A Dichotomy for Real Weighted Holant Problems

Sangxia Huang
KTH Royal Institute of Technology
Stockholm, Sweden
sangxia@csc.kth.se

Pinyan Lu
Microsoft Research Asia
Beijing, China
pinyanl@microsoft.com

***Abstract*—Holant is a framework of counting characterized by local constraints. It is closely related to other well-studied frameworks such as #CSP and Graph Homomorphism. An effective dichotomy for such frameworks can immediately settle the complexity of all combinatorial problems expressible in that framework. Both #CSP and Graph Homomorphism can be viewed as sub-families of Holant with the additional assumption that the equality constraints are always available. Other sub-families of Holant such as Holant* and Holant^c problems, in which we assume some specific sets of constraints to be freely available, were also studied. The Holant framework becomes more expressive and contains more interesting tractable cases with less or no freely available constraint functions, while, on the other hand, it also becomes more challenging to obtain a complete characterization of its time complexity. Recently, complexity dichotomy for a variety of sub-families of Holant such as #CSP, Graph Homomorphism, Holant* and Holant^c were proved. The dichotomy for the general Holant framework, which is the most desirable, still remains open. In this paper, we prove a dichotomy for the general Holant framework where all the constraints are real symmetric functions. This setting already captures most of the interesting combinatorial problems defined by local constraints, such as (perfect) matching, independent set, vertex cover and so on. This is the first time a dichotomy is obtained for general Holant Problems without any auxiliary functions.**

One benefit of working with Holant framework is some powerful new reduction techniques such as Holographic reduction. Along the proof of our dichotomy, we introduce a new reduction technique, namely realizing a constraint function by approximating it. This new technique is employed in our proof in a situation where it seems that all previous reduction techniques fail, thus this new idea of reduction might also be of independent interest. Besides proving dichotomy and developing new technique, we also obtained some interesting by-products. We prove a dichotomy for #CSP restricting to instances where each variable appears a multiple of d times for any d. We also prove that counting the number of Eulerian-Orientations on $2k$-regular graphs is #P-hard for any $k \geq 2$.

***Keywords*-Holant problem; #CSP; complexity dichotomy; reduction technique**

I. INTRODUCTION

In order to study the complexity of counting problems, several interesting frameworks characterizing local properties have been proposed. One is called counting Constraint Satisfaction Problems (#CSP) [1]–[9]. Another well studied framework is called H-coloring or Graph Homomorphism, which can be viewed as a special case of #CSP problems [10]–[17]. Recently, inspired by Valiant's Holographic Algorithms [18], [19], a new refined framework called Holant Problems [20], [21] was proposed. 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 *dichotomy theorems* (i.e., every problem in the class is either in P or #P-hard) [22]. Having a dichotomy is an important property for these languages since in general, Ladner [23] proved that if P $\neq$ NP, or in our case P $\neq$ #P, then such a dichotomy for NP (or #P) is *false*.

We give a brief description of the Holant framework here and a more formal definition is given in Section II. A *signature grid* $\Omega = (G, \mathcal{F}, \pi)$ is a tuple, where $G = (V, E)$ is an undirected graph, $\mathcal{F}$ is a set of functions. In this paper, we study the case where the functions map sets of Boolean variables to some value. Usually, the range of the functions are complex numbers $\mathbb{C}$ or real numbers $\mathbb{R}$ as in [1], [2], [4]–[8], [20], [21], but it is also interesting to consider functions with finite range, such as counting the number of solutions modulo some integer k, as studied in [24]–[28]. The mapping $\pi : V \to \mathcal{F}$ labels each $v \in V(G)$ with a function $f_v \in \mathcal{F}$, where the arity of f_v equals the degree of v. We consider all edge assignments (0-1 assignments in this paper, since we are considering Boolean functions). An assignment σ for every $e \in E$ gives an evaluation $\prod_{v \in V} f_v(\sigma\mid_{E(v)})$, where $\sigma\mid_{E(v)}$ denotes the substring of σ where only bits corresponding to incident edges of v are chosen. The counting problem on the instance Ω is to compute

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

We use the notation Holant($\mathcal{F}$) to denote the class of Holant problems where all functions are taken from $\mathcal{F}$. For example, consider the PERFECT MATCHING problem on G. This problem corresponds to attaching the EXACT-ONE function at every vertex of G — for each 0-1 edge assignment, the product evaluates to 1 when the assignment is a perfect matching, and 0 otherwise, therefore summing over all 0-1 edge assignments gives us the number of perfect matchings in G. If we use the AT-MOST-ONE function at

1093-0159/12 $26.00 © 2012 IEEE
DOI 10.1109/CCC.2012.16

every vertex, then we count all (not necessarily perfect) matchings. This framework can also express the partition function of a system, which is well studied in the statistical physics community, see for example the Ising model [29].

