+3774605184z^8\\ +12935149184z^7+30375187136z^6+49489164080z^5+55372934880z^4\\ +41238374079z^3+19431701370z^2+5465401844z+812262392 \end{pmatrix},$$

which is clearly positive ($z \geq 0$). Thus, $p(x, y_2^-(x)) > 0$. Also

$$p(x, y_2^+(x)) = \\ -2x^{5/2} - \frac{447}{512}x^2 - \frac{193}{512}x^{3/2} - \frac{3185}{16384}x + \frac{20605}{16384}x^{1/2} - \frac{4225}{32768} + \frac{12675}{131072}x^{-1/2} - \frac{274625}{2097152}x^{-3/2} \\ < 0.$$

Hence there is some root $y_2(x)$ in the open interval $(y_2^-(x), y_2^+(x))$. All three terms $y_2^-(x) < y_2(x) < y_2^+(x)$ are positive because $y_2^-(x) > 0$. Hence, for $x \geq 17$,

$$\begin{aligned} -1 < & -4x^{-1/2} - \frac{65}{512}x^{-1} - \frac{1}{2}x^{-3/2} + \frac{4225}{16384}x^{-2} + \frac{65}{32}x^{-5/2} + 4x^{-3} \\ &= g_2(x, y_2^-(x)) < g_2(x, y_2(x)) < g_2(x, y_2^+(x)) = -\frac{65}{512}x^{-1} + \frac{4225}{16384}x^{-2} < 0, \end{aligned}$$

so $g_2(x, y_2(x))$ is not an integer.

Finally, consider

$$\begin{aligned} y_3^-(x) &= -x^{3/2} - \frac{1}{2}x - \frac{1}{8}x^{1/2} + \frac{65}{128}x^{-1/2} - x^{-1} \qquad \text{and} \\ y_3^+(x) &= -x^{3/2} - \frac{1}{2}x - \frac{1}{8}x^{1/2} + \frac{65}{128}x^{-1/2} - \frac{1}{2}x^{-1}. \end{aligned}$$

We have

$$\begin{aligned} p(x, y_3^-(x)) = & -\frac{1471}{512}x^2 - \frac{447}{512}x^{3/2} - \frac{11377}{16384}x - \frac{6013}{16384}x^{1/2} + \frac{94079}{32768} - \frac{339331}{131072}x^{-1/2} - \frac{61}{512}x^{-1} \\ & - \frac{511807}{2097152}x^{-3/2} - \frac{12675}{16384}x^{-2} + \frac{195}{128}x^{-5/2} - x^{-3} \\ & < 0. \end{aligned}$$

Replacing x with $(z+3)^2$ in

$$\begin{aligned} p(x, y_3^+(x)) = x^{5/2} &- \frac{959}{512}x^2 - \frac{127}{512}x^{3/2} - \frac{7281}{16384}x - \frac{13309}{16384}x^{1/2} + \frac{53119}{32768} - \frac{77699}{131072}x^{-1/2} \\ &+ \frac{67}{1024}x^{-1} + \frac{78017}{2097152}x^{-3/2} - \frac{12675}{32768}x^{-2} + \frac{195}{512}x^{-5/2} - \frac{1}{8}x^{-3} \end{aligned}$$

gives

$$\frac{1}{2097152(z+3)^6}\begin{pmatrix} 2097152z^{11}+65277952z^{10}+919728128z^9+7736969088z^8 \\ +43137332608z^7+167175471424z^6+458797435600z^5+889807335920z^4 \\ +1191781601633z^3+1045691960361z^2+537771428331z+121660965323 \end{pmatrix},$$

which is clearly positive ($z \geq 0$). Thus, $p(x, y_3^+(x)) > 0$. Hence there is some root $y_3(x)$ in the open interval $(y_3^-(x), y_3^+(x))$. All three terms $y_3^-(x) < y_3(x) < y_3^+(x)$ are negative because $y_3^+(x) < 0$. Furthermore, the partial derivative $\frac{\partial g_2(x,y)}{\partial y} = 2x^{-1}y+1$ and $\frac{\partial^2 g_2(x,y)}{\partial y^2} = 2x^{-1} > 0$. Thus $\frac{\partial g_2(x,y)}{\partial y} \leq \frac{\partial g_2(x,y)}{\partial y}\big|_{y=y_3^+(x)} = -2x^{1/2} - \frac{1}{4}x^{-1/2} + \frac{65}{64}x^{-3/2} - x^{-2} < 0$, for all $y \in (-\infty, y_3^+(x)]$. Thus, $g_2(x,y)$ is decreasing monotonically in y over the interval $(-\infty, y_3^+(x)]$. Then

$$\begin{aligned} 0 &< x^{-1/2} - \frac{65}{512}x^{-1} + \frac{1}{8}x^{-3/2} + \frac{4225}{16384}x^{-2} - \frac{65}{128}x^{-5/2} + \frac{1}{4}x^{-3} \\ &= g_2(x, y_3^+(x)) < g_2(x, y_3(x)) < g_2(x, y_3^-(x)) \\ &= 2x^{-1/2} - \frac{65}{512}x^{-1} + \frac{1}{4}x^{-3/2} + \frac{4225}{16384}x^{-2} - \frac{65}{64}x^{-5/2} + x^{-3} < 1, \end{aligned}$$

so $g_2(x, y_3(x))$ is not an integer. To complete the proof, notice that the intervals $(y_1^-(x), y_1^+(x))$, $(y_2^-(x), y_2^+(x))$, and $(y_3^-(x), y_3^+(x))$ are disjoint. Therefore, we have shown that none of the three roots is an integer for any integer $x \geq 17$. □

Remark. One can obtain the Puiseux series expansions for $p(x,y)$, which are

$$\begin{aligned} y_1(x) &= x^2 + 2x^{-1} + 2x^{-2} - 6x^{-4} - 18x^{-5} + O(x^{-6}) && \text{for } x \in \mathbb{R}, \\ y_2(x) &= x^{3/2} - \frac{1}{2}x + \frac{1}{8}x^{1/2} - \frac{65}{128}x^{-1/2} - x^{-1} - \frac{1471}{1024}x^{-3/2} - x^{-2} + O(x^{-5/2}) && \text{for } x > 0, \text{ and} \\ y_3(x) &= -x^{3/2} - \frac{1}{2}x - \frac{1}{8}x^{1/2} + \frac{65}{128}x^{-1/2} - x^{-1} + \frac{1471}{1024}x^{-3/2} - x^{-2} + O(x^{-5/2}) && \text{for } x > 0. \end{aligned}$$

These series converge to the actual roots of $p(x,y)$ for large x. The basic idea of the proof—called Runge's method—is that, for example, when we substitute $y_2(x)$ in $g_2(x,y)$, we get $g_2(x, y_2(x)) = O(x^{-1/2})$, where the multiplier in the O-notation is bounded both above and below by a nonzero constant in absolute value. Thus for large x, this cannot be an integer. However, for integer solutions (x, y) of $p(x, y)$, this must be an integer.

We note that the expressions for the $y_i^+(x)$ and $y_i^-(x)$ are the truncated or rounded Puiseux series expansions. The reason we discuss $y_i^+(x)$ and $y_i^-(x)$ is because we want to prove an absolute bound, instead of the asymptotic bound implied by the O-notation.

By Lemma 11.7.6, if $p(x, y)$ is reducible over $\mathbb{Q}$ as a polynomial in x for any integer $y \geq 4$, then the only way it can factor is as a product of an irreducible quadratic and an irreducible cubic. The next lemma handles this possibility.

Lemma 11.7.7. *For any integer $y_0 \geq 4$, if $p(x, y_0)$ is reducible over $\mathbb{Q}$, then the roots of $p(x, y_0)$ satisfy the lattice condition.*

Proof. Let $q(x) = p(x, y_0)$ for a fixed integer $y_0 \geq 4$. Suppose that $q(x) = f(x)g(x)$, where $f(x), g(x) \in \mathbb{Q}[x]$ are monic polynomials of degree at least 1. By Lemma 11.7.6, the degree of each factor must be at least 2. Then without loss of generality, let $f(x)$ and $g(x)$ be quadratic and cubic polynomials respectively, both of which are irreducible over $\mathbb{Q}$. By Gauss' Lemma, we can further assume $f(x), g(x) \in \mathbb{Z}[x]$.

Let $\mathbb{Q}_f$ and $\mathbb{Q}_g$ denote the splitting fields over $\mathbb{Q}$ of f and g respectively. Suppose α, β are the roots of $f(x)$ and γ, δ, ϵ are the roots of $g(x)$. Of course none of these roots are 0. Suppose there exist $i, j, k, m, n \in \mathbb{Z}$ such that

$$\alpha^i \beta^j = \gamma^k \delta^m \epsilon^n \qquad \text{and} \qquad i + j = k + m + n. \tag{11.7.14}$$

We want to show that $i = j = k = m = n = 0$.

