On Holant Problems

Jin-Yi Cai [1]
Computer Sciences Department
University of Wisconsin
Madison, WI 53706. USA
jyc@cs.wisc.edu

Pinyan Lu [2]
Institute for Theoretical Computer Science
Tsinghua University
Beijing, 100084, P. R. China
lpy@mails.tsinghua.edu.cn

Mingji Xia [3]
Computer Sciences Department
University of Wisconsin
Madison, WI 53706. USA
and State Key Laboratory of Computer Science,
Institute of Software, Chinese Academy of Sciences
Beijing 100190, P. R. China
xmjljx@gmail.com

[1]Supported by NSF CCR-0511679.

[2]Supported by the National Natural Science Foundation of China Grant 60553001 and the National Basic Research Program of China Grant 2007CB807900, 2007CB807901.

[3]Supported by Hundred Talent Program of Chinese Academy of Sciences Under Angsheng Li.

Abstract

We propose and explore a novel alternative framework to study the complexity of counting problems, called Holant Problems. Compared to counting Constrained Satisfaction Problems (#CSP), it is a refinement with a more explicit role for the function constraints. Both graph homomorphism and #CSP can be viewed as special cases of Holant Problems. We prove complexity dichotomy theorems in this framework. Because the framework is more stringent, previous dichotomy theorems for #CSP problems no longer apply. Indeed, we discover surprising tractable subclasses of counting problems, which could not have been easily specified in the #CSP framework. The main technical tool we use and develop is holographic reductions. Another technical tool used in combination with holographic reductions is polynomial interpolations.

1 Introduction

In order to study the complexity of counting problems, several interesting frameworks have been proposed. One is called counting Constrained Satisfaction Problems (#CSP) [4, 2, 15, 3]. Another well studied framework is called Graph Homomorphisms or H-coloring problems, which can be viewed as a special case of #CSP problems [5, 6, 18, 19, 16, 17, 21, 7]. One reason such frameworks are interesting is because the language is *expressive* enough so that they can express many natural counting problems, while *specific* enough so that we can prove complete complexity classifications for them [12]. The natural counting problems which can be expressed as graph homomorphism problems include counting the number of vertex covers, the number of k-colorings in a graph, and many others. However, there are some natural and important counting problems, which can not be expressed as a graph homomorphism problem. In [20], it is proved that counting the number of perfect matchings in a graph *cannot* be expressed as a graph homomorphism function. Additionally, sometimes a problem can be expressed in the existing framework, such as #CSP, but only with some contrived restrictions.

In this paper, we propose and explore an alternative framework to study the complexity of counting problems, called Holant Problems. This notion is motivated by holographic reductions proposed by Valiant [29, 30]. Compared to #CSP, it is a refinement with a more explicit role for the function constraints. Both graph homomorphism and #CSP can be viewed as special cases of Holant Problems. We give a brief description here and a more formal definition is given in Section 2. A *signature grid* $\Omega = (G, \mathcal{F}, \pi)$ is a tuple, where $G = (V, E)$ is a graph, and π maps each $v \in V(G)$ to a function $f_v \in \mathcal{F}$. We consider all edge assignments. An assignment σ for every $e \in E$ gives an evaluation $\prod_{v \in V} f_v(\sigma \mid_{E(v)})$, where $E(v)$ denotes the incident edges of v. The counting problem on the instance Ω is to compute

$$\text{Holant}_\Omega = \sum_\sigma \prod_{v \in V} f_v(\sigma \mid_{E(v)}).$$

For example, consider the Perfect Matching problem on G. This problem corresponds to attaching the Exact-One function at every vertex of G, and then consider all 0-1 edge assignments. In this case, Holant_Ω counts the number of perfect matchings. If we use the At-Most-One function at every vertex, then we are counting all (not necessarily perfect) matchings. So this new framework can express some natural counting problems which are not expressible as graph homomorphisms.

To see that Holant is a more expressive framework, we show that every #CSP problem can be simulated by a Holant problem. Represent an instance of a #CSP problem by a bipartite graph where LHS are labeled by variables and RHS are labeled by constraints. Now the signature grid Ω on this bipartite graph is as follows: Every variable node on LHS is attached an Equality function, every constraint node on RHS has the given constraint function. Then Holant_Ω is exactly the answer to the counting #CSP problem. In effect, the Equality function on each variable node forces the incident edges to take the same value; this effectively reduces to assigning values to each variable on LHS as in #CSP. It follows that #CSP problems are precisely the special case of Holant problems on bipartite graphs where every node on LHS is attached an Equality function. We can show that the class of #CSP problems is equivalent to Holant problems where all Equality functions are always assumed to be freely available, and implicitly so. Graph homomorphism is a further special case where not only all Equality functions are freely (and implicitly) available, but the function set $\mathcal{F}$ in our signature grid Ω contains exactly one binary function (other than these Equality functions). It turns out that allowing Equality functions has a major influence on the computational complexity of the problems. By making the presence of these Equality functions explicit, the Holant framework of counting problems can make a finer complexity classification, which is difficult to do in #CSP.

Our Holant Problem framework is strongly influenced by the development of holographic algorithms and holographic reductions [29, 30, 8, 11]. Indeed, we will use and develop holographic reductions here as a primary technique. One advantage of our new framework is that one can naturally consider new subclasses of counting problems as special cases of Holant problems other than #CSP problems. For instance, by assuming all unary functions are freely available, we propose an interesting counting problem family called Holant* Problems. Our first main result is a complexity dichotomy theorem for all Holant* Problems for arbitrary complex valued symmetric functions over Boolean variables: Each problem in the class is either #P-hard or solvable in P. In this dichotomy theorem, all tractable cases are accomplished by holographic algorithms with Fibonacci gates [11]. And what is more interesting and surprising is that the key technique used in the hardness proof is also holographic reductions. Furthermore, we prove that the theorem holds for planar graphs.

Our second main result is a dichotomy theorem for an even more appealing family of counting problems, called Holant^c Problems, where we only assume two special unary functions Δ_0 and Δ_1 are available. These two unary functions simply set a particular edge (variable) to a constant value 0 and 1. We can prove again that every problem in the class is either #P-hard or solvable in P. However here we can only prove it for all real valued symmetric functions over Boolean variables. (We conjecture that it is still true over $\mathbb{C}$.) Note that when we assume fewer functions are freely available in the framework it makes the specification of the family more stringent. It delineates more precisely what functions and in what combinations lead to #P-hardness, or to tractability, respectively. However, the fewer functions are assumed free, the more challenging it is to prove #P-hardness. We make *essential* use of the dichotomy theorem just proved for Holant* Problems, as a launching station to prove our dichotomy theorem for Holant^c Problems.

It is precisely the presence of Equality functions as freely available in #CSP that makes #CSP a less exacting framework than Holant Problems. The Holant^c Problems are basically generic Holant Problems with the ability to fix the assignments of some edges. In many natural counting problems, this is indeed the case, such as counting problems for perfect matchings. By the Pinning Lemma in [15], in any #CSP problem, Δ_0, Δ_1 can be simulated, and as a result can be viewed as freely available. In other words Equality functions are stronger than Δ_0 and Δ_1. Therefore Holant^c Problems already subsume #CSP, and in the meanwhile provide a way for a more exacting account of what makes a problem tractable or #P-hard.

The main technique for the proof of the second dichotomy theorem is polynomial interpolations. Once we can interpolate all the unary functions, we can apply the result for Holant* Problems. Our dichotomy theorems have already paid dividend in the study of classifications of #CSP problems. Since #CSP can be viewed as a special case of Holant^c Problems, the dichotomy theorem for Holant^c Problems automatically implies a dichotomy theorem for Boolean #CSP problems with real symmetric constraints. Motivated by this, we investigated how one might generalize the tractable cases to unsymmetric ones. Surprisingly it turns out that the symmetric tractable cases already supplied the essential ingredients for all possible (including unsymmetric) tractable ones. This led us to a dichotomy theorem for the whole family of complex weighted Boolean #CSP. However the proof there requires substantial new techniques and we report it in a separate paper [10].

2 Definitions and Background

Our functions take values in $\mathbb{C}$ by default. We will mostly be concerned with symmetric functions on Boolean variables, however the framework of Holant Problems is defined for functions mapping any $[q]^k \rightarrow \mathbb{C}$ for a finite q. For brevity we will mostly restrict $q = 2$.

As stated, a *signature grid* $\Omega = (H, \mathcal{F}, \pi)$ consists of a graph $H = (V, E)$ with each vertex labeled by a function $f_v \in \mathcal{F}$. We use $\mathcal{F}_q$ when variables range over $[q]$. The Holant problem on instance Ω is to compute $\text{Holant}_\Omega = \sum_\sigma \prod_{v \in V} f_v(\sigma\ |_{E(v)})$, a sum over all edge assignments. A function f_v can be represented as a truth table, or a tensor in $(\mathbb{C}^2)^{\otimes \deg(v)}$. This is called a *signature*. We denote by $=_k$ the EQUALITY signature of arity k. Δ_0 (respectively Δ_1) denotes the unary signature which takes value 1 on input 0 (respectively 1), and 0 otherwise. A symmetric function f on k Boolean variables can be expressed by $[f_0, f_1, \ldots, f_k]$, where f_j is the value of f on inputs of Hamming weight j. Thus, $(=_k) = [1, 0, \ldots, 0, 1]$, $\Delta_0 = [1, 0]$ and $\Delta_1 = [0, 1]$. A Holant problem is parameterized by a set of signatures.

Definition 2.1. *Given a set of signatures $\mathcal{F}$, we define a counting problem* $\text{Holant}(\mathcal{F})$*:*
Input: *A signature grid $\Omega = (G, \mathcal{F}, \pi)$;*
Output: Holant_Ω*.*

We would like to characterize the complexity of Holant problems in terms of its signature sets. Some special families of Holant problems have already been widely studied. For example, if $\mathcal{F}_q$ contains all EQUALITY signatures $\{=_1, =_2, =_3, \cdots\}$, then this is exactly the weighted #CSP problem. #CSP problems are a special family of Holant problems, where we assume that all EQUALITY functions are freely available. Graph homomorphism is a further special case, where we only allow a single binary function in $\mathcal{F}_q$ other than all the EQUALITY functions.

We now define two more special families of Holant problems by assuming some signatures are freely available. We define them for $q = 2$; they can be easily extended to arbitrary $[q]$.

Definition 2.2. *let $\mathcal{U}$ denote the set of all unary signatures. Given a set of signatures $\mathcal{F}$, we use* $\text{Holant}^*(\mathcal{F})$ *to denote* $\text{Holant}(\mathcal{F} \cup \mathcal{U})$*.*

Definition 2.3. *Given a set of signatures $\mathcal{F}$, we use* $\text{Holant}^c(\mathcal{F})$ *to denote* $\text{Holant}(\mathcal{F} \cup \{\Delta_0, \Delta_1\})$*.*

Replacing a signature $f \in \mathcal{F}$ by a constant multiple cf, where $c \neq 0$, does not change the complexity of $\text{Holant}(\mathcal{F})$. So we view f and cf as the same signature. An important property of a signature is whether it is degenerate.

Definition 2.4. *A signature is degenerate iff it is a tensor product of unary signatures.*

In particular, a symmetric signature in $\mathcal{F}$ is degenerate iff it can be expressed as $\lambda[x, y]^{\otimes k}$. Also a symmetric signature $[x_0, x_1, \ldots, x_n]$ is non-degenerate iff $\text{rank}\begin{bmatrix} x_0 & \cdots & x_{n-1} \\ x_1 & \cdots & x_n \end{bmatrix} = 2$.

To introduce the idea of holographic reductions, it is convenient to consider bipartite graphs. This is without loss of generality. For any general graph, we can make it bipartite by adding an additional vertex on each edge, and giving each new vertex the EQUALITY function $=_2$ on 2 inputs.

We use $\#\mathcal{G}_q|\mathcal{R}_q$ to denote all counting problems, expressed as Holant problems on bipartite graphs $H = (U, V, E)$, where each signature for a vertex in U or V is from $\mathcal{G}_q$ or $\mathcal{R}_q$, respectively. An input instance for the bipartite Holant problem is a bipartite signature grid and is denoted as $\Omega = (H, \mathcal{G}_q|\mathcal{R}_q, \pi)$. Signatures in $\mathcal{G}_q$ are denoted by column vectors (or contravariant tensors); signatures in $\mathcal{R}_q$ are denoted by row vectors (or covariant tensors) [14].

One can perform (contravariant and covariant) tensor transformations on the signatures, which may produce exponential cancelations in tensor spaces. We will define a simple version of holographic reductions, which are invertible. Suppose $\#\mathcal{G}_q|\mathcal{R}_q$ and $\#\mathcal{G}'_q|\mathcal{R}'_q$ are two Holant problems defined for the same family of graphs, and $T \in \mathbf{GL}_q(\mathbb{C})$ is a basis. We say that there is a holographic reduction from $\#\mathcal{G}_q|\mathcal{R}_q$ to $\#\mathcal{G}'_q|\mathcal{R}'_q$, if the *contravariant* transformation $G' = T^{\otimes g}G$ and the *covariant* transformation $R = R'T^{\otimes r}$ map $G \in \mathcal{G}_q$ to $G' \in \mathcal{G}'_q$ and $R \in \mathcal{R}_q$ to $R' \in \mathcal{R}'_q$, where G and R have arity g and r respectively. (Notice the reversal of directions when the transformation $T^{\otimes n}$ is applied. This is the meaning of *contravariance* and *covariance*.)

Theorem 2.1 (Valiant's Holant Theorem)**.** *Suppose there is a holographic reduction from* $\#\mathcal{G}_q|\mathcal{R}_q$ *to* $\#\mathcal{G}'_q|\mathcal{R}'_q$ *mapping signature grid* Ω *to* Ω'*, then* $\text{Holant}_\Omega = \text{Holant}_{\Omega'}$.