The Holant framework is closely related to other frameworks such as #CSP and Graph Homomorphism. In fact, in some sense, the Holant framework provides a unified perspective for different frameworks of counting problems. For example, the #CSP framework can be viewed as a special case of the Holant framework in which equality relations of any arity are always assumed to be available in addition to the stated constraints. A dichotomy for complex weighted Boolean #CSP was discovered and proved with the help from the study of general Holant problem [21]. On the other hand, #CSP excludes the expression of certain important problems such as graph matchings, which, in contrast, are expressible in the Holant framework. Besides #CSP, another two important special families are Holant* Problems, in which we assume that all unary functions are available, and Holant^c Problems, where we only assume two special unary functions — IS-ZERO function $\mathbf{0}$ and IS-ONE function $\mathbf{1}$ — to be available. For all the above families, dichotomy for complex symmetric functions were proved [21], [30]. However, a dichotomy for general Holant family remains open before the current work. The framework becomes more expressive in this general setting and, as we proved in this work, there are more tractable families. On the other hand, the proof for a dichotomy also became more challenging. A major source of difficulty is the lack of flexibility when we construct gadgets for reduction. One exception is a dichotomy for the general Holant framework for symmetric function in the field $\mathbb{Z}_2$ [31]. A couple of recent works studied the complexity of Holant on regular graphs where all the vertices take a same function [30], [32]–[34]. These works can be viewed as a dichotomy for Holant without freely available functions but have the constraint that $\mathcal{F}$ only contains one single function. In these papers, due to the lack of freely available equality functions or unary function, the reduction become much more difficult and even sometimes require assist from computer [32]. The underlying goal of these two sequences of works is to finally get a dichotomy for Holant.

This work achieves this final goal at least partially. We prove a dichotomy theorem for Holant problem where all functions are symmetric (the values of the functions only depend on the Hamming weight of their inputs) and take real values. Real symmetric functions already capture most of the interesting combinatorial problems and physical system problems. This is the first dichotomy for the Holant framework for a broad set of functions without assuming any freely available functions. Our work uses previous results as our starting points. And we believe that it is an important step to finally achieve the goal to characterize the complexity of Holant problems for any function set $\mathcal{F}$ (complex weighted and asymmetric).

One of the main innovation in this work is a new way of doing reduction between counting problems. In previous works as well as our current work, there are three extensively-used reduction methods: (1) gadget construction, (2) polynomial interpolation, and (3) holographic transformation. However, due to the special structure of some functions, we might be in the case that all possible gadget constructions give trivial functions, therefore classical reduction methods might not work well. In Section V, we introduce a new reduction method — realizing a function by approximating it with sufficient precision. The main idea is to construct a series of gadgets that converge to another gadget extremely fast with only a polynomial overhead, so that we would be able to recover the true value in polynomial time by solving a constant dimension integer programming. Although this is still some kind of gadget construction, we do not, and probably cannot, realize the target function precisely. It is also different from polynomial interpolation in that in polynomial interpolation, the number of gadgets one constructs is usually linear in the size of the instance, which is not affordable in our construction because the gadget size grows exponentially, while the new approximation technique only needs a logarithmic number of gadgets due to the fast convergence rate.

Another contribution of this work is a dichotomy for #CSP where each variable appears a multiple of d times, for any positive integer d. We found some tractable cases which are #P-hard for general #CSP. These new cases are still closely related to the tractable cases for the general #CSP, and we characterize them in terms of holographic transformation.

We also prove that counting Eulerian orientation for $2k$-regular graphs is #P-hard. Note that the notion of Eulerian orientation is different from Eulerian circuits in that the former only considers the direction of the edges and thus different Eulerian circuits may correspond to a same Eulerian orientation. Previously, similar problems have been studied, such as counting Eulerian orientation and counting Eulerian circuits in general graphs [35], [36], and Eulerian circuits in regular graphs [37]. All of them were shown to be #P-hard. However, to the best of our knowledge, there is no previous result on counting Eulerian orientation in regular graphs, and we are not aware of any direct reductions between counting Eulerian orientations and counting Eulerian circuits. Instead, we use polynomial interpolation to reduce the calculation of Tutte Polynomial at certain points to counting Eulerian orientation. The construction is easy to analyze in the Holant framework. One of the intriguing part of this problem is that it arises as a very special case along our proof for which all reduction methods — including the approximation approach we introduced here — failed. Hence, this problem may also serve as a new starting point of reduction in future research.