We first show that if $i = j$ and $k = m = n$, then $i = j = k = m = n = 0$. By (11.7.14), we have $(\alpha\beta)^i = (\gamma\delta\epsilon)^k$ and $2i = 3k$. Suppose $i \neq 0$, then also $k \neq 0$. We can write $i = 3t$ and $k = 2t$ for some nonzero $t \in \mathbb{Z}$. Let $A = \alpha\beta$ and $B = \gamma\delta\epsilon$. Then both A and B are integers and $AB = y_0^3$. From $A^{3t} = B^{2t}$, we have $A^3 = \pm B^2$. Then $y_0^6 = A^2B^2 = \pm A^5$, and since $y_0 > 3$, there is a nonzero integer $s > 1$ such that $y_0 = s^5$. This implies $A = \pm s^6$ and $B = \pm s^9$ (with the same $\pm$ sign). Then $f(x) = x^2 + c_1 x \pm s^6$, $g(x) = x^3 + c_2' x^2 + c_1' x \pm s^9$, and $q(x) = x^5 - (2s^5 + 1)x^3 - (s^{10} + 2)x^2 + s^5(s^5 - 1)x + s^{15}$. We consider the coefficient of x in

$q(x) = f(x)g(x)$. This is $s^{10} - s^5 = \pm c_1' s^6 \pm c_1 s^9$. Since $s > 1$, there is a prime p such that $p^u \mid s$ and $p^{u+1} \not| s$, for some $u \geq 1$. But then p^{6u} divides $s^5 = s^{10} \pm c_1' s^6 \pm c_1 s^9$. This is a contradiction. Hence $i = j$ and $k = m = n$ imply $i = j = k = m = n = 0$.

Now we claim that $\omega = \alpha/\beta$ is not a root of unity. For a contradiction, suppose that ω is a primitive dth root of unity. Since $\omega \in \mathbb{Q}_f$, which is a degree 2 extension over $\mathbb{Q}$, we have $\phi(d) \mid 2$, where $\phi(\cdot)$ is Euler's totient function. Hence $d \in \{1, 2, 3, 4, 6\}$. The quadratic polynomial $f(x)$ has the form $x^2 - (1+\omega)\beta x + \omega\beta^2 \in \mathbb{Z}[x]$. Hence $r = \frac{(1+\omega)}{\omega\beta} \in \mathbb{Q}$. We prove the claim separately according to whether $r = 0$ or not.

If $r = 0$, then $\omega = -1$ and $d = 2$. In this case, $f(x)$ has the form $x^2 + a$ for some $a \in \mathbb{Z}$. It is easy to check that $q(x)$ has no such polynomial factor in $\mathbb{Z}[x]$ unless $y_0 = 0$. In fact, suppose $x^2 + a \mid q(x)$ in $\mathbb{Z}[x]$. Then $q(x) = (x^2 + a)(x^3 + bx + c)$ since the coefficient of x^4 in $q(x)$ is 0. Also $a + b = -(2y_0 + 1)$, $c = -(y_0^2 + 2)$, $ab = y_0(y_0 - 1)$ and $ac = y_0^3$. It follows that a and b are the two roots of the quadratic polynomial $X^2 + (2y_0 + 1)X + y_0^2 - y_0 \in \mathbb{Z}[X]$. Since $a, b \in \mathbb{Z}$, the discriminant $8y_0 + 1$ must be a perfect square, and in fact an odd perfect square $(2z - 1)^2$ for some $z \in \mathbb{Z}$. Thus $y_0 = z(z-1)/2$. By the quadratic formula, $a = -y_0 + z - 1$ or $-y_0 - z$. On the other hand, $a = ac/c = -y_0^3/(y_0^2 + 2)$. In both cases, this leads to a polynomial in z in $\mathbb{Z}[z]$ that has no integer solutions other than $z = 0$, which gives $y_0 = 0$.

Now suppose $r \neq 0$. Plugging r back in $f(x)$, we have $f(x) = x^2 - (2 + \omega + \omega^{-1})r^{-1}x + (2 + \omega + \omega^{-1})r^{-2}$. The quantity $2 + \omega + \omega^{-1} = 4, 1, 2, 3$ when $d = 1, 3, 4, 6$ respectively. Since $(2 + \omega + \omega^{-1})r^{-2} \in \mathbb{Z}$, the rational number r^{-1} must be an integer when $d = 3, 4, 6$ and half an integer when $d = 1$. In all cases, it is easy to check that a polynomial $f(x)$ of the specified form does not divide $q(x)$ unless $y = 0$ or $y = -1$. Thus, we have proved the claim that $\omega = \alpha/\beta$ is not a root of unity.

Next consider the case that $f(x)$ is irreducible over $\mathbb{Q}_g$. Let E be the splitting field of f over $\mathbb{Q}_g$. Then $[E : \mathbb{Q}_g] = 2$. Therefore, there exists an automorphism $\tau \in \text{Gal}(E/\mathbb{Q}_g)$ that swaps α and β but fixes $\mathbb{Q}_g$ and thus fixes γ, δ, ϵ pointwise. By applying τ to (11.7.14), we have $\alpha^j\beta^i = \gamma^k\delta^m\epsilon^n$. Dividing by (11.7.14) gives $(\alpha/\beta)^{j-i} = 1$. Since α/β is not a root of unity, we get $i = j$. Hence we have $(\alpha\beta)^i = \gamma^k\delta^m\epsilon^n$. The order of $\text{Gal}(\mathbb{Q}_g/\mathbb{Q})$ is $[\mathbb{Q}_g : \mathbb{Q}]$, which is divisible by 3. Thus

$\mathrm{Gal}(\mathbb{Q}_g/\mathbb{Q}) \subseteq S_3$ contains an element of order 3, which must act as a 3-cycle on γ, δ, ϵ. Since $\alpha\beta \in \mathbb{Q}$, applying this cyclic permutation gives $(\alpha\beta)^i = \gamma^m\delta^n\epsilon^k$. Therefore $\gamma^{k-m}\delta^{m-n}\epsilon^{n-k} = 1$. Notice that $(k-m)+(m-n)+(n-k) = 0$.

It can be directly checked that $q(x)$ is not divisible by any $x^3 + c \in \mathbb{Z}[x]$, and therefore by Lemma 11.3.4, the roots γ, δ, ϵ of the cubic polynomial $g(x)$ satisfy the lattice condition. Therefore, $k = m = n$. Again, we have shown that $i = j$ and $k = m = n$ imply $i = j = k = m = n = 0$.

The last case is when $f(x)$ splits in $\mathbb{Q}_g[x]$. Then $\mathbb{Q}_f$ is a subfield of $\mathbb{Q}_g$, and $2 = [\mathbb{Q}_f : \mathbb{Q}] | [\mathbb{Q}_g : \mathbb{Q}]$. Therefore $[\mathbb{Q}_g : \mathbb{Q}] = 6$ and $\mathrm{Gal}(\mathbb{Q}_g/\mathbb{Q}) = S_3$. Since $\mathbb{Q}_f$ is normal over $\mathbb{Q}$, being a splitting field of a separable polynomial in characteristic 0, by the fundamental theorem of Galois theory, the corresponding subgroup for $\mathbb{Q}_f$ is $\mathrm{Gal}(\mathbb{Q}_g/\mathbb{Q}_f)$, which is a normal subgroup of S_3 with index 2. Such a subgroup of S_3 is unique, namely A_3. In particular, the transposition τ' that swaps γ and δ but fixes ϵ is an element in $\mathrm{Gal}(\mathbb{Q}_g/\mathbb{Q}) = S_3$ but not in $\mathrm{Gal}(\mathbb{Q}_g/\mathbb{Q}_f) = A_3$. This transposition must fix α and β setwise but not pointwise. Hence, it must swap α and β.

By applying τ' to (11.7.14), we have $\alpha^j\beta^i = \gamma^m\delta^k\epsilon^n$. Then dividing these two equations gives $(\alpha/\beta)^{i-j} = (\delta/\gamma)^{m-k}$. Similarly, by considering the transposition that switches γ and ϵ and fixes δ, we get $(\alpha/\beta)^{i-j} = (\gamma/\epsilon)^{k-n}$. By combining these two equations, we have $\gamma^{n-m}\delta^{m-k}\epsilon^{k-n} = 1$. Note that $(n-m)+(m-k)+(k-n) = 0$.

As we noted above, the roots of the irreducible $g(x)$ satisfy the lattice condition, so we conclude that $k = n = m$. From $(\alpha/\beta)^{i-j} = (\delta/\gamma)^{m-k} = 1$, we get $i = j$ since α/β is not a root of unity. We conclude that $i = j = k = m = n = 0$, so the roots of $q(x)$ satisfy the lattice condition. □

Even though $p(x,3) = (x-3)(x^4 + 3x^3 + 2x^2 - 5x - 9)$ is reducible, its roots still satisfy the lattice condition. To show this, we use two standard results from Algebra, Theorem 11.7.8 and Lemma 11.7.9. The first is a well-known theorem of Dedekind.

Theorem 11.7.8 (Theorem 4.37 [84])**.** *Suppose $f(x) \in \mathbb{Z}[x]$ is a monic polynomial of degree n. For a prime p, let $f_p(x)$ be the corresponding polynomial in $\mathbb{Z}_p[x]$. If $f_p(x)$ has distinct roots and factors over $\mathbb{Z}_p[x]$ as a product of irreducible factors with degrees $d_1, d_2, \ldots, d_r$, then the Galois group of f over $\mathbb{Q}$ contains an element with cycle type $(d_1, d_2, \ldots, d_r)$.*

With the second result, we can show that $x^4 + 3x^3 + 2x^2 - 5x - 9$ has Galois group S_4 over $\mathbb{Q}$.