In particular, for invertible holographic reductions from $\#\mathcal{G}_q|\mathcal{R}_q$ to $\#\mathcal{G}'_q|\mathcal{R}'_q$, one problem is in P iff the other one is, and similarly one problem is #P-hard iff the other one is also.

Theorem 2.2. *Let* $\mathcal{F}_q$ *be a set of signatures and* M *be a* $q \times q$ *orthogonal matrix, i.e.,* $MM^{\mathrm{T}} = I_q$*. For any signature grid* $\Omega = (G, \mathcal{F}_q, \pi)$*, replacing every signature* $F \in \mathcal{F}_q$ *by* $M^{\otimes n}F$*, where* n *is the arity of* F*, we can get a new signature grid* Ω'*. Then* $\text{Holant}_\Omega = \text{Holant}_{\Omega'}$.

Proof. First we use a standard technique to reformulate the signature grid $\Omega = (G, \mathcal{F}_q, \pi)$. We insert a new vertex at each edge of G with signature $=_2$. This will not change the value of the signature grid. Then for the new bipartite signature grid $(G', \mathcal{F}_q|\{=_2\}, \pi)$, we apply a holographic reduction with basis M. This will map a signature $F \in \mathcal{F}_q$ to $M^{\otimes n}F$, where n is the arity of F. It is an algebraic fact that $=_2$ will map to itself. Now we can replace each new $=_2$ node back to an edge to revert back to G. This gives the signature grid Ω' as required. By the Holant theorem, its value is the same as Ω. □

This theorem is very useful as a way to normalize a given signature set $\mathcal{F}_q$.

Starting from next section, we will exclusively focus on Boolean variables. A technical issue is the model of computation for $\mathbb{C}$. Strictly speaking we must only use computable numbers [1, 24]. We will state our results for all $\mathbb{C}$, assuming all numbers in a particular instance (signature) are computable.

3 A Dichotomy Theorem for Holant$^*(\mathcal{F})$

Theorem 3.1. *Let* $\mathcal{F}$ *be a set of symmetric signatures over* $\mathbb{C}$*. Then* Holant$^*(\mathcal{F})$ *is computable in polynomial time in the following three* Classes*. In all other cases,* Holant$^*(\mathcal{F})$ *is #P-hard.*

1. *Every signature in* $\mathcal{F}$ *is of arity no more than two;*

2. *There exist two constants* a *and* b *(not both zero, depending only on* $\mathcal{F}$*), such that for all signatures* $[x_0, x_1, \ldots, x_n] \in \mathcal{F}$ *one of the two conditions is satisfied: (1) for every* $k = 0, 1, \ldots, n-2$*, we have* $ax_k + bx_{k+1} - ax_{k+2} = 0$*; (2)* $n = 2$ *and the signature* $[x_0, x_1, x_2]$ *is of form* $[2a\lambda, b\lambda, -2a\lambda]$.

3. *For every signature* $[x_0, x_1, \ldots, x_n] \in \mathcal{F}$ *one of the two conditions is satisfied: (1) For every* $k = 0, 1, \ldots, n-2$*, we have* $x_k + x_{k+2} = 0$*; (2)* $n = 2$ *and the signature* $[x_0, x_1, x_2]$ *is of form* $[\lambda, 0, \lambda]$.

The dichotomy is still true even if the inputs are restricted to planar graphs.

Remark: Since all unary signatures can be used for free, we always assume the arity of every signature in $\mathcal{F}$ is larger than one. And since all the degenerate signatures can be decomposed to unary signatures, we also assume that every signature in $\mathcal{F}$ is non-degenerate.

Proof Outline: The first Class to be computable in P is easy. One can compute the signature of a path by matrix multiplication. The other two polynomial time computable Classes follow from our previous work on Fibonacci gates [11].

Now for the hardness, we first prove in Lemma 3.1 that the theorem holds if $\mathcal{F}$ contains a single symmetric signature of arity three. The main technique is holographic reductions. We make use of the signature theory developed in holographic algorithms [9, 8]. This theory gives us three Categories in a certain parametrization for the signature according to the eigenvalues. For each Category, we choose one #P-hard problem to reduce from, all using holographic reductions. In Lemma 3.2, we prove that if one signature has the form in Class 2 of Theorem 3.1, and we combine it with another signature which is not in this Class, then the Holant* problem is #P-hard. The main idea of the proof is to reduce it to Lemma 3.1 with holographic reductions. In Lemma 3.3, we prove the same thing is true for Class 3. Finally we extend the above proofs to a set of signatures of arbitrary arities and finish the whole proof.

The following lemma is the first important step towards the proof of Theorem 3.1. It says that Theorem 3.1 holds if $\mathcal{F}$ only contains one signature of arity three. Holographic reductions play a decisive role in the proof. This Lemma serves as the foundation for all subsequent lemmas.

Lemma 3.1. *Let* $[x_0, x_1, x_2, x_3]$ *be a symmetric signature with arity 3, then* Holant*$([x_0, x_1, x_2, x_3])$ *is #P-hard unless one of the following two statements is true: (1) there exist two constants* a, b *(not both zero) such that* $ax_0 + bx_1 - ax_2 = 0$ *and* $ax_1 + bx_2 - ax_3 = 0$*; (2)* $x_0 + x_2 = 0$ *and* $x_1 + x_3 = 0$*.*

Proof: Assume $[x_0, x_1, x_2, x_3]$ does not satisfy either of the two statements, we prove that Holant*$([x_0, x_1, x_2, x_3])$ is #P-hard. Our starting point is that $\#[0,1,1]|[1,0,0,1]$ and $\#[1,0,1]|[1,1,0,0]$ are both #P-Complete [31]. The first problem is simply counting the number of vertex covers for 3-regular graphs; while the second is to count the number of (not necessarily perfect) matchings for 3-regular graphs. We remark that both of them remain #P-Complete even for planar graphs.

First we use the signature theory from holographic algorithms to give a better parametrization. Given a non-degenerate signature $[x_0, x_1, x_2, x_3]$, there are three Categories:

- Category 1. $x_i = \alpha_1^{3-i}\alpha_2^i + \beta_1^{3-i}\beta_2^i$, where $\alpha_1\alpha_2 \neq 0, \beta_1\beta_2 \neq 0$ and $\alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$;
- Category 2. $x_i = Ai\alpha^{i-1} + B\alpha^i$, where $A \neq 0$; or
- Category 3. $x_i = A(3-i)\alpha^{2-i} + B\alpha^{3-i}$, where $A \neq 0$.

Category 3 can be viewed as the reversal of Category 2, so we will omit the proof for Category 3. The choices made here in this particular parametrization is informed by the "signature theory" [9, 8] that we have developed in previous work. (But one can directly check that for any non-degenerate signature $[x_0, x_1, x_2, x_3]$, one of these three parameterizations is always possible. Note that, if $\alpha = 0$ then we take the convention that the expression $i\alpha^{i-1} = 0, 1, 0, 0$ for $i = 0, 1, 2, 3$ respectively.)

For Category 1, we have

$$X = [x_0, x_1, x_2, x_3] = \begin{bmatrix}\alpha_1\\ \alpha_2\end{bmatrix}^{\otimes 3} + \begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}^{\otimes 3}.$$

We restate our conditions from the Lemma statement in this new parametrization. The fact that X is non-degenerate implies that $\alpha_1\beta_2 - \alpha_2\beta_1 \neq 0$. The fact that X is not in the case indicated in statement (1) implies that $\alpha_1\beta_1 + \alpha_2\beta_2 \neq 0$. This follows from Prop 5.4. The fact that X is not in the case indicated in statement (2) implies that $\alpha_1^2 + \alpha_2^2 \neq 0$ or $\beta_1^2 + \beta_2^2 \neq 0$. This follows from Prop 5.5. By symmetry, we can assume that $\alpha_1^2 + \alpha_2^2 \neq 0$.

Under the condition $\alpha_1^2 + \alpha_2^2 \neq 0$, we can apply an orthogonal transformation to map the vector (α_1, α_2) to be of the form $(\alpha_1', 0)$, where $\alpha_1' \neq 0$. We may use a (complex orthogonal) Householder matrix for this purpose. Then under this orthogonal basis, the signature becomes

$$X' = [x_0', x_1', x_2', x_3'] = \begin{bmatrix}\alpha_1'\\ 0\end{bmatrix}^{\otimes 3} + \begin{bmatrix}\beta_1'\\ \beta_2'\end{bmatrix}^{\otimes 3}.$$

By Proposition 5.3, this transformation does not change the complexity of the Holant problem. So it suffices to prove the #P-hardness result for this signature. By a scalar multiplication we assume $\alpha_1' = 1$. So, reuse the notation X, we can assume the signature is of this form

$$X = [x_0, x_1, x_2, x_3] = \begin{bmatrix}1\\ 0\end{bmatrix}^{\otimes 3} + \begin{bmatrix}\beta_1\\ \beta_2\end{bmatrix}^{\otimes 3}.$$

The two conditions from the statement of the Lemma become simply $\beta_1\beta_2 \neq 0$.

Now under the basis $T = \begin{bmatrix} 1 & \beta_1 \\ 0 & \beta_2 \end{bmatrix}$, signature $[1,0,0,1]$ becomes $[x_0, x_1, x_2, x_3]$. This is the result of the contravariant transformation (on truth tables) $(x_0, x_1, x_1, x_2, x_1, x_2, x_2, x_3)^{\mathrm{T}} = T^{\otimes 3}(1,0,0,0,0,0,0,1)^{\mathrm{T}}$, namely $X = T^{\otimes 3}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes 3} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}^{\otimes 3}\right)$. Under the same basis, $[0,1,1]$ undergoes a covariant transformation, we have

$$(0,1,1,1)(T^{-1})^{\otimes 2} = \frac{1}{\beta_2^2}(0, \beta_2, \beta_2, 1-2\beta_1).$$

Again, we can ignore the scalar factor $1/\beta_2^2$. So by the holographic reduction defined by T, the complexity of the problem $\#[0,\beta_2,1-2\beta_1]|[x_0,x_1,x_2,x_3]$ is the same as $\#[0,1,1]|[1,0,0,1]$, which is #P-Complete (vertex cover). In order to prove that $\text{Holant}^*([x_0,x_1,x_2,x_3])$ is #P-Complete, we only need to show that the signature $[0,\beta_2,1-2\beta_1]$ can be realized by $[x_0,x_1,x_2,x_3]$ with some unary signatures.

For a binary signature F we can write it in a matrix form $\begin{bmatrix} F(00) & F(01) \\ F(10) & F(11) \end{bmatrix}$. We use the gadget in

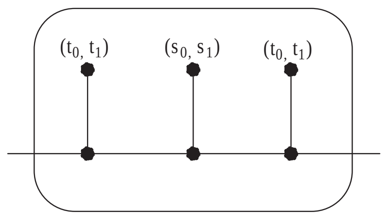

Figure 1: We use this gadget to realize the signature $[0, \beta_2, 1-2\beta_1]$. All (three) nodes of degree 3 in this gadget have the signature $[x_0, x_1, x_2, x_3]$.

Figure 1 to realize $[0,\beta_2,1-2\beta_1]$, where the two unary signatures (t_0,t_1) and (s_0,s_1) will be determined later. Let

$$A = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}\begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix} = \begin{bmatrix} \beta_1^2 & \beta_1\beta_2 \\ \beta_1\beta_2 & \beta_2^2 \end{bmatrix}.$$

In X, if one input is 0, the induced binary signature has matrix form $A + \beta_1 B$. If one input is 1, the induced binary signature has matrix form $\beta_2 B$. It follows that the signature of the above gadget is

$$\begin{aligned} & (t_0(A+\beta_1 B) + t_1\beta_2 B)(s_0(A+\beta_1 B) + s_1\beta_2 B)(t_0(A+\beta_1 B) + t_1\beta_2 B) \\ = \ & (t_0 A + (t_0\beta_1 + t_1\beta_2)B)(s_0 A + (s_0\beta_1 + s_1\beta_2)B)(t_0 A + (t_0\beta_1 + t_1\beta_2)B). \end{aligned}$$

Now we use a new set of variables $x = t_0, y = t_0\beta_1 + t_1\beta_2, z = s_0, w = s_0\beta_1 + s_1\beta_2$, and write the above matrix as $(xA+yB)(zA+wB)(xA+yB)$. We note that for any given x, y, z, w, we can find t_0, t_1, s_0, s_1 to satisfy the above relationships. Then, to realize $[0,\beta_2,1-2\beta_1]$, we just want to choose some x, y, z, w such that

$$(xA+yB)(zA+wB)(xA+yB) = \begin{bmatrix} 0 & \beta_2 \\ \beta_2 & 1-2\beta_1 \end{bmatrix}.$$

We show that we can find some x, y, z, w to satisfy the above condition.