II. PRELIMINARIES

In this section, we recall some basic definitions and results. Let $\mathcal{F}$ be a set of functions. A *signature grid* is a tuple $\Omega = (G, \mathcal{F}, \pi)$, where $G = (V, E)$ is an undirected graph, and $\pi : V \to \mathcal{F}$ labels each $v \in V$ with a function $f_v \in \mathcal{F}$ where the arity of f_v equals the degree of v. The Holant problem on instance Ω is to compute

$$\mathrm{Holant}_\Omega = \sum_{\sigma:E\to\{0,1\}} \prod_{v\in V} f_v(\sigma \mid_{E(v)}),$$

a sum over all 0-1 edge assignments, of the products of the function evaluations at each vertex. Given a set of functions $\mathcal{F}$, we define the problem $\mathrm{Holant}(\mathcal{F})$:

- 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 function set $\mathcal{F}$.

A function f_v can be represented as a truth table. For functions with complex values, it will be more convenient to denote it as a tensor in $(\mathbb{C}^2)^{\otimes \deg(v)}$, or a vector in $\mathbb{C}^{2^{\deg(v)}}$, when we perform holographic transformations. We also call it a *signature*. Replacing a signature $f \in \mathcal{F}$ by a constant multiple cf, where $c \neq 0$, does not change the complexity of $\mathrm{Holant}(\mathcal{F})$, so we view f and cf as the same signature. A function f on k Boolean variables is *symmetric* if the value of the function depends only on the number of inputs that is assigned 1 (also known as the Hamming weight of the input), and 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, for example, we can express the following unary functions IS-ZERO $\mathbf{0} = [1, 0]$, IS-ONE $\mathbf{1} = [0, 1]$. We denote by $=_k$ the EQUALITY signature of arity k, then $(=_k) = [1, 0, \ldots, 0, 1]$ (with $(k-1)$ 0's). The binary disequality could be written as $[0, 1, 0]$. 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 $\lambda[x, y]^{\otimes k}$.

Some special families of Holant problems have already been widely studied. For example, if $\mathcal{F}$ contains all EQUALITY signatures $\{=_1, =_2, =_3, \cdots\}$, then this is exactly the weighted #CSP problem. In [21], the following two special families of Holant problems were introduced by assuming some signatures are freely available. Let $\mathcal{U}$ denote the set of all unary signatures. Then we define $\mathrm{Holant}^*(\mathcal{F}) = \mathrm{Holant}(\mathcal{F} \cup \mathcal{U})$. We use $\mathrm{Holant}^c(\mathcal{F})$ to denote the problem $\mathrm{Holant}(\mathcal{F} \cup \{\mathbf{0}, \mathbf{1}\})$.

There are several special classes of functions. A symmetric signature $[f_0, f_1, \ldots, f_k]$ is called a *generalized Fibonacci signature* if there exist a, b not both zero such that for all $i = 0, \ldots, k-2$ we have $af_i + bf_{i+1} - af_{i+2} = 0$. A k-ary function $f(x_1, \ldots, x_k)$ is *affine* if there exists a $k+1$ column Boolean matrix A, a set of dimension $k+1$ Boolean vectors $\{\alpha_1, \ldots, \alpha_n\}$, some complex number $c \neq 0$, such that f could be represented as $c\chi_{AX=0} i^{\sum_{j=1}^{n}\langle \alpha_j, X\rangle}$ where $X = (x_1, x_2, \ldots, x_k, 1)$, $\langle \cdot, \cdot \rangle$ is the inner product of two vectors, i is the imaginary unit with $i^2 = -1$, and χ is an indicator function such that $\chi_{AX=0}$ is 1 iff $AX = 0$. Note that both the matrix multiplication AX and the inner product are calculated in $\mathbb{Z}_2$. We use $\mathscr{A}$ to denote the set of all affine functions. We use $\mathscr{P}$ to denote the set of functions which can be expressed as a product of unary functions, binary equality functions and binary disequality functions. These two families capture exactly tractable #CSP problems.

Theorem II.1. *[21] Let $\mathcal{F}$ be a set of functions mapping Boolean inputs to complex numbers. Then #CSP($\mathcal{F}$) is #P-hard unless $\mathcal{F} \subseteq \mathscr{A}$ or $\mathscr{F} \subseteq \mathscr{P}$, in which case the problem is in P.*

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 on each edge an additional vertex labeled with the EQUALITY function $=_2$ on 2 inputs.