Lemma 11.7.9 (Lemma on page 98 in [69]). *For $n \geq 2$, let G be a subgroup of S_n. If G is transitive, contains a transposition, and contains a p-cycle for some prime $p > n/2$, then $G = S_n$.*

Theorem 11.7.10. *The roots of $p(x, 3) = (x - 3)(x^4 + 3x^3 + 2x^2 - 5x - 9)$ satisfy the lattice condition.*

Proof. Let $f(x) = x^4 + 3x^3 + 2x^2 - 5x - 9$ and let G_f be the Galois group of f over $\mathbb{Q}$. We claim that $G_f = S_4$. As a polynomial over $\mathbb{Z}_5$, $f(x) \equiv x^4 + 3x^3 + 2x^2 + 1$ is irreducible, so $f(x)$ is also irreducible over $\mathbb{Z}$. By Gauss' Lemma, this implies irreducibility over $\mathbb{Q}$. Over $\mathbb{Z}_{13}$, $f(x)$ factors into the product of irreducibles $(x^2 + 7)(x + 6)(x + 10)$ and clearly has distinct roots, so by Theorem 11.7.8, G_f contains a transposition. Over $\mathbb{Z}_3$, $f(x)$ factors into the product of irreducibles $x(x^3 + 2x + 1)$ and has distinct roots because its discriminant is $1 \not\equiv 0 \pmod 3$, so by Theorem 11.7.8, G_f contains a 3-cycle. Then by Lemma 11.7.9, $G_f = S_4$.

Let $\alpha, \beta, \gamma, \delta$ be the roots of $f(x)$. Suppose there exist $i, j, k, \ell, n \in \mathbb{Z}$ satisfying $n = i + j + k + \ell$ such that $3^n = \alpha^i \beta^j \gamma^k \delta^\ell$. Now $G_f = S_4$ contains the 4-cycle $(1\ 2\ 3\ 4)$ that cyclically permutes the roots of $f(x)$ but fixes $\mathbb{Q}$. We apply it zero, one, two, and three times to get

$$\begin{aligned} 3^n &= \alpha^i \beta^j \gamma^k \delta^\ell, \\ &= \beta^i \gamma^j \delta^k \alpha^\ell, \\ &= \gamma^i \delta^j \alpha^k \beta^\ell, \text{ and} \\ &= \delta^i \alpha^j \beta^k \gamma^\ell. \end{aligned}$$

Then $3^{4n} = (\alpha\beta\gamma\delta)^{i+j+k+\ell} = (-9)^{i+j+k+\ell}$. Since $n = i + j + k + \ell$, this can only hold when $n = 0$.

Thus, it suffices to show that the roots of $f(x)$ satisfy the lattice condition. By the contrapositive of Lemma 5.5.5, the roots of $f(x)$ do not all have the same complex norm. Then we are done by Lemma 11.3.5. ∎

From Lemma 11.7.5, Lemma 11.7.7, and Theorem 11.7.10, we obtain the following Theorem.

Theorem 11.7.11. *For any integer $y_0 \geq 3$, the roots of $p(x, y_0)$ satisfy the lattice condition.*

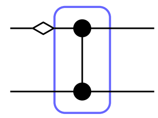


Figure 11.10: Quaternary gadget used in the interpolation construction below. All vertices are assigned $\langle 3(\kappa-1), \kappa-3, -3\rangle$.

We use Theorem 11.7.11 to prove Lemma 11.7.12.

Lemma 11.7.12. *Suppose $\kappa \geq 4$ is the domain size. Then* $\operatorname{Pl-Holant}_\kappa(\langle 3(\kappa-1), \kappa-3, -3\rangle)$ *is* #P*-hard.*

Proof. Let $\langle 2,0,1,0,0,0,1,0,0\rangle$ be a succinct quaternary signature of type τ_4. We reduce from $\operatorname{Pl-Holant}_\kappa(\langle 2,0,1,0,0,0,1,0,0\rangle)$, which is #P-hard by Corollary 11.5.2.

Consider the gadget in Figure 11.10. We assign $\langle 3(\kappa-1), \kappa-3, -3\rangle$ to the vertices. By Lemma 9.2.3, the signature of this gadget is $f = \langle f_{\substack{1\,1\\1\,1}}, f_{\substack{1\,2\\1\,1}}, f_{\substack{1\,2\\1\,2}}, f_{\substack{1\,3\\1\,2}}, f_{\substack{1\,2\\2\,1}}, f_{\substack{1\,3\\2\,1}}, f_{\substack{1\,1\\2\,2}}, f_{\substack{1\,1\\2\,3}}, f_{\substack{1\,4\\2\,3}}\rangle$ up to a nonzero factor of κ, where

$$
\begin{aligned}
f_{\substack{1\,1\\1\,1}} &= (\kappa-1)(\kappa+3),\\
f_{\substack{1\,2\\1\,1}} &= \kappa-3,\\
f_{\substack{1\,2\\1\,2}} &= 2\kappa-3,\\
f_{\substack{1\,3\\1\,2}} &= \kappa-3,\\
f_{\substack{1\,2\\2\,1}} &= 2\kappa-3,\\
f_{\substack{1\,3\\2\,1}} &= \kappa-3,\\
f_{\substack{1\,1\\2\,2}} &= (\kappa-3)(\kappa+1),\\
f_{\substack{1\,1\\2\,3}} &= \kappa-3, \text{ and}\\
f_{\substack{1\,4\\2\,3}} &= -3.
\end{aligned}
$$

Now consider the recursive construction in Figure 11.11. We assign f to every vertex. Up to a nonzero factor of κ^s, let g_s be the succinct signature of type τ_4 for the sth gadget in this construction. Then $g_0 = \langle 1,0,0,0,0,0,1,0,0\rangle$ and $g_s = M^s g_0$, where M is the matrix in Table 11.3.

Table 11.3: Recurrence matrix for the recursive construction in the proof of Lemma 11.7.12.

$$\begin{bmatrix}
(\kappa-1)\left(\kappa^2+9\kappa-9\right) & 12(\kappa-3)(\kappa-1)^2 & (\kappa-3)^2(\kappa-1) & 2(\kappa-3)^2(\kappa-2)(\kappa-1) & (\kappa-3)^2(\kappa-1) & 2(\kappa-3)^2(\kappa-2)(\kappa-1) & (\kappa-1)(2\kappa-3)(4\kappa-3) & 6(\kappa-3)(\kappa-2)(\kappa-1)^2 & (\kappa-3)^3(\kappa-2)(\kappa-1) \\
3(\kappa-3)(\kappa-1) & 3\kappa^3-28\kappa^2+60\kappa-36 & -(\kappa-3)(2\kappa-3) & -2(\kappa-3)(\kappa-2)(2\kappa-3) & -(\kappa-3)(2\kappa-3) & -2(\kappa-3)(\kappa-2)(2\kappa-3) & 3(\kappa-3)(\kappa-1)^2 & (\kappa-2)\left(\kappa^3-14\kappa^2+30\kappa-18\right) & -(\kappa-3)^2(\kappa-2)(2\kappa-3) \\
(2\kappa-3)(4\kappa-3) & 12(\kappa-3)(\kappa-1)^2 & (\kappa-3)^2(\kappa-1) & 2(\kappa-3)^2(\kappa-2)(\kappa-1) & (\kappa-3)^2(\kappa-1) & 2(\kappa-3)^2(\kappa-2)(\kappa-1) & 9\kappa^3-26\kappa^2+27\kappa-9 & 6(\kappa-3)(\kappa-2)(\kappa-1)^2 & (\kappa-3)^3(\kappa-2)(\kappa-1) \\
3(\kappa-3)(\kappa-1) & 2\left(\kappa^3-14\kappa^2+30\kappa-18\right) & -(\kappa-3)(2\kappa-3) & -2(\kappa-3)(\kappa-2)(2\kappa-3) & -(\kappa-3)(2\kappa-3) & -2(\kappa-3)(\kappa-2)(2\kappa-3) & 3(\kappa-3)(\kappa-1)^2 & (\kappa-3)\left(\kappa^3-12\kappa^2+22\kappa-12\right) & -(\kappa-3)^2(\kappa-2)(2\kappa-3) \\
(\kappa-3)^2 & -4(\kappa-3)(2\kappa-3) & 3(\kappa-3) & 6(\kappa-3)(\kappa-2) & \kappa^3+3\kappa-9 & 6(\kappa-3)(\kappa-2) & (\kappa-3)^2(\kappa-1) & -2(\kappa-3)(\kappa-2)(2\kappa-3) & 3(\kappa-3)^2(\kappa-2) \\
(\kappa-3)^2 & -4(\kappa-3)(2\kappa-3) & 3(\kappa-3) & 6(\kappa-3)(\kappa-2) & 3(\kappa-3) & \kappa^3+6\kappa^2-30\kappa+36 & (\kappa-3)^2(\kappa-1) & -2(\kappa-3)(\kappa-2)(2\kappa-3) & 3(\kappa-3)^2(\kappa-2) \\
(\kappa-3)^2 & -4(\kappa-3)(2\kappa-3) & \kappa^3+3\kappa-9 & 6(\kappa-3)(\kappa-2) & 3(\kappa-3) & 6(\kappa-3)(\kappa-2) & (\kappa-3)^2(\kappa-1) & -2(\kappa-3)(\kappa-2)(2\kappa-3) & 3(\kappa-3)^2(\kappa-2) \\
(\kappa-3)^2 & -4(\kappa-3)(2\kappa-3) & 3(\kappa-3) & \kappa^3+6\kappa^2-30\kappa+36 & 3(\kappa-3) & 6(\kappa-3)(\kappa-2) & (\kappa-3)^2(\kappa-1) & -2(\kappa-3)(\kappa-2)(2\kappa-3) & 3(\kappa-3)^2(\kappa-2) \\
(\kappa-3)^2 & -4(\kappa-3)(2\kappa-3) & 3(\kappa-3) & 6(\kappa-3)(\kappa-2) & 3(\kappa-3) & 6(\kappa-3)(\kappa-2) & (\kappa-3)^2(\kappa-1) & -2(\kappa-3)(\kappa-2)(2\kappa-3) & (2\kappa-3)\left(2\kappa^2-9\kappa+18\right)
\end{bmatrix}$$

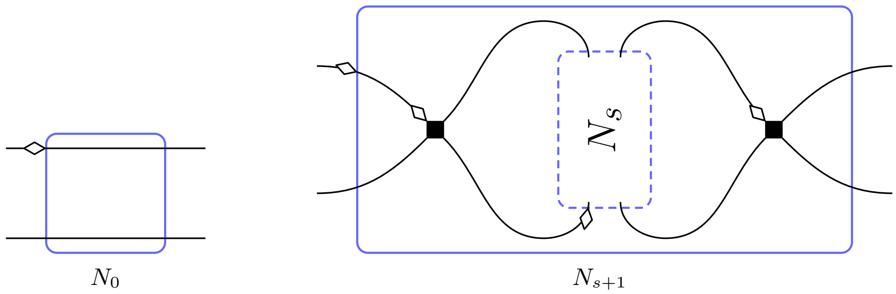


Figure 11.11: Recursive construction to interpolate the weighted Eulerian partition signature. The vertices are assigned the signature of the gadget in Figure 11.10.