Substituting A and B, and denote by $\gamma = \beta_1^2 + \beta_2^2$, we have the following:

$$\begin{aligned} & (xA+yB)(zA+wB)(xA+yB) \\ = \ & w\begin{bmatrix} \beta_1^2(x+y\gamma)^2 & y\beta_1\beta_2\gamma(x+y\gamma) \\ y\beta_1\beta_2\gamma(x+y\gamma) & y^2\beta_2^2\gamma^2 \end{bmatrix} + z\begin{bmatrix} (x+y\beta_1^2)^2 & y\beta_1\beta_2(x+y\beta_1^2) \\ y\beta_1\beta_2(x+y\beta_1^2) & y^2\beta_1^2\beta_2^2 \end{bmatrix} \end{aligned}$$

We may choose $w = (x + y\beta_1^2)^2$ and $z = -\beta_1^2(x + y\gamma)^2$ to make the $(1, 1)$ entry zero. The $(1, 2)$ (and $(2, 1)$) entry is

$$g_1 = xy\beta_1\beta_2^3(x + \beta_1^2 y)(x + y\gamma);$$

and the $(2, 2)$ entry is

$$g_2 = xy^2\beta_2^4(x(2\beta_1^2 + \beta_2^2) + 2y(\beta_1^4 + \beta_1^2\beta_2^2)).$$

We want to choose some x, y such that $[g_1, g_2] = [\beta_2, 1 - 2\beta_1]$. We have $\beta_2 \neq 0$. We will choose $xy \neq 0$. As both g_1 and g_2 are homogenous in x and y, we can ignore the common factor $xy\beta_2^3$ of g_1 and g_2. It follows that we only have to satisfy that $g_2/g_1 = (1 - 2\beta_1)/\beta_2$ with $y = 1$. We need the following

$$0 = \beta_2 g_2 - (1 - 2\beta_1)g_1 = \beta_1(2\beta_1 - 1)x^2 + (2\beta_1^2 - \beta_1 + \beta_2^2)(2\beta_1^2 + \beta_2^2)x + \beta_1^2(\beta_1^2 + \beta_2^2)(2\beta_1^2 - \beta_1 + 2\beta_2^2). \quad (1)$$

What we have to prove is that at least one of the roots to the equation in (1) is not a root of $g_1 = g_1(x, 1) = 0$. The roots of $g_1 = 0$ are $x = 0, x = -\beta_1^2$ and $x = -\gamma$. Firstly we can verify that $x = -\beta_1^2$ can not be a root of (1). This is because when $x = -\beta_1^2$, the expression in (1) can be simplified to $\beta_1^2\beta_2^4 \neq 0$. Secondly if $x = -\gamma$ is a root of (1), the expression in (1) can be simplified to $-\beta_2^4\gamma$, and this would force $\gamma = 0$. So, assuming the expression in (1) as a polynomial in x is indeed of degree 2, then the only case we need to worry is that $x = 0$ is a double root of (1). In fact, suppose (1) is indeed quadratic, and $x = 0$ is not a double root, then we may let $\xi \neq 0$ be a root of (1). This $\xi \neq -\beta_1^2$, because $-\beta_1^2$ is not a root of (1); ξ can't be $-\gamma$ either, for otherwise $-\gamma$ would be a root of (1) which we had proved it would force $\gamma = 0$, and thus $\xi = -\gamma = 0$, a contradiction. Thus ξ is a root of (1) but not a root of g_1, as is needed.

Now let's consider the exceptional cases: either $x = 0$ is a double root of (1), or (1) has degree less than 2. If $x = 0$ is a double root of (1), we have

$$(2\beta_1^2 - \beta_1 + \beta_2^2)(2\beta_1^2 + \beta_2^2) = \beta_1^2(\beta_1^2 + \beta_2^2)(2\beta_1^2 - \beta_1 + 2\beta_2^2) = 0.$$

To satisfy this, there are only four exceptional cases (A1 to A4): $\beta_1 = 1, \beta_2 = \pm i$ or $\beta_1 = -\frac{1}{2}, \beta_2 = \pm\frac{i}{\sqrt{2}}$. On the other hand, if the polynomial in (1) has degree less than 2, then $\beta_1 = \frac{1}{2}$. In this case, the polynomial becomes

$$(1/2 + \beta_2^2)x + (1/4 + \beta_2^2)/2 = 0.$$

This gives us four additional exceptional cases (B1 to B4): $\beta_1 = \frac{1}{2}, \beta_2 = \pm\frac{i}{2}$, in which case the polynomial is linear with root $x = 0$; or $\beta_1 = \frac{1}{2}, \beta_2 = \pm\frac{i}{\sqrt{2}}$, in which case the polynomial degenerates to a (non-zero) constant. In all other cases, there is a root of (1) which is not a root of g_1, which completes the #P-completeness proof.

For the cases A1 and A2, we use a new starting problem #[1, 1, 0]|[1, 0, 0, 1], which is the reversal of the previous problem and therefore it is also #P-Complete. Then all previous part of the proof is still valid, except that the signature of arity two we would like to realize is

$$(1, 1, 1, 0)(T^{-1})^{\otimes 2} = (1, \frac{1 - \beta_1}{\beta_2}, \frac{1 - \beta_1}{\beta_2}, \frac{\beta_1^2 - 2\beta_1}{\beta_2^2}).$$

Substituting $\beta_1 = 1, \beta_2 = \pm i$, the signature is $[1, 0, 1]$ which is trivially realizable by one edge. So we have proved that it is #P-Complete in the cases A1 and A2. Now consider the case A3 and A4, $\beta_1 = -\frac{1}{2}, \beta_2 = \pm\frac{i}{\sqrt{2}}$. We will give a different parametrization. For case A3, we apply an orthogonal transformation $M = \begin{bmatrix} -i & -\sqrt{2} \\ \sqrt{2} & -i \end{bmatrix}$ and a scalar multiplier $2i$ on the signature and it becomes $\begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes 3} + \begin{bmatrix} 2 \\ 2\sqrt{2}i \end{bmatrix}^{\otimes 3}$. This is not one of the exceptional cases and we have proved that it is #P-Complete. For case A4, we apply another orthogonal transformation $M' = \begin{bmatrix} i & -\sqrt{2} \\ \sqrt{2} & i \end{bmatrix}$ and a scalar multiplier $-2i$ on the signature and it becomes $\begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes 3} + \begin{bmatrix} 2 \\ -2\sqrt{2}i \end{bmatrix}^{\otimes 3}$.

The cases B3 and B4 can be shown by the same method as in A4 and A3, using M' and M respectively. The only cases left are B1 and B2. Here we will use another gadget similar to the one in Figure 1 except we remove the middle edge (including the node labeled (s_0, s_1) and the middle node of degree 3). For B1, the signature of this gadget is

$$(t_0(A + \beta_1 B) + t_1\beta_2 B)^2 = (xA + yB)^2,$$

where A and B are as before, and with the specific values of β_1, β_2, $B = \frac{1}{4}\begin{bmatrix} 1 & i \\ i & -1 \end{bmatrix}$. By setting $x = i$ and $y = -2i$, we have $(xA + yB)^2 = \begin{bmatrix} 0 & i/2 \\ i/2 & 0 \end{bmatrix}$, which is the matrix form of the target signature $[0, \beta_2, 1 - 2\beta_1] = [0, \frac{i}{2}, 0]$. This finishes case B1. The case B2 can be done with $x = 1$ and $y = -2$.

Now we prove for Category 2. In this case $x_i = Ai\alpha^{i-1} + B\alpha^i$, the condition that it does not satisfy statement (2) in Lemma 3.1 implies that $\alpha \neq \pm i$. This is because $\operatorname{rank}\begin{bmatrix} x_0 - x_2 & x_1 \\ x_1 - x_3 & x_2 \end{bmatrix} = 2$ and its determinant can be shown to be $-A^2(1 + \alpha^2)$. Under this condition, we can choose some orthogonal transformation to make it in the form $[x, y, 0, 0]$ where $y \neq 0$. In fact, if we let $T = \begin{bmatrix} 1 & \frac{B-1}{3} \\ \alpha & A + \frac{B-1}{3}\alpha \end{bmatrix}$, then the signature $[x_0, x_1, x_2, x_3]$ can be expressed as

$$(x_0, x_1, x_1, x_2, x_1, x_2, x_2, x_3)^{\mathrm{T}} = T^{\otimes 3}(1, 1, 1, 0, 1, 0, 0, 0)^{\mathrm{T}}.$$

(We chose these basis transformations based on an underlying signature theory of holographic algorithms, not "out of blue". But for brevity of exposition we state these transformations *as is* without discussing the background. They can be directly verified, albeit a bit tedious.) Let $T = QR$ be its QR factorization, where Q is orthogonal and R is upper triangular. In fact if we denote $T = \begin{bmatrix} 1 & * \\ \alpha & * \end{bmatrix}$, then we can choose our Q as the (orthogonal) Householder matrix, which is a (complex) reflection, $Q = Q^{\mathrm{T}} = \frac{1}{\sqrt{1+\alpha^2}}\begin{bmatrix} 1 & \alpha \\ \alpha & -1 \end{bmatrix}$. Then $QT = R = \begin{bmatrix} u & w \\ 0 & v \end{bmatrix}$ is upper triangular, where $u = \sqrt{1 + \alpha^2}$. As $\det Q = -1$, $\det R = -\det T = -A \neq 0$, we have $uv \neq 0$. This Q is our choice of the orthogonal transformation. It follows that

$$\begin{aligned}
& Q^{\otimes 3}(x_0, x_1, x_1, x_2, x_1, x_2, x_2, x_3)^{\mathrm{T}} \\
= \; & (QT)^{\otimes 3}(1, 1, 1, 0, 1, 0, 0, 0)^{\mathrm{T}} \\
= \; & R^{\otimes 3}(1, 1, 1, 0, 1, 0, 0, 0)^{\mathrm{T}} \\
= \; & R^{\otimes 3}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes 3} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} \\
= \; & \begin{bmatrix} u \\ 0 \end{bmatrix}^{\otimes 3} + \begin{bmatrix} u \\ 0 \end{bmatrix} \otimes \begin{bmatrix} u \\ 0 \end{bmatrix} \otimes \begin{bmatrix} w \\ v \end{bmatrix} + \begin{bmatrix} u \\ 0 \end{bmatrix} \otimes \begin{bmatrix} w \\ v \end{bmatrix} \otimes \begin{bmatrix} u \\ 0 \end{bmatrix} + \begin{bmatrix} w \\ v \end{bmatrix} \otimes \begin{bmatrix} u \\ 0 \end{bmatrix} \otimes \begin{bmatrix} u \\ 0 \end{bmatrix}
\end{aligned}$$

This can be written as a symmetric signature form $[u^3 + 3u^2w, u^2v, 0, 0]$. Note that the entry $u^2v \neq 0$.

By a scalar multiplication, we can make the entry u^2v equal to 1. So we only have to deal with a signature of the form $[v, 1, 0, 0]$ for an arbitrary given v.

For this signature, we can apply a holographic transformation defined by the matrix $T' = \begin{bmatrix} 1 & \frac{v-1}{3} \\ 0 & 1 \end{bmatrix}$ with inverse $T'^{-1} = \begin{bmatrix} 1 & -\frac{v-1}{3} \\ 0 & 1 \end{bmatrix}$. To prove #P-hardness, we will reduce from the matching problem $\#[1, 0, 1] \mid [1, 1, 0, 0]$. Under a contravariant transformation $(v, 1, 1, 0, 1, 0, 0, 0)^{\mathrm{T}} = T'^{\otimes 3}(1, 1, 1, 0, 1, 0, 0, 0)^{\mathrm{T}}$, the signature $[1, 1, 0, 0]$ becomes $[v, 1, 0, 0]$. Under the same basis, $[1, 0, 1]$

undergoes the covariant transformation to become $(1,0,0,1)(T'^{-1})^{\otimes 2} = ((1,0)^{\otimes 2} + (0,1)^{\otimes 2})(T'^{-1})^{\otimes 2} = (1, \frac{1-v}{3}, \frac{1-v}{3}, 1+\frac{(1-v)^2}{9})$. I.e., the signature $[1,0,1]$ becomes a new symmetric signature $[1, \frac{1-v}{3}, 1+\frac{(1-v)^2}{9}]$. The proof is then to use the same gadget as in Figure 1 to realize this signature, using unary signatures and $[v,1,0,0]$.

We will rename the values $x = t_0$, $y = t_1$, $z = s_0$ and $w = s_1$ in Figure 1. The signature of this gadget in matrix form is $(xA + yB)(zA + wB)(xA + yB)$, where $A = \begin{bmatrix} v & 1 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. After some calculations we found that this signature in matrix form is

$$\begin{bmatrix} w \cdot (x^2v^2 + 2xyv + y^2) + z \cdot (x^2(v^3 + 2v) + 2xy(v^2 + 1) + y^2v) & w \cdot (x^2v + xy) + z \cdot (x^2(v^2 + 1) + xyv) \\ w \cdot (x^2v + xy) + z \cdot (x^2(v^2 + 1) + xyv) & w \cdot x^2 + z \cdot x^2v \end{bmatrix}.$$

Our goal is to choose x, y, z and w such that it is equal to $\begin{bmatrix} 1 & \frac{1-v}{3} \\ \frac{1-v}{3} & 1+\frac{(1-v)^2}{9} \end{bmatrix}$. We can write this as a system of three linear equations in z and w. Then we can complete the proof, if we can choose x and y such that the following matrix has determinant 0, yet the first two columns have rank 2.

$$\begin{bmatrix} x^2v^2 + 2xyv + y^2 & x^2(v^3 + 2v) + 2xy(v^2 + 1) + y^2v & 1 \\ x^2v + xy & x^2(v^2 + 1) + xyv & \frac{1-v}{3} \\ x^2 & x^2v & 1 + \frac{(1-v)^2}{9} \end{bmatrix}.$$

After some row operations it becomes $\begin{bmatrix} y^2 & 2xy + y^2v & f_3 \\ xy & x^2 + xyv & f_2 \\ x^2 & x^2v & f_1 \end{bmatrix}$, where f_1, f_2, f_3 are polynomials in v, and explicitly, $f_1 = (10 - 2v + v^2)/9$ and $f_2 = (3 - 13v + 2v^2 - v^3)/9$. Subtracting from the second column the first column multiplied by v, we get $\begin{bmatrix} y^2 & 2xy & f_3 \\ xy & x^2 & f_2 \\ x^2 & 0 & f_1 \end{bmatrix}$. We will set $x = 1$; this guarantees that the first two columns have rank 2, and gives the matrix $\begin{bmatrix} y^2 & 2y & f_3 \\ y & 1 & f_2 \\ 1 & 0 & f_1 \end{bmatrix}$. Now the determinant is easily calculated, (subtract the first row by the second row multiplied by y, and the second from the third multiplied by y). The determinant is $-(f_1y^2 - 2f_2y + f_3)$. As long as f_1 and f_2 are not simultaneously 0, we can always choose a y to make this determinant 0.

However it is easy to show that f_1 and f_2 have no common zero in v, as $3(f_2 + vf_1) = 1 - v$ and $v = 1$ is not a zero of either f_1 or f_2. This completes the proof.