The row vectors

$$\begin{gathered}(0,0,0,0,-1,0,0,0,1),\\(0,-1,0,1,-1,0,0,1,0),\\(-1,0,1,0,-1,0,1,0,0),\ \text{and}\\(0,0,0,0,-1,1,0,0,0)\end{gathered}$$

are linearly independent row eigenvectors of M, all with eigenvalue κ^3, that are orthogonal to the initial signature g_0. Note that our target signature $\langle 2,0,1,0,0,0,1,0,0\rangle$ is also orthogonal to these four row eigenvectors.

Up to a factor of $(x-\kappa^3)^4$, the characteristic polynomial of M is

$$h(x,\kappa) = x^5 - \kappa^6(2\kappa-1)x^3 - \kappa^9(\kappa^2-2\kappa+3)x^2 + (\kappa-2)(\kappa-1)\kappa^{12}x + (\kappa-1)^3\kappa^{15}.$$

Since $h(\kappa^3,\kappa) = (\kappa-3)\kappa^{17}$ and $\kappa \geq 4$, we know that κ^3 is not a root of $h(x,\kappa)$ as a polynomial in x. Thus, none of the remaining eigenvalues are κ^3. The roots of $h(x,\kappa)$ satisfy the lattice condition iff the roots of

$$\tilde{h}(x,\kappa) = \frac{1}{\kappa^{15}} h(\kappa^3 x,\kappa) = x^5 - (2\kappa-1)x^3 - (\kappa^2-2\kappa+3)x^2 + (\kappa-2)(\kappa-1)x + (\kappa-1)^3$$

satisfy the lattice condition. In $\tilde{h}(x,\kappa)$, we replace κ by $y+1$ to get $p(x,y) = x^5 - (2y+1)x^3 -$

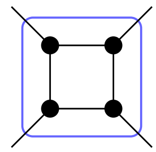


Figure 11.12: Square gadget used to construct the weighted Eulerian partition signature.

$(y^2+2)x^2+(y-1)yx+y^3$. By Theorem 11.7.11, the roots $p(x,y_0)$ satisfy the lattice condition for any integer $y_0 \geq 3$. Thus, the roots of $\tilde{h}(x,\kappa)$ satisfy the lattice for any $\kappa \geq 4$. In particular, this means that the five eigenvalues of M different from κ^3 are distinct, so M is diagonalizable.

The 5-by-5 matrix in the upper-left corner of $[g_0 \; Mg_0 \; \dots \; M^8 g_0]$ is

$$\begin{bmatrix} 1 & 9(\kappa-1)^2\kappa & (\kappa-1)\kappa^4\left(\kappa^3-3\kappa^2+11\kappa+3\right) & (\kappa-1)\kappa^7\left(\kappa^3+12\kappa^2-11\kappa+6\right) & (\kappa-1)\kappa^{10}\left(\kappa^4+4\kappa^3-4\kappa^2+44\kappa-33\right) \\ 0 & 3(\kappa-3)(\kappa-1)\kappa & -(\kappa-3)\kappa^4\left(\kappa^2-2\kappa-1\right) & (\kappa-3)\kappa^7\left(3\kappa^2-3\kappa+2\right) & (\kappa-3)\kappa^{10}\left(\kappa^3-4\kappa^2+16\kappa-11\right) \\ 0 & 9(\kappa-1)^2\kappa & \kappa^4\left(\kappa^4-4\kappa^3+6\kappa^2+4\kappa-3\right) & \kappa^7\left(15\kappa^3-28\kappa^2+11\kappa-6\right) & \kappa^{10}\left(\kappa^5+3\kappa^4-22\kappa^3+72\kappa^2-83\kappa+33\right) \\ 0 & 3(\kappa-3)(\kappa-1)\kappa & -(\kappa-3)(\kappa-1)\kappa^4(\kappa+1) & 2(\kappa-3)\kappa^7\left(2\kappa^2-\kappa+1\right) & (\kappa-3)(\kappa-1)\kappa^{10}\left(\kappa^2-6\kappa+11\right) \\ 0 & (\kappa-3)^2\kappa & (\kappa-3)\kappa^4(\kappa+1) & (\kappa-3)\kappa^7\left(\kappa^2-\kappa+2\right) & (\kappa-3)\kappa^{10}\left(\kappa^3-2\kappa^2+10\kappa-11\right) \end{bmatrix}.$$

Its determinant is $(\kappa-3)^3(\kappa-1)^2\kappa^{26}(\kappa^4+\kappa^3+17\kappa^2+3\kappa+2)$, which is nonzero since $\kappa \geq 4$. Thus $[g_0 \; Mg_0 \; \dots \; M^8 g_0]$ has rank at least 5, so by Lemma 11.3.2, g_0 is not orthogonal to the five remaining row eigenvectors of M.

Therefore, by Lemma 11.3.6, we can interpolate $\langle 2,0,1,0,0,0,1,0,0\rangle$, so we are done. □

When $\kappa = 3$, $\langle 3(\kappa-1), \kappa-3, -3\rangle$ simplifies to $-3\langle -2,0,1\rangle$. We have a much simpler proof that this signature is #P-hard.

Lemma 11.7.13. *Suppose the domain size is* 3. *Then* Pl-Holant$_3(\langle -2,0,1\rangle)$ *is* #P*-hard.*

Proof. Let $g = \langle 2,0,1,0,0,0,1,0\rangle$ be a succinct quaternary signature of type τ_4. We reduce from Pl-Holant$_3(g)$, which is #P-hard by Corollary 11.5.2.

Consider the gadget in Figure 11.12. The vertices are assigned $\langle -2,0,1\rangle$. Up to a factor of 9, the signature of this gadget is g, as desired. □

We summarize this section with the following result. With all succinct binary signatures of type τ_2 available as well as the succinct unary signature $\langle 1\rangle$ of type τ_1, any succinct ternary signature $\langle a,b,c\rangle$ of type τ_3 satisfying $\mathfrak{B} \neq 0$ is #P-hard.

Lemma 11.7.14. *Suppose $\kappa \ge 3$ is the domain size and $a, b, c \in \mathbb{C}$. Let $\mathcal{F}$ be a signature set containing the succinct ternary signature $\langle a, b, c\rangle$ of type τ_3, the succinct unary signature $\langle 1\rangle$ of type τ_1, and the succinct binary signature $\langle x, y\rangle$ of type τ_2 for all $x, y \in \mathbb{C}$. If $\mathfrak{B} \neq 0$, then* Pl-Holant$_\kappa(\mathcal{F})$ *is* #P*-hard.*

Proof. Suppose $\mathfrak{A} \neq 0$. By Lemma 11.7.1, we have a succinct ternary signature $\langle a', b', b'\rangle$ of type τ_3 with $a' \neq b'$. Then we are done by Corollary 10.3.7.

Otherwise, $\mathfrak{A} = 0$. Since $\mathfrak{B} \neq 0$, we have $b \neq c$. By Lemma 11.7.3, we have the signature $\langle 3(\kappa-1), \kappa-3, -3\rangle$. If $\kappa \ge 4$, then we are done by Lemma 11.7.12. Otherwise, $\kappa = 3$ and we are done by Lemma 11.7.13. □

11.8 Main Result

Now we can prove our main dichotomy theorem.

Theorem 11.8.1. *Suppose $\kappa \ge 3$ is the domain size and $a, b, c \in \mathbb{C}$. Let $\langle a, b, c\rangle$ be a succinct ternary signature of type τ_3. Then* Pl-Holant$_k$ *appa$(\langle a, b, c\rangle)$ is* #P*-hard unless at least one of the following holds:*

1. *$a = b = c$;*
2. *$a = c$ and $\kappa = 3$;*

there exists $\lambda \in \mathbb{C}$ and $T \in \{I_\kappa, \kappa I_\kappa - 2J_\kappa\}$ such that

3. *$\langle a, b, c\rangle = T^{\otimes 3}\lambda\langle 1, 0, 0\rangle$;*
4. *$\langle a, b, c\rangle = T^{\otimes 3}\lambda\langle 1, 0, \omega\rangle$ and $\kappa = 3$ where $\omega^3 = 1$;*
5. *$\langle a, b, c\rangle = T^{\otimes 3}\lambda\langle \mu^2, 1, \mu\rangle$ and $\kappa = 4$ where $\mu = -1 \pm 2i$;*

in which case, the computation can be done in polynomial time.

Proof. The signature in case 1 is degenerate, which is trivially tractable. Case 2 is tractable by case 3 of Corollary 11.1.3. Case 3 is tractable by Corollary 11.1.4. Case 4 is tractable by Corollary 11.1.5. Case 5 is tractable by Lemma 11.1.6.