□

Lemma 3.1 shows us what happens when there is a single non-degenerate symmetric signature of arity 3. It explicitly lists two exceptional cases for being *not* #P-hard. The next Lemma addresses what happens if one signature of arity 3 happens to be in the first exceptional case, but some other signature *does not quite fit*.

Lemma 3.2. *Let $[x_0, x_1, x_2, x_3]$ and $[y_0, y_1, y_2]$ be non-degenerate symmetric signatures with arity three and two respectively. Suppose there exist two constants a, b (not both zero), such that $ax_0 + bx_1 - ax_2 = 0$ and $ax_1 + bx_2 - ax_3 = 0$, but $ay_0 + by_1 - ay_2 \neq 0$ and $[y_0, y_1, y_2]$ is not of the form $[2a\lambda, b\lambda, -2a\lambda]$. Then* Holant*($\{[x_0, x_1, x_2, x_3], [y_0, y_1, y_2]\}$) *is #P-hard.*

Lemma 3.3 does the same thing as Lemma 3.2 for the other exceptional case of the arity 3 signature.

Lemma 3.3. *Let $[x, y, -x, -y]$ be a symmetric signature with arity three and $[y_0, y_1, y_2]$ be a symmetric signature with arity two. Suppose they are both non-degenerate. If $y_0 + y_2 \neq 0$ and $[y_0, y_1, y_2]$ is not of the form $[\lambda, 0, \lambda]$, then* Holant*($\{[x, y, -x, -y], [y_0, y_1, y_2]\}$) *is #P-hard.*

The proofs of the above two lemmas can be found in the appendix. The following lemma extends the result to a signature with arbitrary arity. The proof is also given in the appendix.

Lemma 3.4. *Let* $[x_0, x_1, x_2, \ldots, x_n]$ *be a non-degenerate symmetric signature with arity* $n > 3$*, then* $\text{Holant}^*([x_0, x_1, x_2, \ldots, x_n])$ *is #P-hard unless one of the following two statements is true: (1) there exist two constants* a, b *(not both zero), such that for all* $k = 0, 1, \ldots, n-2$*, we have* $ax_k + bx_{k+1} - ax_{k+2} = 0$ *(the pair* (a, b) *is unique up to a scalar factor); (2) for all* $k = 0, 1, \ldots, n-2$*, we have* $x_k + x_{k+2} = 0$*.*

Finally we further extend this result to a set of signatures and finish the proof for Theorem 3.1. The details can be found in the appendix.

4 A Dichotomy Theorem for $\text{Holant}^c(\mathcal{F})$

Theorem 4.1. *Let* $\mathcal{F}$ *be a set of real symmetric signatures, and let* $\mathcal{F}_1, \mathcal{F}_2$ *and* $\mathcal{F}_3$ *be three families of signatures defined as*

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

Then $\text{Holant}^c(\mathcal{F})$ *is computable in polynomial time if (1)* $\text{Holant}^*(\mathcal{F})$ *is computable in polynomial time or (2)* $\mathcal{F} \subseteq \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$*. Otherwise,* $\text{Holant}^c(\mathcal{F})$ *is #P-hard.*

We only give a proof outline here. The complete proof is given as an appendix in Section 11.

Proof Outline: By definition, every instance of $\text{Holant}^c(\mathcal{F})$ is also an instance of $\text{Holant}^*(\mathcal{F})$. So it is obvious that if $\text{Holant}^*(\mathcal{F})$ is computable in polynomial time then so is $\text{Holant}^c(\mathcal{F})$. The polynomial time algorithm for $\mathcal{F} \subseteq \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$ is non-trivial. The tractability crucially depends on algebraic cancelations. The crux of the matter is a polynomial time algorithm to evaluate

$$\sum_{x_1, x_2, \ldots, x_k = 0,1} i^{L_1(X) + L_2(X) + \cdots + L_n(X)},$$

where each $L_i(X)$ is a 0-1 indicator function of a mod 2 sum. Thus the exponent on i can be viewed as a mod 4 sum of mod 2 sums. It turns out that this tractability leads to a more general P-time algorithm which includes unsymmetric signatures as well. In [10], we extend these families to a unified algorithm which eventually leads a dichotomy theorem for complex valued Boolean #CSP.

The main result here is hardness. We want to prove that aside from these tractable cases, all remaining problems are #P-hard. Here the main technique is polynomial interpolation. We prove the second dichotomy theorem (Theorem 4.1) by a reduction to the first (Theorem 3.1). We will show how to interpolate all the unary signatures. The interpolation method used here is briefly described in Section 10. Once we can interpolate all unary signatures by Lemma 11.1, we can make use of the dichotomy theorem for $\text{Holant}^*(\mathcal{F})$. The whole proof is organized as a sequence of lemmas (Lemma 11.3 to Lemma 11.7). In each lemma, we prove the theorem for a larger family of $\mathcal{F}$, and the remaining unproved ones are the beginning of the next lemma. Finally we prove the theorem for all possible signature sets $\mathcal{F}$. In some cases, the attempt to interpolate all unary signatures does not work. In these cases, we employ yet another (the third) starting point of #P-hardness, which is the problem of counting Perfect Matchings on 3-regular graphs [13]. We reduce the Perfect Matching problem also by polynomial interpolation, which is done in Lemma 11.2. However, note that counting Perfect Matchings is computable in polynomial time for planar graphs [22, 23, 25], therefore our dichotomy theorem for Holant^c problems here does not extend to planar graphs as our dichotomy theorem for Holant^* problems does.

References

[1] Lenore Blum, Felipe Cucker, Michael Shub, and Steve Smale. *Complexity and real computation*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1998.

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

[3] Andrei A. Bulatov. The complexity of the counting constraint satisfaction problem. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, *ICALP (1)*, volume 5125 of *Lecture Notes in Computer Science*, pages 646–661. Springer, 2008.

[4] Andrei A. Bulatov and Víctor Dalmau. Towards a dichotomy theorem for the counting constraint satisfaction problem. In *FOCS*, pages 562–571. IEEE Computer Society, 2003.

[5] Andrei A. Bulatov and Martin Grohe. The complexity of partition functions. In Josep Díaz, Juhani Karhumäki, Arto Lepistö, and Donald Sannella, editors, *ICALP*, volume 3142 of *Lecture Notes in Computer Science*, pages 294–306. Springer, 2004.

[6] Andrei A. Bulatov and Martin Grohe. The complexity of partition functions. *Theor. Comput. Sci.*, 348(2-3):148–186, 2005.

[7] Jin-Yi Cai, Xi Chen, and Pinyan Lu. Graph homomorphisms with complex values: A dichotomy theorem. *Submitted to STOC 09*.

[8] Jin-Yi Cai and Pinyan Lu. Holographic algorithms: from art to science. In *STOC '07: Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 401–410, New York, NY, USA, 2007. ACM.

[9] Jin-Yi Cai and Pinyan Lu. On symmetric signatures in holographic algorithms. In Wolfgang Thomas and Pascal Weil, editors, *STACS*, volume 4393 of *Lecture Notes in Computer Science*, pages 429–440. Springer, 2007.

[10] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. The complexity of complex weighted boolean #csp. *Submitted to STOC 09*.

[11] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Holographic algorithms by fibonacci gates and holographic reductions for hardness. In *FOCS '08: Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, Washington, DC, USA, 2008. IEEE Computer Society.

[12] N. Creignou, S. Khanna, and M. Sudan. *Complexity classifications of boolean constraint satisfaction problems*. SIAM Monographs on Discrete Mathematics and Applications, 2001.

[13] P. Dagum and M. Luby. Approximating the permanent of graphs with large factors. *Theor. Comput. Sci.*, 102:283–305, 1992.

[14] C. T. J. Dodson and T. Poston. *Tensor Geometry*. Graduate Texts in Mathematics 130. Springer-Verlag, New York, 1991.

[15] Martin E. Dyer, Leslie Ann Goldberg, and Mark Jerrum. The complexity of weighted boolean #csp. *CoRR*, abs/0704.3683, 2007.

[16] Martin E. Dyer, Leslie Ann Goldberg, and Mike Paterson. On counting homomorphisms to directed acyclic graphs. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *ICALP (1)*, volume 4051 of *Lecture Notes in Computer Science*, pages 38–49. Springer, 2006.

[17] Martin E. Dyer, Leslie Ann Goldberg, and Mike Paterson. On counting homomorphisms to directed acyclic graphs. *J. ACM*, 54(6), 2007.

[18] Martin E. Dyer and Catherine S. Greenhill. The complexity of counting graph homomorphisms (extended abstract). In *SODA*, pages 246–255, 2000.

[19] Martin E. Dyer and Catherine S. Greenhill. Corrigendum: The complexity of counting graph homomorphisms. *Random Struct. Algorithms*, 25(3):346–352, 2004.

[20] M. Freedman, L. Lovász, and A. Schrijver. Reflection positivity, rank connectivity, and homomorphism of graphs. *J. AMS*, 20:37–51, 2007.

[21] Leslie Ann Goldberg, Martin Grohe, Mark Jerrum, and Marc Thurley. A complexity dichotomy for partition functions with mixed signs. *CoRR*, abs/0804.1932, 2008.

[22] P. W. Kasteleyn. The statistics of dimers on a lattice. *Physica*, 27:1209–1225, 1961.

[23] P. W. Kasteleyn. Graph theory and crystal physics. In (F. Harary, editor, *Graph Theory and Theoretical Physics*, pages 43–110. Academic Press, London, 1967.

[24] Ker-I Ko. *Complexity theory of real functions*. Birkhauser Boston Inc., Cambridge, MA, USA, 1991.

[25] H. N. V. Temperley and M. E. Fisher. Dimer problem in statistical mechanics c an exact result. *Philosophical Magazine*, 6:1061C 1063, 1961.

[26] Salil P. Vadhan. The complexity of counting in sparse, regular, and planar graphs. *SIAM J. Comput.*, 31(2):398–427, 2001.

[27] Leslie G. Valiant. The complexity of enumeration and reliability problems. *SIAM J. Comput.*, 8(3):410–421, 1979.

[28] Leslie G. Valiant. Quantum circuits that can be simulated classically in polynomial time. *SIAM J. Comput.*, 31(4):1229–1254, 2002.

[29] Leslie G. Valiant. Holographic algorithms (extended abstract). In *FOCS '04: Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 306–315, Washington, DC, USA, 2004. IEEE Computer Society.

[30] Leslie G. Valiant. Accidental algorthims. In *FOCS '06: Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 509–517, Washington, DC, USA, 2006. IEEE Computer Society.

[31] 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.

Appendix

5 Some Preliminary Results

A signature from $\mathcal{F}_q$ at a vertex is considered a basic realizable function. Instead of a single vertex, we can use graph fragments to generalize this notion. An $\mathcal{F}_q$-gate Γ is also a tuple $(H, \mathcal{F}_q, \pi)$, where $H = (V, E, D)$ is a graph with some dangling edges D. (See Figure 2 for one example.) Other than

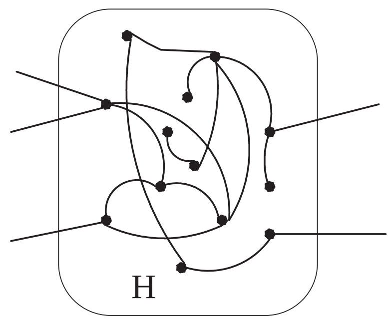

Figure 2: An $\mathcal{F}$-gate with 5 dangling edges.

these dangling edges, an $\mathcal{F}_q$-gate is the same as a signature grid. The role of dangling edges is similar to that of external nodes in Valiant's notion [28, 29], however we allow more than one dangling edges for a node. In $H = (V, E, D)$ each node is assigned a function in $\mathcal{F}_q$ by the mapping π (we do not consider "dangling" leaf nodes at the end of a dangling edge among these), E is the set of regular edges, denoted as $1, 2, \ldots, m$, and D is the set of dangling edges, denoted as $m+1, m+2, \ldots, m+n$. Then we can define a function for this $\mathcal{F}$-gate $\Gamma = (H, \mathcal{F}_q, \pi)$,

$$\Gamma(y_1, y_2, \ldots, y_n) = \sum_{x_1 x_2 \cdots x_m} H(x_1 x_2 \cdots x_m y_1 y_2 \cdots y_n),$$

where $(y_1, y_2, \ldots, y_n) \in \{0,1\}^n$ denotes an assignment on the dangling edges and $H(x_1 x_2 \cdots x_m y_1 y_2 \cdots y_n)$ denotes the value of the signature grid on an assignment of all edges. We will also call this function the signature of the $\mathcal{F}_q$-gate Γ. An $\mathcal{F}_q$-gate can be used in a signature grid as if it is just a single node with the particular signature. We note that even for a very simple signature set $\mathcal{F}_q$, the signatures for all $\mathcal{F}_q$-gates can be quite complicated and expressive. Matchgate signatures are an example.

Using the idea of $\mathcal{F}_q$-gates, we can reduce one holant problem to another. Let g be the signature of some $\mathcal{F}_q$-gate Γ. Then $\text{Holant}(\mathcal{F}_q \cup \{g\}) \leq_T \text{Holant}(\mathcal{F}_q)$. The reduction is quite simple. Given an instance of $\text{Holant}(\mathcal{F}_q \cup \{g\})$, by replacing every appearance of g by an $\mathcal{F}_q$-gate Γ, we get an instance of $\text{Holant}(\mathcal{F}_q)$. Since the signature of Γ is g, the values for these two signature grids are identical.

We give some propositions which are useful in the proof. We first give one more definition.

Definition 5.1. *Given a symmetric signature* $[x_0, x_1, \ldots, x_n]$ *and any* l, h *where* $0 \leq l < h \leq n$, *we call* $[x_l, x_{l+1}, \ldots, x_h]$ *a sub-signature of* $[x_0, x_1, \ldots, x_n]$ *with arity* $h - l$.

Proposition 5.1. *If* $[x_l, x_{l+1}, \ldots, x_h]$ *is a sub-signature of* $[x_0, x_1, \ldots, x_n]$ *and* $\text{Holant}^*([x_l, x_{l+1}, \ldots, x_h])$ *is #P-Complete, then* $\text{Holant}^*([x_0, x_1, \ldots, x_n])$ *is #P-Complete. This is also true for Holant*^{c}.

Proof. We use unary signatures $[1, 0]$, $[0, 1]$ and $[x_0, x_1, \ldots, x_n]$ to simulate a sub-signature $[x_l, x_{l+1}, \ldots, x_h]$. We connect l dangling edges to $[x_0, x_1, \ldots, x_n]$ with the unary signature $[0, 1]$, and connect $n-h$ dangling edges with the unary signature $[1, 0]$. □

Proposition 5.2. *Let $n \geq 3$ and let $[x_0, x_1, \ldots, x_n]$ be a non-degenerate symmetric signature. Then for any $m = 2, 3, \ldots, n-1$, there exists a non-degenerate sub-signature of $[x_0, x_1, \ldots, x_m]$ with arity m, unless the signature is of the form $[x_0, 0, \ldots, 0, x_n]$.*

Proof. Since $\operatorname{rank} \begin{bmatrix} x_0 & \ldots & x_{n-1} \\ x_1 & \ldots & x_n \end{bmatrix} = 2$, there must be non-zero entries among $x_0, x_1, \ldots, x_n$. If all entries are non-zero, then either $\operatorname{rank} \begin{bmatrix} x_0 & \ldots & x_{n-2} \\ x_1 & \ldots & x_{n-1} \end{bmatrix} = 2$, or $\operatorname{rank} \begin{bmatrix} x_1 & \ldots & x_{n-1} \\ x_2 & \ldots & x_n \end{bmatrix} = 2$. Otherwise they are both of rank 1, and being non-zero, the second row is a multiple of the first row in both matrices. Since they share at least one column, and being non-zero, this multiplier must be the same, which says that $\operatorname{rank} \begin{bmatrix} x_0 & \ldots & x_{n-1} \\ x_1 & \ldots & x_n \end{bmatrix} = 1$, a contradiction. Then we use induction to complete the proof.

Now suppose there are zero entries. Consider $x_1, \ldots, x_{n-1}$. Since $[x_0, \ldots, x_n]$ is not of the form $[x_0, 0, \ldots, 0, x_n]$, there must be some $1 \leq i \leq n-1$ such that $x_i \neq 0$. Find a $x_i \neq 0$, for some $1 \leq i \leq n-1$, such that a neighbor $x_{i-1} = 0$ or $x_{i+1} = 0$. Now any submatrix containing $\begin{bmatrix} \ldots & x_{i-1} & x_i & \ldots \\ \ldots & x_i & x_{i+1} & \ldots \end{bmatrix}$, has rank 2. □

Proposition 5.3. *Let $\mathcal{F}$ be a set of signatures and M be a 2×2 orthogonal matrix, i.e., $MM^{\mathrm{T}} = I_2$. We define a new set of signatures $\mathcal{F}_M$ as follows:*

$$\mathcal{F}_M = \{Y | Y = M^{\otimes n} X, \text{ where } X \text{ is a signature in } \mathcal{F} \text{ with arity } n.\}$$

Then Holant$^*(\mathcal{F})$ *and* Holant$^*(\mathcal{F}_M)$ *have the same complexity.*

Proof. This is a direct corollary of Theorem 2.2, given the fact that the unary signature set $\mathcal{U}$ is invariant under any transformation M. □

Proposition 5.4. *Let $[x_0, x_1, x_2, x_3]$ be a symmetric signature of arity three, expressed as*

$$[x_0, x_1, x_2, x_3] = A(\alpha_1, \alpha_2)^{\otimes 3} + B(\beta_1, \beta_2)^{\otimes 3}.$$

Then

$$\det \begin{bmatrix} x_0 & x_1 \\ x_1 & x_2 \end{bmatrix} = AB \det \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}^2 \alpha_1 \beta_1,$$

$$\det \begin{bmatrix} x_1 & x_2 \\ x_2 & x_3 \end{bmatrix} = AB \det \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}^2 \alpha_2 \beta_2,$$

and

$$\det \begin{bmatrix} x_0 - x_2 & x_1 \\ x_1 - x_3 & x_2 \end{bmatrix} = AB \det \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}^2 (\alpha_1 \beta_1 + \alpha_2 \beta_2).$$

In particular, if $AB \neq 0$ and $\det \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} \neq 0$, then, $[x_0, x_1, x_2, x_3]$ is non-degenerate and $\det \begin{bmatrix} x_0 - x_2 & x_1 \\ x_1 - x_3 & x_2 \end{bmatrix} = 0$ iff the inner product $\alpha_1 \beta_1 + \alpha_2 \beta_2 = 0$. In this case, the unique (upto a scalar) non-zero solution (a, b) to

$$\begin{bmatrix} x_0 - x_2 & x_1 \\ x_1 - x_3 & x_2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = 0$$

is $a = \alpha_1 \beta_1 = -\alpha_2 \beta_2$ and $b = \alpha_1 \beta_2 + \alpha_2 \beta_1$.

Proof: A straightforward calculation. ∎

Proposition 5.5. *Let $[x_0, x_1, x_2, x_3]$ be a symmetric signature of arity three, expressed as*

$$[x_0, x_1, x_2, x_3] = A(\alpha_1, \alpha_2)^{\otimes 3} + B(\beta_1, \beta_2)^{\otimes 3},$$

where $AB \neq 0$ and $\det \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} \neq 0$. Then, $[x_0, x_1, x_2, x_3]$ is of the form $[x, y, -x, -y]$ implies that $\alpha_1^2 + \alpha_2^2 = \beta_1^2 + \beta_2^2 = 0$.

Proof: The equations $x_0 + x_2 = 0$ and $x_1 + x_3 = 0$ are respectively $A\alpha_1(\alpha_1^2 + \alpha_2^2) + B\beta_1(\beta_1^2 + \beta_2^2) = 0$ and $A\alpha_2(\alpha_1^2 + \alpha_2^2) + B\beta_2(\beta_1^2 + \beta_2^2) = 0$. Viewed as a linear equation system on A and B, we obtain that its determinant is zero:

$$\det \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} (\alpha_1^2 + \alpha_2^2)(\beta_1^2 + \beta_2^2) = 0.$$

It follows that at least one of the factors $\alpha_1^2 + \alpha_2^2 = 0$ or $\beta_1^2 + \beta_2^2 = 0$. In the first case, back to the linear equation system, since (β_1, β_2) is a non-zero vector, we get the second factor $\beta_1^2 + \beta_2^2 = 0$ as well. Similarly starting with $\beta_1^2 + \beta_2^2 = 0$ we also get $\alpha_1^2 + \alpha_2^2 = 0$.

∎

6 Proof of Lemma 3.2

We remark that, since $[x_0, x_1, x_2, x_3]$ is non-degenerate, the constant pair (a, b) is unique upto a scalar factor.

Proof: Our plan of the proof is as follows: We will show that the counting problem Holant* $(\{[x_0, x_1, x_2, x_3], [y_0, y_1, y_2]\})$ when restricted to instances where the input graph is bipartite and all degree three nodes are on one side and given the signature $[x_0, x_1, x_2, x_3]$, and all degree two nodes are on the other side and given the signature $[y_0, y_1, y_2]$, is still #P-complete. (There might be any number of degree one nodes, on either side of the bipartite graph, assigned any unary signatures. We may denote this problem Holant$^*([x_0, x_1, x_2, x_3] \mid [y_0, y_1, y_2])$.) We show its #P-complete by a holographic reduction where $[y_0, y_1, y_2]$ is (covariantly) transformed to $[1, 0, 1]$ and $[x_0, x_1, x_2, x_3]$ is (contravariantly) transformed to some $[u_0, u_1, u_2, u_3]$. Note that $[1, 0, 1]$ can be replaced by an edge, and the unary signatures are transformed to some other unary signatures. Thus the complexity of the problem Holant$^*([x_0, x_1, x_2, x_3] \mid [y_0, y_1, y_2])$ is the same as Holant$^*([u_0, u_1, u_2, u_3])$. We then apply Lemma 3.1 to $[u_0, u_1, u_2, u_3]$.

Our first step is to show that there exists a non-singular $T = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}$ such that $(T^{-1})^{\otimes 2}(y_0, y_1, y_1, y_2)^{\mathrm{T}}$ $= (1, 0, 0, 1)^{\mathrm{T}}$, but $[u_0, u_1, u_2, u_3] = [x_0, x_1, x_2, x_3]T^{\otimes 3}$ is not of the form $[x, y, -x, -y]$ (exception (2) in Lemma 3.1). We note that $[u_0, u_1, u_2, u_3]$ is non-degenerate, for otherwise, $[x_0, x_1, x_2, x_3] = [u_0, u_1, u_2, u_3](T^{-1})^{\otimes 3}$ would also be degenerate.

Since $[y_0, y_1, y_2]$ is non-degenerate, clearly there exists a basis $T = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}$ such that $(y_0, y_1, y_1, y_2)^{\mathrm{T}} = T^{\otimes 2}(1, 0, 0, 1)^{\mathrm{T}}$. Let $[u_0, u_1, u_2, u_3] = [x_0, x_1, x_2, x_3]T^{\otimes 3}$.

We assume for a contradiction that $[u_0, u_1, u_2, u_3]$ is of the form $[x, y, -x, -y]$. Then we can find scalars A and B, such that

$$[u_0, u_1, u_2, u_3] = A(1, i)^{\otimes 3} + B(1, -i)^{\otimes 3}.$$

Let $\Delta = \det(T) \neq 0$, we have $T^{-1} = \Delta^{-1} \begin{bmatrix} \beta_2 & -\beta_1 \\ -\alpha_2 & \alpha_1 \end{bmatrix}$. Then it follows that

$$[x_0, x_1, x_2, x_3] = [u_0, u_1, u_2, u_3](T^{-1})^{\otimes 3} = A\Delta^{-3}(\beta_2 - \alpha_2 i, -\beta_1 + \alpha_1 i)^{\otimes 3} + B\Delta^{-3}(\beta_2 + \alpha_2 i, -\beta_1 - \alpha_1 i)^{\otimes 3}.$$

Since $[x_0, x_1, x_2, x_3]$ is non-degenerate, $AB \neq 0$. By the assumption on $[x_0, x_1, x_2, x_3]$ and Proposition 5.4, we have

$$0 = (\beta_2 - \alpha_2 i)(\beta_2 + \alpha_2 i) + (-\beta_1 + \alpha_1 i)(-\beta_1 - \alpha_1 i) = \alpha_1^2 + \alpha_2^2 + \beta_1^2 + \beta_2^2.$$

And after a scaling, we also have $a = \alpha_2^2 + \beta_2^2 = -\alpha_1^2 - \beta_1^2$ and $b = -2(\alpha_1\alpha_2 + \beta_1\beta_2)$. Now

$$(y_0, y_1, y_1, y_2)^{\mathrm{T}} = T^{\otimes 2}(1, 0, 0, 1)^{\mathrm{T}} = (\alpha_1^2 + \beta_1^2, \alpha_1\alpha_2 + \beta_1\beta_2, \alpha_1\alpha_2 + \beta_1\beta_2, \alpha_2^2 + \beta_2^2)^{\mathrm{T}}.$$

Then $[y_0, y_1, y_2] = [-a, -b/2, a]$ is of the form $[2a\lambda, b\lambda, -2a\lambda]$, a contradiction.

This completes our first step. Our second step is to take care of the other exceptional case (exception (1) in Lemma 3.1).

This exceptional case is that there exist two constants a', b' (not both zero), such that $a'u_0 + b'u_1 - a'u_2 = 0$ and $a'u_1 + b'u_2 - a'u_3 = 0$. We split this exceptional case into two cases depending on whether $a = 0$ or $a \neq 0$.

If $a = 0$ then $b \neq 0$ and $x_1 = x_2 = 0$. As $[x_0, 0, 0, x_3]$ is non-degenerate, $x_0x_3 \neq 0$. We can write

$$[x_0, 0, 0, x_3] = x_0 \begin{pmatrix}1 & 0\end{pmatrix}^{\otimes 3} + x_3 \begin{pmatrix}0 & 1\end{pmatrix}^{\otimes 3},$$

and

$$[u_0, u_1, u_2, u_3] = [x_0, 0, 0, x_3]T^{\otimes 3} = x_0(\begin{pmatrix}1 & 0\end{pmatrix} T)^{\otimes 3} + x_3(\begin{pmatrix}0 & 1\end{pmatrix} T)^{\otimes 3}.$$

The existence of a', b' (not both zero), such that $a'u_0 + b'u_1 - a'u_2 = 0$ and $a'u_1 + b'u_2 - a'u_3 = 0$, implies that the matrix $\begin{bmatrix} u_2 - u_0 & u_1 \\ u_3 - u_1 & u_2 \end{bmatrix}$ is degenerate. By Proposition 5.4 we have the inner product of $(1, 0)T$ and $(0, 1)T$ is zero, i.e.,

$$\alpha_1\alpha_2 + \beta_1\beta_2 = 0.$$

But noticing that $y_1 = \alpha_1\alpha_2 + \beta_1\beta_2$ and $a = 0$, we have $ay_0 + by_1 - ay_2 = 0$, a contradiction.

Finally we consider the case $a \neq 0$. By a scaling we can assume $a = 1$. Then $x_2 = bx_1 + x_0$, $x_3 = bx_2 + x_1 = b^2x_1 + bx_0 + x_1$. Assume temporarily that $b \neq \pm 2i$, then the recurrence has two distinct eigenvalues, and therefore Proposition 5.4 applies. We can then calculate that the determinant

$$\det \begin{bmatrix} u_0 - u_2 & u_1 \\ u_1 - u_3 & u_2 \end{bmatrix} = \det \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}^2 (\alpha_1^2 + \beta_1^2 + b(\alpha_1\alpha_2 + \beta_1\beta_2) - \alpha_2^2 - \beta_2^2)(bx_0x_1 - x_1^2 + x_0^2) = 0.$$

However now we dispense with the temporary assumption that $b \neq \pm 2i$; for this is a polynomial identity valid for all values except for a finite number of exceptional points of b. Thus it holds for all values.