Otherwise, $\langle a, b, c\rangle$ is none of these tractable cases. If $\mathfrak{B} = 0$, then we are done by Corollary 11.5.5, so assume that $\mathfrak{B} \neq 0$. If $a + (\kappa-1)b = 0$ and $b^2 - 4bc - (\kappa-3)c^2 = 0$, then we are done by Lemma 11.5.3, so assume that $a + (\kappa - 1)b \neq 0$ or $b^2 - 4bc - (\kappa - 3)c^2 \neq 0$.

If $a + (\kappa - 1)b \neq 0$, then we have the succinct unary signature $\langle 1 \rangle$ of type τ_1 by Lemma 11.5.1. Otherwise, $a + (\kappa - 1)b = 0$ and $b^2 - 4bc - (\kappa - 3)c^2 \neq 0$. Since $\mathfrak{B} \neq 0$, we have $2b + (\kappa - 2)c \neq 0$. Then again we have $\langle 1 \rangle$ by Lemma 11.5.1. Thus, in either case, we have $\langle 1 \rangle$.

By Corollary 11.6.14, we have all binary succinct signatures $\langle x, y \rangle$ for any $x, y \in \mathbb{C}$. Then we are done by Lemma 11.7.14. □

11.9 Closing Thoughts

Lattice condition with two terms When $\ell = 2$, Definition 11.3.3, the lattice condition, says that $\frac{\lambda_1}{\lambda_2}$ is not a root of unity. It is often difficult to satisfy the lattice condition in this case. See, for example, the proof of Lemma E.2 in [27]. Instead, it is often easier to show that the norm of $\frac{\lambda_1}{\lambda_2}$ is not 1.

To compensate for this weaker conclusion though, one can try a few recursive gadget constructions, each of which define their own λ_1 and λ_2 values (cf. Section 6.2 and Section 11.3), and consider when all ratios are of norm 1. Specifically, it often suffices to consider three recursive constructions. See, for example, the applications of Lemma 11.6.2, Lemma 11.6.3, and Lemma 11.6.11.

Further generalizing interpolation I am quite satisfied with the level of generality in the interpolation result of Lemma 11.3.6. However, there is still room for improvement.

The first improvement is simple. Intuitively, in the proof of Lemma 11.3.6, the condition that s is orthogonal to a certain row eigenvector r is functionally equivalent to r having eigenvalue 0. Thus, condition 2 can be relaxed to

> s is not orthogonal to exactly ℓ of these linearly independent row eigenvectors of M with *nonzero* eigenvalues $\lambda_1, \ldots, \lambda_\ell$

with the corresponding conclusion being

> for any succinct signature f of type τ that is orthogonal to the $n - \ell$ of these linearly independent eigenvectors of M to which *either* s is also orthogonal *or the corresponding*

eigenvalue is 0.

Note that the changes here are only the addition of text, which have been emphasized.

The second improvement is more complicated. For simplicity, consider the case where s is orthogonal to every row eigenvector. Then the three conditions imply that all eigenvalues are distinct. However, this is not necessary. If some eigenvalues are the same, then the same conclusion might hold if their eigenvectors are linearly dependent (i.e. the eigenvalues are in the same Jordan block). For example, see Section 6.2, especially Lemma 6.2.4 and the paragraph that precedes it. I am not sure what the proper generalization should be; I leave this as an open problem.

Even with these generalizations in mind, I consider Lemma 11.3.6 to be sufficiently strong for now. Instead, I believe that an important tool for future progress on the complexity of Holant problems will be new ways to satisfy condition 3, the lattice condition. The proof of Lemma E.2 in [27] is already one such example.

Computing Puiseux series The proof of Lemma 11.7.6 used (truncations of) the Puiseux series

$$\begin{aligned}
y_1(x) &= x^2 + 2x^{-1} + 2x^{-2} - 6x^{-4} - 18x^{-5} + O(x^{-6}) && \text{for } x \in \mathbb{R},\\
y_2(x) &= x^{3/2} - \frac{1}{2}x + \frac{1}{8}x^{1/2} - \frac{65}{128}x^{-1/2} - x^{-1} - \frac{1471}{1024}x^{-3/2} - x^{-2} + O(x^{-5/2}) && \text{for } x > 0, \text{ and}\\
y_3(x) &= -x^{3/2} - \frac{1}{2}x - \frac{1}{8}x^{1/2} + \frac{65}{128}x^{-1/2} - x^{-1} + \frac{1471}{1024}x^{-3/2} - x^{-2} + O(x^{-5/2}) && \text{for } x > 0.
\end{aligned}$$

for

$$p(x,y) = x^5 - 2x^3y - x^3 - x^2y^2 - 2x^2 + xy^2 - xy + y^3.$$

Mathematica can compute these (using SERIES), but let me explain how one can do this by hand.

We want to understand the roots of $p(x,y)$ for large x and y. Since series expansions are simplest when done at 0 (or $(0,0)$ in this case), we invert x and y and clear denominators to get

$$q(x,y) = x^5y^3p(x^{-1},y^{-1}) = x^5 - x^4y^2 + x^4y - 2x^3y^3 - x^3y - x^2y^3 - 2x^2y^2 + y^3.$$

To obtain our initial approximation for each root, we use Newton's polygon. My understanding

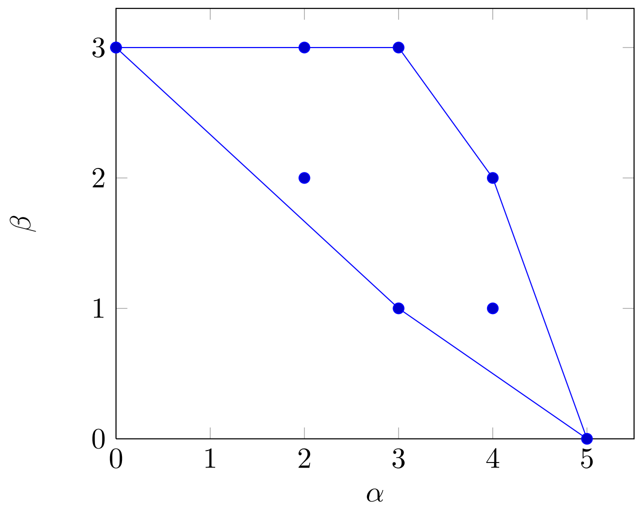


Figure 11.13: Carrier and Newton polygon of q.

for how to do this comes from [11, Section 8.3]. Let f be a polynomial in x and y. The *carrier* of f, denoted by $\Delta(f)$, is the set of pairs of exponents (α, β) for its monomials $c_{\alpha,\beta}x^{\alpha}y^{\beta}$ with nonzero coefficients (i.e. $c_{\alpha,\beta} \neq 0$). Then the Newton polygon of f is the convex hull of $\Delta(f)$.

The carrier of q is

$$\Delta(q) = \{(5,0),(4,2),(4,1),(3,3),(3,1),(2,3),(2,2),(0,3)\},$$

which is plotted in Figure 11.13 along with the Newton polygon of q. We use the lower envelope to obtain the initial approximations.[1] The lower envelope is comprised of the line segment from $(0,3)$ to $(3,1)$ and the line segment from $(3,1)$ to $(5,0)$. The slope of the first line segment is $-\frac{2}{3}$, which says that an initial approximation for a root is $y = tx^{3/2}$ for some $t \in \mathbb{C}$. We solve for t by summing the monomials involved in this line segment and setting them equal to 0. In this case, we get

$$y^3 - x^3y = (tx^{3/2})^3 - x^3(tx^{3/2}) = tx^{9/2}(t^2 - 1) = 0,$$

for which the nonzero solutions are $t = \pm 1$. Therefore, initial approximations for two roots are $y = \pm x^{3/2}$. The slope of the second line segment is $-\frac{1}{2}$, which says that an initial approximation a

[1]Alternatively, we could have used the upper envelope for the Newton polygon of p.

root is $y = tx^2$ for some $t \in \mathbb{C}$. Then

$$x^5 - x^3 y = x^5 - x^3(tx^2) = x^5(t-1) = 0,$$

for which the only solution is $t = 1$. Therefore, a third initial approximation for a root is $y = x^2$.

Now we improve each approximation using Newton's method. For this setting with two variables, Newton's method takes an initial approximation $y_{\text{old}}(x)$ and obtains a better approximation $y_{\text{new}}(x)$ according to the equation

$$y_{\text{new}}(x) = y_{\text{old}}(x) - \frac{p(x, y_{\text{old}}(x))}{\partial_y p(x, y_{\text{old}}(x))}.$$

Since the numerator and denominator in the fraction are polynomials[2] in x, we do not expect to have cancellations in the fraction. However, we are only looking for an approximation, so we can ignore all terms of lower order in the denominator so that the fraction can be expressed with denominator 1. Using the initial approximation $y_{\text{old}}(x) = x^2$ gives

$$\begin{aligned} y_{\text{new}}(x) &= y_{\text{old}}(x) - \frac{p(x, y_{\text{old}}(x))}{\partial_y p(x, y_{\text{old}}(x))} \\ &= y_{\text{old}}(x) - \frac{x^5 - 2x^3 y_{\text{old}}(x) - x^3 - x^2 y_{\text{old}}(x)^2 - 2x^2 + x y_{\text{old}}(x)^2 - x y_{\text{old}}(x) + y_{\text{old}}(x)^3}{-2x^3 - 2x^2 y_{\text{old}}(x) + 2x y_{\text{old}}(x) - x + 3y_{\text{old}}(x)^2} \\ &= x^2 - \frac{-2x^3 - 2x^2}{x^4 - x} \\ &\approx x^2 + 2x^{-1} + 2x^{-2}. \end{aligned}$$

Our initial approximation was the first term of $y_1(x)$, and now we have the first three terms of $y_1(x)$. We knew that we would get the second term from this application of Newton's method, but I would say that we were lucky by also getting the third term. This does not happen in general. For example, applying Newton's method starting with $y_{\text{old}}(x) = x^2 + 2x^{-1}$ (and ignoring terms of lower order in the denominator) gives

$$y_{\text{new}}(x) \approx x^2 + 2x^{-1} + 2x^{-2} - 6x^{-4} - 4x^{-5} - 8x^{-7},$$

[2] Actually, Laurent polynomials in $x^{1/n}$ for some positive integer n.

but the correct coefficient of x^{-5} is 18 (and the coefficients for x^{-6} and x^{-7} are also incorrect). To be certain about each term in the new approximation, I prefer to ignore terms of lower order in the numerator of the fraction as well. Throwing away this information will drastically slow down the convergence rate of Newton's method, but I only want to compute the first few terms, so this is not a problem.

Chapter 12

Conclusion

This dissertation contains dichotomy theorems for several classes of Holant problems, but the complexity remains unknown for many counting problems. One way to make fundamental progress would be to obtain a better understanding of which points satisfy the lattice condition (Definition 11.3.3). For example, it is extremely easy to apply Lemma 11.3.4. Is there a similarly useful characterization of the irreducible quartic polynomials with rational coefficients?

12.1 Future Progress on the Complexity of Holant Problems

The most pertinent question about the complexity of Holant problems involves asymmetric signatures. Symmetric signatures have many nice properties, but they lack the property of being closed under gadget constructions. Thus far, the majority of progress on Holant problems with symmetric signatures has managed to avoid addressing the complexity of Holant problems with asymmetric signatures.[1]

It is natural to consider Holant problems over regular graphs (by, for example, considering a Holant problem in which only a single signature is available). The existing proof techniques also work quite well for planar graphs. (Intuitively, the reason for this is that the simplest gadget constructions are often the most useful in proving the hardness of a dichotomy theorem, and simple gadgets are usually planar.) In contrast, little is known about the complexity of bipartite Holant

[1] A notable exception is Lemma 6.5.2.

problems (since simple gadgets are usually not bipartite).

A tough road lies ahead for anyone who considers Holant problems over higher domain sizes (than the Boolean domain). The success we had in Chapter 10 and Chapter 11 with higher domain Holant problems can be partially attributed to the fact that we considered a single signature with higher degree of symmetry. Here are two specific difficulties. First, it might be that a higher domain Holant problem is properly viewed as a Holant problem with a smaller domain size via a holographic transformation. See the proof of Lemma 11.5.4 for an example of this. Second, higher domain signatures can express tractable signatures that are a sum of tractable signatures. The proof of Corollary 11.5.5 contains a simple example of this. In special cases, this can also happen via a sum of (two) tractable signatures defined over disjoint domain elements. I also expect there to be further difficulties that we have yet to encounter.

12.2 Future Progress on the Complexity of Graph Polynomials

One way I envision progress being made on the complexity of Holant problems is through progress on the complexity of evaluating related graph polynomials. Chief among them is the transition polynomial [85] (for 4-regular graphs) and its generalizations [65, 63] (to Eulerian graphs and all graphs respectively). In particular, Lemma 10.2.2 has been partially generalized by Theorem 4.2 in [62]. Specifically, this result can be used to show (with a proof similar to that of Lemma 10.2.5) that the transition polynomial has a significant overlap with Holant problems. By extending these results to the two generalizations of the transition polynomial, one would find that the overlap with Holant problems increases all the more.

There are a handful of results about the complexity of evaluating graph polynomials. I conclude by listing in Table 12.1 all such results of which I am aware.

Table 12.1: Summary of results about the complexity of evaluating graph polynomials.

Graph Polynomial	Field	Graph Class Restrictions	Complete Dichotomy?	Reference
Tutte	$\mathbb{C}$	-	Yes	[86]
		planar	Yes	[138]
		planar bipartite	Yes	[137]
Weighted Matching	$\mathbb{R}$	-	No	[3]
Bollobás-Riordan	$\mathbb{Q}$	-	No	[8]
Edge Elimination	$\mathbb{Q}$	-	No	[78]
trivariate Ising	$\mathbb{Q}$	simple planar bipartite	No	[90]
Cover	$\mathbb{Q}$	-	Yes	[7]
	$\mathbb{Q}$	Planar	No	[6]
Geometric Cover	$\mathbb{Q}$	-	Yes	[7]
	$\mathbb{Q}$	Planar	Yes	[6]
Domination	$\mathbb{Q}$	-	No	[123]
Interlace	$\mathbb{Q}$	-	No	[9, 10]

Bibliography

[1] Richard Arratia, Béla Bollobás, and Gregory B. Sorkin. The interlace polynomial: A new graph polynomial. In *SODA*, pages 237–245. Society for Industrial and Applied Mathematics, 2000.

[2] J. Ashkin and E. Teller. Statistics of two-dimensional lattices with four components. *Phys. Rev.*, 64(5-6):178–184, 1943.

[3] Ilia Averbouch and J.A.Makowsky. The complexity of multivariate matching polynomials. http://www.cs.technion.ac.il/admlogic/TR/2007/wg07.pdf, 2007.

[4] Régis Barbanchon. On unique graph 3-colorability and parsimonious reductions in the plane. *Theor. Comput. Sci.*, 319(1-3):455–482, 2004.

[5] Rodney J. Baxter. *Exactly solved models in statistical mechanics.* Academic press London, 1982.

[6] Markus Bläser and Radu Curticapean. The complexity of the cover polynomials for planar graphs of bounded degree. In *MFCS*, pages 96–107. Springer Berlin Heidelberg, 2011.

[7] Markus Bläser, Holger Dell, and Mahmoud Fouz. Complexity and approximability of the cover polynomial. *Comput. Complex.*, 21(3):359–419, 2012.

[8] Markus Bläser, Holger Dell, and Johann A. Makowsky. Complexity of the bollobsriordan polynomial. exceptional points and uniform reductions. *Theor. Comput. Syst.*, 46(4):690–706, 2010.

[9] Markus Bläser and Christian Hoffmann. On the complexity of the interlace polynomial. *CoRR*, abs/0707.4565, 2007.

[10] Markus Bläser and Christian Hoffmann. On the complexity of the interlace polynomial. In *STACS*, pages 97–108. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2008.

[11] Egbert Brieskorn and Horst Knörrer. *Plane Algebraic Curves*. Modern Birkhäuser Classics. Springer Basel, 1986. Translated by John Stillwell.

[12] Thomas Brylawski and James Oxley. The Tutte polynomial and its applications. In Neil White, editor, *Matriod Applications*, pages 123–225. Cambridge University Press, 1992.

[13] Andrei Bulatov, Martin Dyer, Leslie Ann Goldberg, Markus Jalsenius, and David Richerby. The complexity of weighted Boolean #CSP with mixed signs. *Theor. Comput. Sci.*, 410(38-40):3949–3961, 2009.

[14] Andrei Bulatov and Martin Grohe. The complexity of partition functions. *Theor. Comput. Sci.*, 348(2):148–186, 2005.

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

[16] Andrei A. Bulatov. The complexity of the counting constraint satisfaction problem. *J. ACM*, 60(5):34:1–34:41, 2013.

[17] Andrei A. Bulatov and Víctor Dalmau. Towards a dichotomy theorem for the counting constraint satisfaction problem. *Inform. and Comput.*, 205(5):651–678, 2007.

[18] Robert Burton and Jeffrey E. Steif. Non-uniqueness of measures of maximal entropy for subshifts of finite type. *Ergodic Theory Dyn. Syst.*, 14(2):213–235, 1994.

[19] Robert Burton and Jeffrey E. Steif. New results on measures of maximal entropy. *Israel J. Math.*, 89(1-3):275–300, 1995.

[20] Jin-Yi Cai and Xi Chen. Complexity of counting CSP with complex weights. In *STOC*, pages 909–920. ACM, 2012.

[21] Jin-Yi Cai, Xi Chen, Richard J. Lipton, and Pinyan Lu. On tractable exponential sums. In *FAW*, pages 148–159. Springer Berlin Heidelberg, 2010. CoRR, abs/1307.7430, 2010.

[22] Jin-Yi Cai, Xi Chen, and Pinyan Lu. Non-negatively weighted #CSP: An effective complexity dichotomy. In *CCC*, pages 45–54. IEEE Computer Society, 2011.

[23] Jin-Yi Cai, Xi Chen, and Pinyan Lu. Graph homomorphisms with complex values: A dichotomy theorem. *SIAM J. Comput.*, 42(3):924–1029, 2013.

[24] Jin-Yi Cai and Vinay Choudhary. Some results on matchgates and holographic algorithms. *Int. J. Software and Informatics*, 1(1):3–36, 2007.

[25] Jin-Yi Cai and Vinay Choudhary. Valiant's Holant theorem and matchgate tensors. *Theor. Comput. Sci.*, 384(1):22–32, 2007.

[26] Jin-Yi Cai, Vinay Choudhary, and Pinyan Lu. On the theory of matchgate computations. *Theory Comput. Syst.*, 45(1):108–132, 2009.

[27] Jin-Yi Cai, Zhiguo Fu, Heng Guo, and Tyson Williams. A holant dichotomy: Is the FKT algorithm universal? In submission, 2015.