We know that $\det \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} \neq 0$; the second factor $\alpha_1^2 + \beta_1^2 + b(\alpha_1\alpha_2 + \beta_1\beta_2) - \alpha_2^2 - \beta_2^2$ is exactly $y_0 + by_1 - y_2$ which is non-zero. So we have $bx_0x_1 - x_1^2 + x_0^2 = 0$. Since $bx_0x_1 + x_0^2 - x_1^2 = x_0x_2 - x_1^2$, this says that $[x_0, x_1, x_2, x_3]$ is degenerate, a contradiction. □

7 Proof of Lemma 3.3

Proof: The overall plan of the proof of this Lemma is the same as for Lemma 3.2.

We can choose a basis $T = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}$ such that $(y_0, y_1, y_1, y_2)^{\mathrm{T}} = T^{\otimes 2}(1, 0, 0, 1)^{\mathrm{T}}$. Then under the same basis T, $[x, y, -x, -y]$ will become $[u_0, u_1, u_2, u_3] = [x, y, -x, -y]T^{\otimes 3}$. Using a new parametrization, we have

$$[x, y, -x, -y] = A(1, i)^{\otimes 3} + B(1, -i)^{\otimes 3},$$

and

$$[u_0, u_1, u_2, u_3] = [x, y, -x, -y]T^{\otimes 3} = A(\alpha_1 + \alpha_2 i, \beta_1 + \beta_2 i)^{\otimes 3} + B(\alpha_1 - \alpha_2 i, \beta_1 - \beta_2 i)^{\otimes 3}.$$

Since we assumed that $[x, y, -x, -y]$ is non-degenerate, we have $AB \neq 0$.

By this holographic reduction, we only need to prove that Holant$^*(\{[u_0, u_1, u_2, u_3]\})$ is #P-Complete. Since $[u_0, u_1, u_2, u_3]$ is non-degenerate, by Lemma 3.1, we have to consider two cases.

If $[u_0, u_1, u_2, u_3]$ is of the form $[x', y', -x', -y']$, then by Proposition 5.5

$$0 = (\alpha_1 + \alpha_2 i)^2 + (\beta_1 + \beta_2 i)^2 = (\alpha_1 - \alpha_2 i)^2 + (\beta_1 - \beta_2 i)^2.$$

This gives $\alpha_1\alpha_2 + \beta_1\beta_2 = 0$ and $\alpha_1^2 + \beta_1^2 = \alpha_2^2 + \beta_2^2$. Then by $(y_0, y_1, y_1, y_2)^{\mathrm{T}} = T^{\otimes 2}(1, 0, 0, 1)^{\mathrm{T}}$, we know that $[y_0, y_1, y_2]$ is of the form $[\lambda, 0, \lambda]$, a contradiction.

Now we consider the second exceptional case from Lemma 3.1: there exist two constants a, b (not both zero) such that $au_0 + bu_1 - au_2 = 0$ and $au_1 + bu_2 - au_3 = 0$. By Proposition 5.4 we have

$$0 = (\alpha_1 + \alpha_2 i)(\alpha_1 - \alpha_2 i) + (\beta_1 + \beta_2 i)(\beta_1 + \beta_2 i) = \alpha_1^2 + \alpha_2^2 + \beta_1^2 + \beta_2^2.$$

This gives $y_0 + y_2 = \alpha_1^2 + \beta_1^2 + \alpha_2^2 + \beta_2^2 = 0$, a contradiction. $\square$

8 Proof of Lemma 3.4

Proof: We assume that Holant$^*([x_0, x_1, x_2, \ldots, x_n])$ is not #P-Complete. Then we prove that it must be in one of the two cases.

Case A: We first consider the case where for every $k = 0, 1, \ldots, n-2$, the sub-signature $[x_k, x_{k+1}, x_{k+2}]$ is non-degenerate. Then by Lemma 3.1, there are the following two exceptional cases (1) and (2) to be considered for $[x_0, x_1, x_2, x_3]$.

(1) There is a non-zero pair (a, b) such that $ax_0 + bx_1 - ax_2 = 0$ and $ax_1 + bx_2 - ax_3 = 0$. Such a non-zero pair (a, b) must be unique up to a scalar factor, since otherwise $[x_0, x_1, x_2, x_3]$ is all 0. Then by Lemma 3.2, for every $k = 0, 1, \ldots, n-2$, $ax_k + bx_{k+1} - ax_{k+2} = 0$ or $[x_k, x_{k+1}x_{k+2}]$ is of the form $[2a\lambda, b\lambda, -2a\lambda]$.

First we claim $a \neq 0$. Suppose otherwise $a = 0$, then $x_2 = x_1 = 0$, from the linear relation. But then surely $x_3 \neq 0$, since $[x_0, x_1, x_2, x_3]$ is non-degenerate. Consider when $k = 2$. If $ax_2 + bx_3 - ax_4 = 0$, then we have a contradiction. Hence $[x_2, x_3, x_4]$ is of the form $[2a\lambda, b\lambda, -2a\lambda]$. But then $x_4 = x_2 = 0$. It follows that $[x_1, x_2, x_3, x_4] = [0, 0, x_3, 0]$. This signature satisfies Lemma 3.1, then it follows that Holant$^*([x_1, x_2, x_3, x_4])$ is #P-Complete, contrary to assumption.

So we can assume $a = 1$. Now we show that the form $[2a\lambda, b\lambda, -2a\lambda]$ can not appear among all $[x_k, x_{k+1}, x_{k+2}]$. This will conclude that we are in the first of two cases in Lemma 3.3.

Suppose the form $[2a\lambda, b\lambda, -2a\lambda]$ does occur. Such a λ certainly is non-zero, otherwise it is degenerate.

After a scaling, we may take this form $[2, b, -2]$. If it occurs as $[x_k, x_{k+1}, x_{k+2}]$ for $k = 0$ or 1, then together with $ax_0 + bx_1 - ax_2 = 0$ and $ax_1 + bx_2 - ax_3 = 0$, we have $b = \pm 2i$. But then $[2, b, -2]$ is degenerate. Thus, if this form $[2, b, -2]$ appears as some $[x_k, x_{k+1}, x_{k+2}]$, consider the minimal k where it appears, then $k \geq 2$, and there is a sub-signature $[x, y, 2, b, -2]$ with arity 4, where $ax + by - 2a = 0$ and $ay + 2b - ab = 0$. So the sub-signature is $[b^2 + 2, -b, 2, b, -2]$. There is a sub-signature of $[-b, 2, b, -2]$. Applying Lemma 3.3 to $[-b, 2, b, -2]$ and $[b^2 + 2, -b, 2]$, we have $b^2 + 2 = -2$ or $b^2 + 2 = 2$. The first case gives that $[2, b, -2]$ is degenerate. So we have $b^2 + 2 = 2$ and thus $b = 0$. In this case, the sub-signature is $[2, 0, 2, 0, -2]$. Using a unary signature $[1, 1]$, we will have $[2, 2, 2, -2]$. By Lemma 3.1, this is #P-Complete, again contrary to assumption.

(2) If $[x_0, x_1, x_2, x_3]$ is of the form $[x, y, -x, -y]$, then by Lemma 3.3, for every $k = 0, 1, \ldots, n-2$, we have $x_k + x_{k+2} = 0$ or $[x_k, x_{k+1}, x_{k+2}]$ is of the form $[\lambda, 0, \lambda]$. We prove that $[\lambda, 0, \lambda]$ can not appear. This will conclude that we are in the second of two cases in Lemma 3.3.

Suppose the form $[\lambda, 0, \lambda]$ does appear among $[x_k, x_{k+1}, x_{k+2}]$. It is easy to see that if it occurred t $k = 0$ or 1, then $[x_0, x_1, x_2, x_3] = [x, y, -x, -y]$ is degenerate. Then look at the first $k \geq 2$ where it occurred, and to its left we have $x_k + x_{k+2} = 0$. It follows that there must be a sub-signature (after a scaling) of the form $[1, 0, -1, 0, -1]$. Now we can use a similar trick with a unary signature $[1, 1]$ and get $[1, -1, -1, -1]$, which is #P-Complete.

This completes the proof of Case A.

Case B: There exists some $0 \leq k \leq n-2$, such that the sub-signature $[x_k, x_{k+1}, x_{k+2}]$ is degenerate.

If the signature is of the form $[x_0, 0, 0, \ldots, 0, x_n]$, we can choose $(a, b) = (0, 1)$. In the following we assume the signature is not of this form.

By Proposition 5.2, there exist a $s = 0, 1, \ldots, n-3$, such that a sub-signature $\sigma_3 = [x_s, x_{s+1}, x_{s+2}, x_{s+3}]$ of arity three is non-degenerate. Starting from a degenerate sub-signature $\tau_2 = [x_k, x_{k+1}, x_{k+2}]$ of arity two, we want to find a degenerate sub-signature τ_2' of arity two which is a sub-signature of σ_3. If $s \leq k \leq s+1$, then τ_2 is already a sub-signature of σ_3. Otherwise, $k < s$ or $k > s+1$. W.l.o.g, suppose $k < s$. Consider the sub-signature $[x_{k+1}, x_{k+2}, x_{k+3}]$. If it is degenerate, we can replace it for τ_2, and continue. If it is non-degenerate, then we can replace τ_2' by $[x_k, x_{k+1}, x_{k+2}, x_{k+3}]$, and it will also be non-degenerate.

Thus we can find a degenerate sub-signature of arity two which is a sub-signature of a non-degenerate sub-signature of arity three. This must be of the form $[s^2, sr, r^2, x]$ or $[y, s^2, sr, r^2]$. By symmetry, we only consider the first case. By Lemma 3.1, we have two cases.

For the first case, $(r^2 - s^2)r^2 - (x - sr)sr = 0$. This implies that $srx = r^4$. If $sr \neq 0$, $srx = r^4$ implies that $[s^2, sr, r^2, x]$ is degenerate, a contradiction. If $sr = 0$, then $r = 0$ by $srx = r^4$. Since $[s^2, sr, r^2, x]$ is non-degenerate, we must have $s^2 \neq 0$ and and $x \neq 0$. It is of form $[s^2, 0, 0, x]$. As $n > 3$, there must be entries to its left or to its right, say $[s^2, 0, 0, x, z]$. Consider the pair $[s^2, 0, 0, x]$ and $[0, x, z]$. By Lemma 3.2, the (upto scale) unique pair for $[s^2, 0, 0, x]$ is $(a, b) = (0, 1)$. If $z \neq 0$, then Lemma 3.2 would imply #P-completeness. If $z = 0$, then we have a sub-signature $[0, 0, x, 0]$. This also implies #P-completeness by Lemma 3.1.

Finally for the other case: $s^2 + r^2 = 0$ and $sr + x = 0$. Then the signature must be $[s^2, s^2 i, -s^2, -s^2 i]$ or $[s^2, -s^2 i, -s^2, s^2 i]$. Both are degenerate, a contradiction. $\square$

9 Proof of Theorem 3.1

Proof. The first easy class is obvious.

For the second easy class, if $a = 0$, all functions in $\mathcal{F}$ have form $[x_1, 0, \ldots, 0, x_n]$ or $[0, x_2, 0]$, so Holant$^*(\mathcal{F})$ is obviously easy, and if $a \neq 0$, we consider two subcases, that $a + b\lambda - a\lambda^2 = 0$ has one double root or two distinct roots. Suppose λ_1 and λ_2 are the two roots. If they are equal, all functions in $\mathcal{F}$ with arity larger than 2, have form $(1, \lambda_1)^{\otimes n}$, which is equal to n functions $[1, \lambda_1]$ applied to each inputs, so Holant$^*(\mathcal{F})$ is easy. If $\lambda_1 \neq \lambda_2$, because $\lambda_1 \lambda_2 = -1$, neither of them is i or $-i$. All functions in $\mathcal{F}$ have form $F = u \begin{bmatrix} 1 \\ \lambda_1 \end{bmatrix}^{\otimes n} + v \begin{bmatrix} 1 \\ \lambda_2 \end{bmatrix}^{\otimes n}$ or $G = [2a, b, -2a]$. We use nonsingular matrix $M = \begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{bmatrix}$ as base to do holographic reduction on $\#\mathcal{F}|\{=_2\}$, which is the bipartite form of problem Holant$^*(\mathcal{F})$. Under this base, F is turned into $M^{\otimes n} F = u \begin{bmatrix} 1 + \lambda_1^2 \\ 0 \end{bmatrix}^{\otimes n} + v \begin{bmatrix} 0 \\ 1 + \lambda_2^2 \end{bmatrix}^{\otimes n}$, and G is turned into $M^{\otimes 2} G = [0, (4a^2 + b^2)/a, 0]$, and $=_2 = [1, 0, 1]$ is turned into $[1, 0, 1](M^{-1})^{\otimes 2} = [1 + \lambda_2^2, 0, 1 + \lambda_1^2]$. This reduction reduces Holant$^*(\mathcal{F})$ to an obviously easy problem.