[28] Jin-Yi Cai and Aaron Gorenstein. Matchgates revisited. *Theory Comput.*, 10(7):167–197, 2014.

[29] Jin-Yi Cai, Heng Guo, and Tyson Williams. A complete dichotomy rises from the capture of vanishing signatures. *CoRR*, abs/1204.6445, 2012.

[30] Jin-Yi Cai, Heng Guo, and Tyson Williams. A complete dichotomy rises from the capture of vanishing signatures (extended abstract). In *STOC*, pages 635–644. ACM, 2013.

[31] Jin-Yi Cai, Heng Guo, and Tyson Williams. Holographic algorithms beyond matchgates. *CoRR*, abs/1307.7430, 2013.

[32] Jin-Yi Cai, Heng Guo, and Tyson Williams. The complexity of counting edge colorings and a dichotomy for some higher domain Holant problems. *CoRR*, abs/1404.4020, 2014.

[33] Jin-Yi Cai, Heng Guo, and Tyson Williams. The complexity of counting edge colorings and a dichotomy for some higher domain Holant problems (extended abstract). In *FOCS*, pages 601–610. IEEE, 2014.

[34] Jin-Yi Cai, Heng Guo, and Tyson Williams. Holographic algorithms beyond matchgates. In *ICALP*, pages 271–282. Springer Berlin Heidelberg, 2014.

[35] Jin-Yi Cai, Sangxia Huang, and Pinyan Lu. From Holant to #CSP and back: Dichotomy for Holant^c problems. *Algorithmica*, 64(3):511–533, 2012.

[36] Jin-Yi Cai and Michael Kowalczyk. Spin systems on k-regular graphs with complex edge functions. *Theor. Comput. Sci.*, 461:2–16, 2012.

[37] Jin-Yi Cai and Michael Kowalczyk. Partition functions on k-regular graphs with $\{0,1\}$-vertex assignments and real edge functions. *Theor. Comput. Sci.*, 494(0):63–74, 2013.

[38] Jin-Yi Cai, Michael Kowalczyk, and Tyson Williams. Gadgets and anti-gadgets leading to a complexity dichotomy. *CoRR*, abs/1108.3383, 2011.

[39] Jin-Yi Cai, Michael Kowalczyk, and Tyson Williams. Gadgets and anti-gadgets leading to a complexity dichotomy. In *ITCS*, pages 452–467. ACM, 2012.

[40] Jin-Yi Cai and Pinyan Lu. On block-wise symmetric signatures for matchgates. In *FCT*, pages 187–198. Springer Berlin Heidelberg, 2007.

[41] Jin-Yi Cai and Pinyan Lu. Holographic algorithms with unsymmetric signatures. In *SODA*, pages 54–63. Society for Industrial and Applied Mathematics, 2008.

[42] Jin-Yi Cai and Pinyan Lu. On symmetric signatures in holographic algorithms. *Theory Comput. Syst.*, 46(3):398–415, 2010.

[43] Jin-Yi Cai and Pinyan Lu. Holographic algorithms: From art to science. *J. Comput. Syst. Sci.*, 77(1):41–61, 2011.

[44] Jin-Yi Cai and Pinyan Lu. Signature theory in holographic algorithms. *Algorithmica*, 61(4):779–816, 2011.

[45] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Holant problems and counting CSP. In *STOC*, pages 715–724. ACM, 2009.

[46] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Holographic algorithms with matchgates capture precisely tractable planar #CSP. In *FOCS*, pages 427–436. IEEE Computer Society, 2010.

[47] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Computational complexity of Holant problems. *SIAM J. Comput.*, 40(4):1101–1132, 2011.

[48] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. A computational proof of complexity of some restricted counting problems. *Theor. Comput. Sci.*, 412(23):2468–2485, 2011.

[49] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Dichotomy for Holant* problems of Boolean domain. In *SODA*, pages 1714–1728. SIAM, 2011.

[50] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Holographic reduction, interpolation and hardness. *Computational Complexity*, 21(4):573–604, 2012.

[51] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Dichotomy for Holant* problems with domain size 3. In *SODA*, pages 1278–1295. SIAM, 2013.

[52] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Holographic algorithms by Fibonacci gates. *Linear Algebra Appl.*, 438(2):690–707, 2013.

[53] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. The complexity of complex weighted Boolean #CSP. *J. Comput. System Sci.*, 80(1):217–236, 2014.

[54] Matthew Cook and Jehoshua Bruck. Implementability among predicates. Caltech Paradise ETR 067, http://www.paradise.caltech.edu/papers/etr067.pdf, 2005.

[55] Nadia Creignou, Sanjeev Khanna, and Madhu Sudan. *Complexity Classifications of Boolean Constraint Satisfaction Problems*. Society for Industrial and Applied Mathematics, 2001.

[56] Radu Curticapean. Counting perfect matchings in graphs that exclude a single-crossing minor. *CoRR*, abs/1406.4056, 2014.

[57] C. T. J. Dodson and T. Poston. *Tensor Geometry*, volume 130 of *Graduate Texts in Mathematics*. Springer-Verlag, second edition, 1991.

[58] Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. The complexity of weighted Boolean #CSP. *SIAM J. Comput.*, 38(5):1970–1986, 2009.

[59] Martin Dyer and Catherine Greenhill. The complexity of counting graph homomorphisms. *Random Struct. Algorithms*, 17(3-4):260–289, 2000.

[60] Martin Dyer and David Richerby. On the complexity of #CSP. In *STOC*, pages 725–734. ACM, 2010.

[61] Jack Edmonds. Paths, trees, and flowers. *Canad. J. Math.*, 17:449–467, 1965.

[62] J. Ellis-Monaghan and I. Moffatt. Evaluations of topological Tutte polynomials. *Combin. Probab. Comput.*, 24(3):556–583, 2015.

[63] Joanna A. Ellis-Monaghan. New results for the Martin polynomial. *J. Comb. Theory Ser. B*, 74(2):326–352, 1998.

[64] Joanna A. Ellis-Monaghan. Identities for circuit partition polynomials, with applications to the Tutte polynomial. *Adv. Appl. Math.*, 32(1-2):188–197, 2004.

[65] Joanna A. Ellis-Monaghan and Irasema Sarmiento. Generalized transition polynomials. *Congr. Numer.*, 155:57–69, 2002.

[66] Gerd Faltings. Endlichkeitssätze für abelsche Varietäten über Zahlkörpern. *Invent. Math.*, 73(3):349–366, 1983.

[67] 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.

[68] Michael E. Fisher. On the dimer solution of planar ising models. *J. Math. Phys.*, 7(10):1776–1781, 1966.

[69] P. X. Gallagher. The large sieve and probabilistic Galois theory. In *Proc. Symp. Pure Math.*, volume 24 of *Analytic Number Theory*, pages 91–101. Amer. Math. Soc., 1973.

[70] Leslie Ann Goldberg, Martin Grohe, Mark Jerrum, and Marc Thurley. A complexity dichotomy for partition functions with mixed signs. *SIAM J. Comput.*, 39(7):3336–3402, 2010.

[71] Heng Guo, Sangxia Huang, Pinyan Lu, and Mingji Xia. The complexity of weighted Boolean #CSP modulo k. In *STACS*, pages 249–260. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2011.

[72] Heng Guo, Pinyan Lu, and Leslie G. Valiant. The complexity of symmetric Boolean parity Holant problems. *SIAM J. Comput.*, 42(1):324–356, 2013.

[73] Heng Guo and Tyson Williams. The complexity of planar Boolean #CSP with complex weights. *CoRR*, abs/1212.2284, 2012.

[74] Heng Guo and Tyson Williams. The complexity of planar Boolean #CSP with complex weights. In *ICALP*, pages 516–527. Springer Berlin Heidelberg, 2013.

[75] Per Hallberg. *Gibbs Measures and Phase Transitions in Potts and Beach Models*. PhD thesis, Chalmers Tekniska Högskola, 2004. http://www.dissertations.se/dissertation/9e3d6872b0/.

[76] Johan Håstad. Tensor rank is NP-complete. *J. Algorithm.*, 11(4):644–654, 1990.

[77] Brian Hayes. Accidental algorithms. *American Scientist*, 96(1), 2008.

[78] Christian Hoffmann. A most general edge elimination polynomial - thickening of edges. *Fundam. Inform.*, 98(4):373–378, 2010.

[79] Ian Holyer. The NP-completeness of edge-coloring. *SIAM J. Comput.*, 10(4):718–720, 1981.

[80] Sangxia Huang and Pinyan Lu. A dichotomy for real weighted Holant problems. In *CCC*, pages 96–106. IEEE Computer Society, 2012.

[81] Sangxia Huang and Pinyan Lu. A dichotomy for real weighted Holant problems. Preliminary version appeared at CCC 2012, http://www.csc.kth.se/~sangxia/papers/2012-ccc.pdf, 2013.

[82] Tamás Hubai. The chromatic polynomial. Master's thesis, Eötvös LorÁnd University, 2009. https://www.cs.elte.hu/blobs/diplomamunkak/mat/2009/hubai_tamas.pdf.

[83] Ernst Ising. Beitrag zür theorie des ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, 1925.

[84] Nathan Jacobson. *Basic Algebra I*. W. H. Freeman & Co., second edition, 1985.

[85] F. Jaeger. On transition polynomials of 4-regular graphs. In Geňa Hahn, Gert Sabidussi, and Robert E. Woodrow, editors, *Cycles and Rays*, volume 301 of *NATO ASI Series*, pages 123–150. Springer Netherlands, 1990.

[86] F. Jaeger, D. L. Vertigan, and D. J. A. Welsh. On the computational complexity of the Jones and Tutte polynomials. *Math. Proc. Cambridge*, 108(1):35–53, 1990.

[87] P. W. Kasteleyn. The statistics of dimers on a lattice. *Physica*, 27(12):1209–1225, 1961.

[88] P. W. Kasteleyn. Graph theory and crystal physics. In F. Harary, editor, *Graph Theory and Theoretical Physics*, pages 43–110. Academic Press, London, 1967.

[89] P.W. Kasteleyn. Dimer statistics and phase transitions. *J. Math. Phys.*, 4(2):287–293, 1963.

[90] Tomer Kotek. Complexity of ising polynomials. *Combin. Probab. Comput.*, 21(5):743–772, 2012.

[91] Michael Kowalczyk. Classification of a class of counting problems using holographic reductions. In *COCOON*, pages 472–485. Springer, 2009.

[92] Michael Kowalczyk. *Dichotomy theorems for Holant problems*. PhD thesis, University of Wisconsin—Madison, 2010. http://cs.nmu.edu/~mkowalcz/research/main.pdf.

[93] Michael Kowalczyk and Jin-Yi Cai. Holant problems for regular graphs with complex edge functions. In *STACS*, pages 525–536. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2010.

[94] Michael Kowalczyk and Jin-Yi Cai. Holant problems for regular graphs with complex edge functions. *CoRR*, abs/1001.0464, 2010.

[95] J. M. Landsberg, Jason Morton, and Serguei Norine. Holographic algorithms without matchgates. *Linear Algebra Appl.*, 438(2):782–795, 2013.

[96] J.M. Landsberg. *Tensors: Geometry and Applications*, volume 128 of *Graduate Studies in Mathematics*. American Mathematical Society, 2011.

[97] T. D. Lee and C. N. Yang. Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model. *Phys. Rev.*, 87(3):410–419, 1952.

[98] Gaëtan Leurent. The On-Line Encyclopedia of Integer Sequences. https://oeis.org/A182176, Apr 2012.

[99] Daniel Leven and Zvi Galil. NP completeness of finding the chromatic index of regular graphs. *J. Algorithm.*, 4(1):35–44, 1983.

[100] Aaron Levin. private communication, 2013.

[101] Elliott H. Lieb. Residual entropy of square ice. *Phys. Rev.*, 162(1):162–172, 1967.

[102] Elliott H. Lieb and Alan D. Sokal. A general Lee-Yang theorem for one-component and multicomponent ferromagnets. *Comm. Math. Phys.*, 80(2):153–179, 1981.

[103] Nathan Linial. Hard enumeration problems in geometry and combinatorics. *SIAM J. Alg. Discr. Meth.*, 7(2):331–335, 1986.

[104] Charles H.C. Little. An extension of Kasteleyn's method of enumerating the 1-factors of planar graphs. In Derek A. Holton, editor, *Combinatorial Mathematics, Proceedings of the Second Australian Conference*, volume 403 of *Lecture Notes in Mathematics*, pages 63–72. Springer Berlin Heidelberg, 1974.

[105] Pierre Martin. *Enumérations Eulériennes dans les multigraphes et invariants de Tutte-Grothendieck*. PhD thesis, Joseph Fourier University, 1977. http://tel.archives-ouvertes.fr/tel-00287330.

[106] Jason Morton. Pfaffian circuits. *CoRR*, abs/1101.0129, 2011.

[107] Jason Morton and Susan Margulies. Polynomial-time solvable #CSP problems via algebraic models and Pfaffian circuits. *CoRR*, abs/1311.4066, 2013. To appear in Journal of Symbolic Computation.

[108] Peter Müller. Hilbert's irreducibility theorem for prime degree and general polynomials. *Israel J. Math.*, 109(1):319–337, 1999.

[109] Lars Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev.*, 65(3-4):117–149, 1944.

[110] Julius Petersen. Die theorie der regulären graphs. *Acta Mathematica*, 15(1):193–220, 1891.

[111] Bjorn Poonen. private communication, 2013.

[112] R. B. Potts. Some generalized order-disorder transformations. *Math. Proc. Cambridge*, 48(1):106–109, 1952.

[113] J. Scott Provan and Michael O. Ball. The complexity of counting cuts and of computing the probability that a graph is connected. *SIAM J. Comput.*, 12(4):777–788, 1983.

[114] Guus Regts. *Graph parameters and invariants of the orthogonal group*. PhD thesis, Universiteit van Amsterdam, 2013. http://dare.uva.nl/record/1/399823.

[115] Carl Ludwig Siegel. Über einige anwendungen diophantischer approximationen. *Abh. Pruess. Akad. Wiss. Phys. Math. Kl.*, pages 41–69, 1929.

[116] Ian Stewart. *Galois Theory*. Chapman Hall/CRC Mathematics Series. Taylor & Francis, third edition, 2003.

[117] Michael Stiebitz, Diego Scheide, Bjarne Toft, and Lene M. Favrholdt. *Graph Edge Coloring: Vizing's Theorem and Goldberg's Conjecture*. Wiley, 2012.

[118] Simon Straub, Thomas Thierauf, and Fabian Wagner. Counting the number of perfect matchings in k5-free graphs. In *CCC*, pages 66–77. IEEE Computer Society, 2014.

[119] Peter Tait. Remarks on the colourings of maps. *Proc. R. Soc. Edinburgh*, 10:729, 1880.

[120] H. N. V. Temperley and Michael E. Fisher. Dimer problem in statistical mechanics—an exact result. *Philosophical Magazine*, 6(68):1061–1063, 1961.

[121] Marc Thurley. *The Complexity of Partition Functions*. PhD thesis, Humboldt-Universität zu Berlin, 2009. https://sites.google.com/site/marcthurley/diss_thurley.pdf.

[122] Marc Thurley. The complexity of partition functions on Hermitian matrices. *CoRR*, abs/1004.0992, 2010.

[123] Peter Tittmann Tomer Kotek, James Preen. Domination polynomials of graph products. *CoRR*, abs/1305.1475, 2013.

[124] Salil P. Vadhan. The complexity of counting in sparse, regular, and planar graphs. *SIAM J. Comput.*, 31(2):398–427, 2001.

[125] Leslie G. Valiant. The complexity of computing the permanent. *Theor. Comput. Sci.*, 8(2):189–201, 1979.

[126] Leslie G. Valiant. The complexity of enumeration and reliability problems. *SIAM J. Comput.*, 8(3):410–421, 1979.

[127] Leslie G. Valiant. Expressiveness of matchgates. *Theor. Comput. Sci.*, 289(1):457–471, 2002.

[128] Leslie G. Valiant. Quantum circuits that can be simulated classically in polynomial time. *SIAM J. Comput.*, 31(4):1229–1254, 2002.

[129] Leslie G. Valiant. Corrigendum to "Expressiveness of matchgates": [Theor. Comput. Sci. 289(1) (2002) 457471]. *Theor. Comput. Sci.*, 299(1-3):795, 2003.

[130] Leslie G. Valiant. Holographic circuits. In *ICALP*, pages 1–15. Springer Berlin Heidelberg, 2005.

[131] Leslie G. Valiant. Accidental algorthims. In *FOCS*, pages 509–517. IEEE Computer Society, 2006.

[132] Leslie G. Valiant. Holographic algorithms. *SIAM J. Comput.*, 37(5):1565–1594, 2008.

[133] Leslie G. Valiant. Some observations on holographic algorithms. In *LATIN*, pages 577–590. Springer Berlin Heidelberg, 2010.

[134] Vijay V. Vazirani. NC algorithms for computing the number of perfect matchings in $k_{3,3}$-free graphs and related problems. *Inform. and Comput.*, 80(2):152–164, 1989.

[135] Michel Las Vergnas. Le polynôme de Martin d'un graphe Eulerien. *Ann. Discrete Math.*, 17:397–411, 1983.

[136] Michel Las Vergnas. On the evaluation at (3, 3) of the Tutte polynomial of a graph. *J. Comb. Theory, Ser. B*, 45(3):367–372, 1988.

[137] D. L. Vertigan and D. J. A. Welsh. The computational complexity of the Tutte plane: the bipartite case. *Combin. Probab. Comput.*, 1(2):181–187, 1992.

[138] Dirk Vertigan. The computational complexity of Tutte invariants for planar graphs. *SIAM J. Comput.*, 35(3):690–712, 2005.

[139] Vadim G. Vizing. Critical graphs with given chromatic class. *Metody Diskret. Analiz.*, 5:9–17, 1965.

[140] P. G. Walsh. A quantitative version of Runge's theorem on Diophantine equations. *Acta Arith.*, 62(2):157–172, 1992.

[141] Dominic Welsh. *Complexity: Knots, Colourings and Countings.* London Mathematical Society Lecture Note Series. Cambridge University Press, 1993.

[142] B. Widom and J. S. Rowlinson. New model for the study of liquidvapor phase transitions. *J. Chem. Phys*, 52(4):1670–1684, 1970.

[143] Mingji Xia. Holographic reduction: A domain changed application and its partial converse theorems. *Int. J. Software and Informatics*, 5(4):567–577, 2011.

[144] Mingji Xia, Peng Zhang, and Wenbo Zhao. Computational complexity of counting problems on 3-regular planar graphs. *Theor. Comput. Sci.*, 384(1):111–125, 2007.

[145] C. N. Yang. The spontaneous magnetization of a two-dimensional Ising model. *Phys. Rev.*, 85(5):808–816, 1952.

[146] C. N. Yang and T. D. Lee. Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Phys. Rev.*, 87(3):404–409, 1952.