For the last easy class, we will prove that the Holant is zero unless the input graph is bipartite, and there is a holographic algorithm for bipartite graphs. If there is a function $[1,0,1]$ applied to two variables x and y, we just merge them into one variable and remove this function. Hence, all functions on vertices of input graphs have form $[u,v,-u,-v,\ldots]$. If the input graph $G(V,E)$ is not bipartite, there is a circle $v_1,e_1,\ldots,v_k,e_k,v_1$ of odd length. We perfectly match all assignments for E into pairs (T,T'), such that, T and T' assign the same values on $E-\{e_1,\ldots,e_k\}$, and opposite values on $\{e_1,\ldots,e_k\}$. Under T and T', all functions on $V-\{v_1,\ldots,v_k\}$ give the same value, and if $T(e_j)\neq T(e_{j+1})$, the function on v_j give the same value, and if $T(e_j)=T(e_{j+1})$, the function on v_j give the opposite values. Consider $T(e_1),T(e_2),\ldots,T(e_k),T(e_1)$, there must be even many times value change in this sequence. Since k is odd, there are odd many v_js, whose functions give opposite values under T and T'. Hence, in the summation, the contributions of T and T' are canceled. If there input graph is bipartite, the problem is $\#\mathcal{F}|\mathcal{F}$. $\#\mathcal{F}|\mathcal{F}$ is turned into an easy problem $\#\{[x_1,0,\ldots,0,x_n]\}|\{[x_1,0,\ldots,0,x_n]\}$ by holographic reduction under base $M=\begin{bmatrix}1 & i\\ 1 & -i\end{bmatrix}$.

Now we prove that if $\mathcal{F}$ does not belong to any of the easy classes, Holant$^*(\mathcal{F})$ is #P-hard. Assume for a contradiction that $\mathcal{F}$ fits in none of the exceptional cases yet it is still not #P-hard. We consider two cases. First, if all the sub-signatures with arity 2 of every signature in $\mathcal{F}$ are non-degenerate, then the same proof as in Lemma 3.4 works. Second, suppose this is not the case, again by the proof of Lemma 3.4, there exists a signature of the form $[x_0,0,0,\ldots,0,x_n]$ for some arity $n\geq 3$. By connecting to $n-3$ unary signatures of $(1,1)$, we have a new signature $[x_0,0,0,x_n]$ with arity 3. Then by Lemma 3.2 (with $a=0,b=1$), we know that all the signatures in $\mathcal{F}$ have this form $[x,0,0,\ldots,0,y]$. But this is of the second easy case with $a=0,b=1$, a contradiction. □

10 Polynomial Interpolation

In this section, we discuss the interpolation method we use for dichotomy theorem for Holant^c. Polynomial interpolation is a powerful tool in the study of counting problems initiated by Valiant [27] and further developed by Vadhan, Dyer and Greenhill [26, 18]. The method we use here is essentially the same as Vadhan [26].

For some set of signatures $\mathcal{F}$, we want to show that for all unary signatures $f=[x,y]$, we have Holant$(\mathcal{F}\cup\{[x,y]\})\leq_T$ Holant$(\mathcal{F})$. Let $\Omega=(G,\mathcal{F}\cup\{[x,y]\},\pi)$. We want to compute Holant$_\Omega$ in polynomial time using an oracle for Holant$(\mathcal{F})$.

Let V_f be the subset of vertices in G assigned f in Ω. Suppose $|V_f|=n$. We can classify all 0-1 assignments σ in the holant sum according to how many vertices in V_f whose incident edge is assigned a 0 or a 1. Then the holant value can be expressed as

$$\text{Holant}_\Omega=\sum_{0\leq i\leq n} c_i x^i y^{n-i}, \tag{2}$$

where c_i is the sum over all edge assignments σ, of products of evaluations at all $v\in V(G)-V_f$, where σ is such that exactly i vertices in V_f have their incident edges assigned 0 (and $n-i$ have their incident edges assigned 1.) If we can evaluate these c_i, we can evaluate Holant$_\Omega$.

Now suppose $\{G_s\}$ is a sequence of $\mathcal{F}$-gates, and each G_s has one dangling edge. Denote the signature of G_s by $f_s=[x_s,y_s]$, for $s=0,1,\ldots$. If we replace each occurrence of f by f_s in Ω we get a new signature grid Ω_s, which is a instance of Holant$(\mathcal{F})$, with

$$\text{Holant}_{\Omega_s}=\sum_{0\leq i\leq n} c_i x_s^i y_s^{n-i}. \tag{3}$$

One can evaluate Holant_{Ω_s} by oracle access to $\text{Holant}(\mathcal{F})$. Note that the same set of values c_i occurs. We can treat c_i in (3) as a set of unknowns in a linear system. The idea of interpolation is to find a suitable sequence $\{f_s\}$ such that the evaluation of Holant_{Ω_s} gives a linear system (3) of full rank, from which we can solve all c_i.

In this paper, the sequence $\{G_s\}$ will be constructed recursively using suitable gadgetry. There are two gadgets in a recursive construction: one gadget has arity 1, giving the initial signature $g = [x_0, y_0]$; the other has arity 2, giving the recursive iteration. It is more convenient to use a 2×2 matrix A to denote it. So we can recursively connect them as in Figure 3 and get $\{G_s\}$.

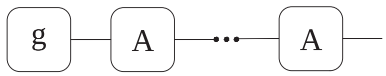

Figure 3: Recursive construction.

The signatures of $\{G_s\}$ have the following relation,

$$\begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_{s-1} \\ y_{s-1} \end{bmatrix}, \tag{4}$$

where $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ and $g = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$.

We call this gadget pair (A, g) a recursive construction. It follows from lemma 6.1 in [26] that

Lemma 10.1. *Let α, β be the two eigenvalues of A. If the following three conditions are satisfied*

1. $\det(A) \neq 0$;
2. *g is not a column eigenvector of A (nor the zero vector);*
3. *α/β is not a root of unity.*

Then the recursive construction (A, g) can be used to interpolate all the unary signatures.

This similar interpolation method also works for signatures with larger arity but only have two dimensions of freedom. For example, all the signatures are of form $[0, x, 0, y]$. This example is used in the proof of Lemma 11.2.

11 Proof of Theorem 4.1

In the dichotomy theorem for $\text{Holant}^*(\mathcal{F})$, we assume the arity of every signature in $\mathcal{F}$ is larger than one and all the signatures in $\mathcal{F}$ are non-degenerate. In $\text{Holant}^c(\mathcal{F})$, not all the unary signatures are freely available, so we can not assume that. In some case, the present of some unary signatures or degenerate signatures does change the complexity of the problem. However, we can still do some normalization here to make the proof clear. Since any degenerate signature $[x, y]^{\otimes k}$ can be replaced by the corresponding unary signature $[x, y]$ without change the complexity of the problem, we always assume that all the signatures in $\mathcal{F}$, whose arity is large than 1, is non-degenerate. Since $[1, 0]$ and $[0, 1]$ are freely available, we can construct any sub signature of the original signatures as well as any signatures realizable by some $\mathcal{F}$-gate.

The polynomial algorithm for $\mathcal{F} \subseteq \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$ is very non-trivial. In [10], we extend this family to asymmetric cases and give a unified algorithm there. So we omit the algorithm here. In spirited by the tractable algorithm for this generalized family, we prove a dichotomy for all the boolean #CSP in [10]. Here, $\mathcal{F}$ is a set of real signatures, so technically, we only need to consider the real signatures in the three families $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$. We give the form of these three families as above because we believe that it is the right form to express them and we also conjecture that this statement is also true for complex symmetric signatures. The main idea of the proof is to interpolate all the unary function. A brief description of this polynomial interpolation method is given in Section 10 . If we can do that, we can reduce the problem Holant$^*(\mathcal{F})$ to Holant$^c(\mathcal{F})$ and finish the proof. In some cases, we can not do that and we prove the theorem separately.

Lemma 11.1. *If we can construct a gadget with signature $[a, b, c]$, where $b^2 \neq ac$, $b \neq 0$ and $a + c \neq 0$, then we can interpolate all the unary function. (And hence Theorem 4.1 holds.)*

Proof. we use the interpolation method as described in Section 10. We consider two recursive constructions ($\begin{bmatrix} a & b \\ b & c \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$) and ($\begin{bmatrix} a & b \\ b & c \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$), and argue that at least one of them will success given the conditions on a, b, c. We use A to denote $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$. Since $b^2 \neq ac$, A is non-degenerate, the first condition of Lemma 10.1 is satisfied for both of the two recursive constructions. If both $[1, 0]$ and $[0, 1]$ are column eigenvectors of A, then $b = 0$, a contradiction. So at least for one of the two recursive constructions, the second condition of Lemma 10.1 is satisfied. Since A is a real symmetric matrix, so both its eigenvalues are real. If the ratio of two real numbers is a root of unity, they must be the same or opposite to each other. If the two eigenvalues are the same, we will have $b = 0$ and $a = c$, a contradiction. If the two eigenvalues are opposite to each other, then we have $a+c = 0$, also a contradiction. Therefore, the third condition of Lemma 10.1 is also satisfied for both of the two recursive constructions. To sum up, at least one of the two recursive constructions satisfies all the conditions of Lemma 10.1. This completes the proof. □

If we can construct a gadget with binary symmetric signature $[a, b, c]$, which satisfies all the conditions in Lemma 11.1, then, we are done. Most of the cases, we prove the theorem by interpolating all the unary signatures. However, in some cases, we are not able to do that. For example, if all the signature have the parity condition, then all the unary signatures we can realize have form $[a, 0]$ or $[0, a]$, so we can not interpolate all the unary signatures. For these cases, our start point is the following lemma.

Lemma 11.2. *If $a \neq \pm 1$, Holant$^c([0, 1, 0, a])$ is #P Hard.*

Proof. Our start point here is that Holant$([0, 1, 0, 0])$ is #P-Hard. This is exactly the perfect matching problem in 3-regular graph [13]. So the problem is #P-Hard if $a = 0$.

Now assume that $a \notin \{-1, 0, 1\}$, and we use this signature to interpolate all the signature of form $[0, 1, 0, x]$, in particular, we can interpolate $[0, 1, 0, 0]$ and finish the hardness reduction.

The recursive construction is depicted by Figure 4. By a simple parity argument, every $\mathcal{F}$-gate N_i has a signature of form $[0, x_i, 0, y_i]$. After some calculation, we can get that they satisfy the following recursive relation:

$$\begin{bmatrix} x_{i+1} \\ y_{i+1} \end{bmatrix} = \begin{bmatrix} 3(a^2+1) & (a^3+a) \\ 3(a^3+a) & a^6+1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix}.$$

In this case, the signature we want to interpolate are of arity 3. But since all of them are of form $[0, x_i, 0, y_i]$ with two dimensions freedom. We can also use the interpolation method as in Section 10. Let $A = \begin{bmatrix} 3(a^2+1) & (a^3+a) \\ 3(a^3+a) & a^6+1 \end{bmatrix}$, then $(A, [1, a]^{\mathrm{T}})$ forms a recursive construction. Since $\det(A) = 3(a^4-1)^2 \neq 0$, the first condition holds. Its characterize equation is $X^2-(a^6+3a^2+4)X+3(a^4-1)^2 = 0$.

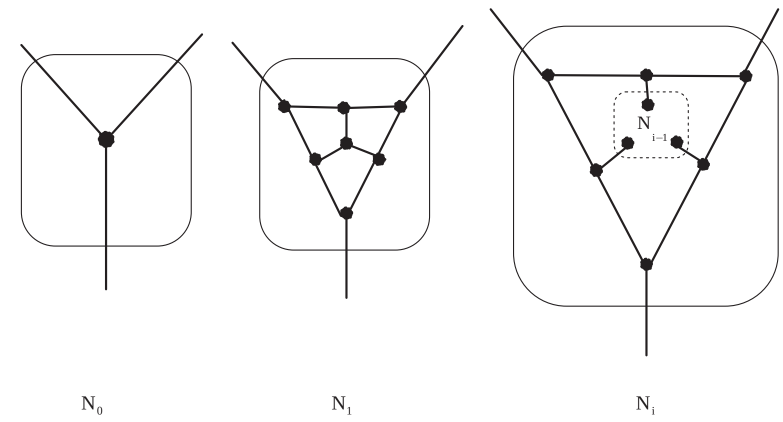

Figure 4: The recursive construction. The signature of every vertex in the gadget is $[0, 1, 0, a]$.

For this quadratic equation, $\Delta = (a^6 - 3a^2 - 2)^2 + 12(a + a^3)^2 > 0$. So A has two distinct real eigenvalues. The sum of the two eigenvalues is $a^6 + 3a^2 + 4$ which is larger than zero. So they are not opposite to each other. Therefore, the ratio of these two eigenvalues is not a root of unity and the third condition holds. Consider the second condition, if the initial vector $[1, a]$ is a column eigenvectors of A. We have $A \begin{bmatrix} 1 \\ a \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ a \end{bmatrix}$, where λ is one eigenvalue of A. From this, we will conclude that $a(a^2 - 1)(a^4 - 1) = 0$, which will not happen given $a \notin \{-1, 0, 1\}$. To sum up, this recursive relation satisfies all the three conditions of Lemma 10.1 and can be used to interpolate all the signatures of form $[0, 1, 0, x]$. This completes the proof. □

We define some families of symmetric signatures, which will be used in our proof.

$$\begin{aligned} \mathcal{G}_1 &= \{[a, 0, 0, \cdots, 0, b] | ab \neq 0\} \\ \mathcal{G}_2 &= \{[x_0, x_1, \cdots, x_k] | \forall i \text{ is even}, x_i = 0 \text{ or } \forall i \text{ is odd}, x_i = 0\} \\ \mathcal{G}_3 &= \{[x_0, x_1, \cdots, x_k] | \forall i, x_i + x_{i+2} = 0\} \end{aligned}$$

We note that $\mathcal{G}_1$, $\mathcal{G}_2$ and $\mathcal{G}_3$ are supersets of $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$ respectively. The following several lemmas all have the form "If $\mathcal{F} \not\subseteq \mathcal{A}$, then Theorem 4.1 holds." After proving that, in the later lemma, we only need to consider the case that $\mathcal{F} \subseteq \mathcal{A}$.

Lemma 11.3. *If $\mathcal{F} \not\subseteq \mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3$, then Theorem 4.1 holds.*

Proof. Since $\mathcal{F} \not\subseteq \mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3$, there exists a $f \in \mathcal{F}$ and $f \notin \mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3$. Since all the unary signatures are in $\mathcal{G}_3$, the arity of f is larger than 1 and f is non-degenerate. There are two cases according to whether f has a zero entry or not.

(1) f has some zero entries. If there exists a sub signature of f has the form $[0, a, b]$ or $[a, b, 0]$, where $ab \neq 0$, then we are done by Lemma 11.1. Otherwise, we can conclude that there is no two successive non-zero entries. So the signature f is like this $[0^{i_0} x_1 0^{i_1} x_2 0^{i_2} \cdots x_k 0^{i_k}]$, where $x_j \neq 0$ and for all $1 \leq j \leq k-1$, $i_j \geq 1$. If for all $1 \leq j \leq k-1$, i_j is odd, then $f \in \mathcal{G}_2$, a contradiction. Otherwise there exists a sub signature of form $[x, 0, 0, \cdots, 0, y]$, where $xy \neq 0$ and there are even number of 0s between x and y. If this is the whole f, then $f \in \mathcal{G}_1$, a contradiction. So there is one 0 before x or after y. By symmetric, we assume there is a 0 before x, so we have a sub signature $[0, x, 0, 0, \cdots, 0, y]$, whose arity

is even and larger than 3. We call its dangling edges $1, 2, \cdots, 2k$. Then for every $i = 1, 2, \cdots, k-1$, we connect dangling edges $2i+1$ and $2i+2$ together to a regular edge. After that, we have a $\mathcal{F}$-gate with arity 2, and its signature is $[0, x, y]$. Then we are done by Lemma 11.1.

(2) f has no zero entry. We only need to prove that we can construct a function $[a', b', c']$ satisfying the three conditions in Lemma 11.1. Suppose all sub signatures of f with arity 2 do not satisfy all the three conditions. For each sub-signature $[a', b', c']$, either $a' + c' = 0$, or $b'^2 = a'c'$. If all of them satisfy $a' + c' = 0$, then $f \in \mathcal{G}_3$. A contradiction. If all of them satisfy $b'^2 = a'c'$, then f is degenerate. A contradiction. W.l.o.g, we can assume there is a sub-signature $[a, b, c, d]$ of f, such that $a + c = 0$, $b + d \neq 0$, and $c^2 = bd$. Combining two $[a, b, c, d]$, we can get a function $[a', b', c'] = [a^2 + 2b^2 + c^2, ab + 2bc + cd, b^2 + 2c^2 + d^2]$. $b' = c(b + d) \neq 0$. $a' + c' = a^2 + 3b^2 + 3c^2 + d^2 > 0$. Because $c^2 = bd$, $a'c' - b'^2 = a^2b^2 + 2a^2c^2 + a^2d^2 + 2b^4 + 4b^2c^2 + 2b^2d^2 \neq 0$. We are done by Lemma 11.1. ∎

Lemma 11.4. *If $\mathcal{F} \not\subseteq \mathcal{G}_1 \cup \mathcal{F}_2 \cup \mathcal{G}_3$, then Theorem 4.1 holds.*

Proof. If $\mathcal{F} \not\subseteq \mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3$, then by Lemma 11.3, we are done. Otherwise, there exists a signature $f \in \mathcal{F} \subseteq \mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3$ and $f \notin \mathcal{G}_1 \cup \mathcal{F}_2 \cup \mathcal{G}_3$. Then it must be the case that $f \in \mathcal{G}_2$. Note that every signature with arity less than 3 in $\mathcal{G}_2$ is also contained in $\mathcal{G}_1 \cup \mathcal{G}_3$, so f is of arity larger than 2. Since $f \notin \mathcal{G}_1$, there is some non-zero in the middle of the signature f, after normalization, we can assume there is a sub signature of form $[0, 1, 0, x]$ (or $[x, 0, 1, 0]$). If $x \neq \pm 1$, then by Lemma 11.2, we know the problem is #P-Hard and we are done. Otherwise, for every such pattern , we have $x = \pm 1$. Since $f \notin \mathcal{F}_2$, then there is some sub signature $[0, 1, 0, -1]$ and because $f \notin \mathcal{G}_3$, there is some sub signature $[0, 1, 0, 1]$. Therefore, there is a sub signature $[1, 0, 1, 0, -1]$ of f. Then we can construct a $\mathcal{F}$-gate as Figure 5 , whose signature is $[8, 0, 4, 0]$. So by Lemma 11.2, we know that the problem is #P-Hard and

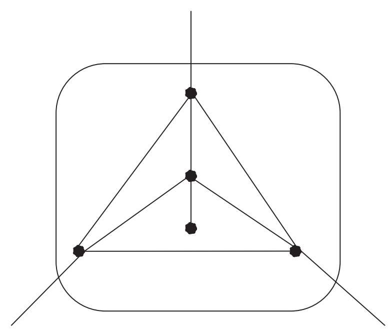

Figure 5: The signature of every degree 4 vertex in the gadget is $[1, 0, 1, 0, -1]$. And the signature of the degree 1 vertex in the gadget is $[1, 0]$.

we are done. This completes the proof. ∎

Lemma 11.5. *If $\mathcal{F} \not\subseteq \mathcal{G}_1 \cup \mathcal{G}_3$, then Theorem 4.1 holds.*

Proof. If $\mathcal{F} \not\subseteq \mathcal{G}_1 \cup \mathcal{F}_2 \cup \mathcal{G}_3$, then by Lemma 11.4, we are done. Otherwise, there exists a signature $f \in \mathcal{F} \subseteq \mathcal{G}_1 \cup \mathcal{F}_2 \cup \mathcal{G}_3$ and $f \notin \mathcal{G}_1 \cup \mathcal{G}_3$. Then it must be the case that $f \in \mathcal{F}_2$. Note that every signature with arity less than 3 in $\mathcal{F}_2$ is also contained in $\mathcal{G}_1 \cup \mathcal{G}_3$, so f is of arity larger than 2. Then f has a sub signature $[1, 0, 1, 0]$ or $[0, 1, 0, 1]$. By symmetry, we assume it is $[1, 0, 1, 0]$. If $\mathcal{F} \subseteq \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$, then Theorem 4.1 trivially holds and there is nothing to prove. If not, there exists a signature g is in $\mathcal{G}_1 - \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$ ($\subseteq \mathcal{G}_1 - \mathcal{F}_1$) or $\mathcal{G}_3 - \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$ ($\subseteq \mathcal{G}_3 - \mathcal{F}_3$).

For the first case, $g \in (\mathcal{G}_1 - \mathcal{F}_1)$, then after a scale, g is of form $[1, 0, 0, \cdots, b]$, where $b \notin \{-1, 0, 1\}$. If the arity of g is odd, we can realize $[1, b]$. (We connect its every two dangling edges into one edge and left one dangling edge.) Then connecting this unary signature to one dangling edge of $[1, 0, 1, 0]$, we can realize a binary signature $[1, b, 1]$. Then by Lemma 11.1, Theorem 4.1 holds. If the arity of g is even, we can realize $[1, 0, b]$ (left two dangling edges). By connecting one of its dangling edge to one dangling edge of $[1, 0, 1, 0]$, we can have a new ternary signature $[1, 0, b, 0]$. By lemma 11.2, we know the problem is #P-Hard.

For the second case $g \in (\mathcal{G}_3 - \mathcal{F}_3)$, then g has a sub signature of form $[1, b]$, where $b \notin \{-1, 0, 1\}$. By the same argument as above, Theorem 4.1 holds. This completes the proof. □

Lemma 11.6. *If $\mathcal{F} \not\subseteq \mathcal{G}_1 \cup \mathcal{F}_3$, then Theorem 4.1 holds.*

Proof. If $\mathcal{F} \not\subseteq \mathcal{G}_1 \cup \mathcal{G}_3$, then by Lemma 11.5, we are done. Otherwise, there exists a signature $f \in \mathcal{F} \subseteq \mathcal{G}_1 \cup \mathcal{G}_3$ and $f \notin \mathcal{G}_1 \cup \mathcal{F}_3$. Then it must be the case that $f \in \mathcal{G}_3$, and f has a sub signature of form $[1, a, -1]$, where $a \notin \{-1, 0, 1\}$.

If $\mathcal{F} \subseteq (\{[1, 0, 1]\} \cup \mathcal{G}_3)$, then Holant$^*(\mathcal{F})$ is polynomial time computable by Theorem 3.1 and as a result Theorem 4.1 trivially holds and there is nothing to prove.

If not, there exists a signature $g \in \mathcal{F} \subseteq \mathcal{G}_1 \cup \mathcal{G}_3$ and $g \notin \{[1, 0, 1]\} \cup \mathcal{G}_3$. Then it mush be the case that $g \in \mathcal{G}_1$. The arity of g is large than 1.

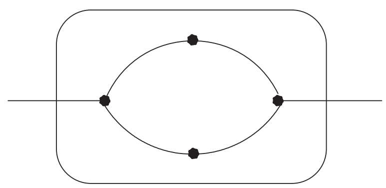

Figure 6: The function on degree 2 nodes is $[1, a, -1]$, and the function on degree 3 nodes is $[1, 0, 0, b]$.

If the arity of g is 2, then g is of form $[1, 0, b]$, where $b \notin \{-1, 0, 1\}$. Connecting two signatures $[1, 0, b]$ to the both sides of one binary signature $[1, a, -1]$, we can get a new binary signature $[1, ab, -b^2]$. It satisfies all the conditions of Lemma 11.1, and we are done. If the arity of g is larger than 2, then we can always realize a signature $[1, 0, 0, b]$, where $b \neq 0$. (We connect the unary signature $[1, a]$ to all its dangling edges except the three ones.) Then we can use an $\mathcal{F}$-gate as Figure 6. Its signature is $[1, a^2 b, b^2]$, and by Lemma 11.1, we are done. This completes the proof. □

By the above lemmas, the only case we have to handle is that $\mathcal{F} \subseteq \mathcal{G}_1 \cup \mathcal{F}_3$. This is done by the following lemma. which completes the proof of Theorem 4.1.

Lemma 11.7. *If $\mathcal{F} \subseteq \mathcal{G}_1 \cup \mathcal{F}_3$, then Theorem 4.1 holds.*

Proof. If $\mathcal{F} \subseteq \mathcal{F}_1 \cup \mathcal{F}_3$ or $\mathcal{F} \subseteq \mathcal{U} \cup \mathcal{F}_3 \cup \{[1, 0, 1]\}$, then Holant$^c(\mathcal{F})$ is computable in polynomial time and Theorem4.1 holds. Otherwise, there are two cases. There exists $f \in \mathcal{F}$ such that $f \in \mathcal{G}_1$ but $f \notin \mathcal{F}_1 \cup \mathcal{U} \cup \mathcal{F}_3$. So the arity of f is larger than 1. W.l.o.g, we can assume f has form $[1, 0, a]$ or $[1, 0, 0, a]$, where $a \notin \{-1, 0, 1\}$. (By connecting its dangling edges together except two or three depends on the parity of the arity of f.) The other case is that there exists $f_1, f_2 \in \mathcal{G}_1$ such that $f_1 \in \mathcal{F}_1$ but $f_1 \notin \mathcal{U} \cup \mathcal{F}_3 \cup \{[1, 0, 1]\}$, and $f_2 \in \mathcal{U}$ but $f_2 \notin \mathcal{F}_1$. So the arity of f_1 is larger than 2 and with form $[1, 0, 0, \cdots, \pm 1]$, and f_2 is of form $[1, a']$, where $a' \notin \{-1, 0, 1\}$. By connecting all the dangling edges of f_1 except two with f_2, we can construct a $\mathcal{F}$-gate with signature of form $[1, 0, a]$, where $a \notin \{-1, 0, 1\}$.This is one of the above two forms.

If $\mathcal{F} \subseteq \mathcal{G}_1 \cup \{[0,1,0]\}$, then $\operatorname{Holant}^c(\mathcal{F})$ is computable in polynomial time and Theorem 4.1 holds. Otherwise, there exists $g \in \mathcal{F} \subseteq \mathcal{G}_1 \cup \mathcal{F}_3$, and $g \notin \mathcal{G}_1 \cup \{[0,1,0]\}$. Then g must have one of the following sub signatures: $[1,1,-1], [1,-1,-1], [1,0,-1,0], [0,1,0,-1]$. By symmetric, we only need to consider two cases $[1,1,-1]$ and $[1,0,-1,0]$.

According to f and g, we have four cases. If $f = [1,0,a]$ and $g = [1,1,-1]$, then connecting them together into a chain fgf, we can realize $[1,a,-a^2]$. By Lemma 11.1, we are done. If $f = [1,0,a]$ and $g = [1,0,-1,0]$, for each of dangling edges of g, we extend it by one copy of f. Then we can realize $[1,0,-a^2,0]$. So by lemma 11.2, the problem is #P-hard. If $f = [1,0,0,a]$ and $g = [1,1,-1]$, we can connect a unary signature $[1,1]$ (sub signature of g) to one dangling edge of f, and realize a binary signature $f = [1,0,a]$. This reduce to the first case, which has been proved. If $f = [1,0,0,a]$ and $g = [1,0,-1,0]$, we can realize a unary signature $[1,a]$ from f and then connect this unary signature to one dangling edge of g to realize $[1,-a,-1]$. Note that $[1,-a,-1] \notin \mathcal{G}_1 \cup \mathcal{F}_3$, by Lemma 11.6, we are done. □

In the above proof, especially in Lemma 11.1, we use the fact that $\mathcal{F}$ is a set of real signatures. However, we believe that this statement is also true for complex symmetric signatures. So we have the following conjecture.

Conjecture 11.1. *Let $\mathcal{F}$ be a set of symmetric signatures over $\mathbb{C}$. Then* $\operatorname{Holant}^c(\mathcal{F})$ *is computable in polynomial time if (1)* $\operatorname{Holant}^*(\mathcal{F})$ *is computable in polynomial time or (2) $\mathcal{F} \subseteq \mathcal{F}_1 \cup \mathcal{F}_2 \cup \mathcal{F}_3$. Otherwise,* $\operatorname{Holant}^c(\mathcal{F})$ *is #P-Complete.*