We use $\mathrm{Holant}(\mathcal{G}|\mathcal{R})$ 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 $\mathcal{G}$ or $\mathcal{R}$, respectively. An input instance for the bipartite Holant problem is a bipartite signature grid and is denoted as $\Omega = (H, \mathcal{G}|\mathcal{R}, \pi)$. Signatures in $\mathcal{G}$ are denoted by column vectors (or contravariant tensors); signatures in $\mathcal{R}$ are denoted by row vectors (or covariant tensors) [38].

One can perform (contravariant and covariant) tensor transformations on the signatures. We will define a simple version of holographic reductions, which are invertible. Suppose $T \in \mathbf{GL}_2(\mathbb{C})$ is a basis. We say that there is an (invertible) holographic reduction from $\mathrm{Holant}(\mathcal{G}|\mathcal{R})$ to $\mathrm{Holant}(\mathcal{G}'|\mathcal{R}')$, if the *contravariant* transformation $G' = T^{\otimes g} G$ and the *covariant* transformation $R = R' T^{\otimes r}$ map $G \in \mathcal{G}$ to $G' \in \mathcal{G}'$ and $R \in \mathcal{R}$ to $R' \in \mathcal{R}'$, and vice versa, where G and R have arity g and r respectively. Suppose that there is a holographic reduction from $\mathrm{Holant}(\mathcal{G}|\mathcal{R})$ to $\mathrm{Holant}(\mathcal{G}'|\mathcal{R}')$, mapping signature grid Ω to Ω', then $\mathrm{Holant}_\Omega = \mathrm{Holant}_{\Omega'}$. In particular, for invertible holographic reductions from $\mathrm{Holant}(\mathcal{G}|\mathcal{R})$ to $\mathrm{Holant}(\mathcal{G}'|\mathcal{R}')$, one problem is in P iff the other one is in P, and similarly one problem is #P-hard iff the other one is also #P-hard.

The following theorem is very useful as a way to normalize the given signature set $\mathcal{F}$.

Theorem II.2. *Let $\mathcal{F}$ be a set of signatures and M be a 2×2 orthogonal matrix. For any signature grid $\Omega = (G, \mathcal{F}, \pi)$, replacing every signature $F \in \mathcal{F}$ by $M^{\otimes n} F$, where n is the arity of F, we can get a new signature grid Ω'. Then*

$$\mathrm{Holant}_\Omega = \mathrm{Holant}_{\Omega'}.$$

Proof: First we use the standard technique to reformulate the signature grid $\Omega = (G, \mathcal{F}, \pi)$. 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 $\mathcal{F} \mid =_2$, we apply a holographic reduction on basis M. This will map a signature $F \in \mathcal{F}$ to $M^{\otimes n}F$, where n is the arity of F. It is an algebraic fact that the $=_2$ will map to itself. Then we view these (new) $=_2$ as an edge and ignore these vertices. This gives the signature grid Ω' as required. Due to the Holant theorem, its value is the same as Ω. ■

In the study of Holant problems, we will often transfer between bipartite and non-bipartite settings. When this does not cause confusion, we do not distinguish signatures between column vectors (or contravariant tensors) and row vectors (or covariant tensors). Whenever we write a transformation as $T^{\otimes n}F$ or $T\mathcal{F}$, we view the signatures as column vectors (or contravariant tensors); whenever we write a transformation as $FT^{\otimes n}$ or $\mathcal{F}T$, we view the signatures as row vectors (or covariant tensors).

Regarding models of computation for real numbers, strictly speaking we should restrict it to computable numbers [39], [40], or algebraic numbers. However this issue seems not essential for our result, and we will state our theorems assuming that we can compute $+$, $\times$ and solve linear equations in polynomial time for all real numbers used. If restricted to algebraic numbers, our proof in Section V can be simplified. But we do not restrict our result by exploiting the special properties of algebraic numbers.

III. Main Dichotomy Theorem and Proof Outline

For simplicity of statement, we define the following property for signature sets.

Definition III.1. *A set of signatures $\mathcal{F}$ is called* good *if there exists a 2×2 complex matrix T such that one of the following conditions is satisfied: $\mathcal{F}T^{-1} \subseteq \mathscr{A}$ and $T^{\otimes 2}[1,0,1]^T \in \mathscr{A}$; or $\mathcal{F}T^{-1} \subseteq \mathscr{P}$ and $T^{\otimes 2}[1,0,1]^T \in \mathscr{P}$.*

Our main theorem is the following.

Theorem III.2. *Let $\mathcal{F}$ be a set of symmetric signatures on Boolean variables with real values. Then $\operatorname{Holant}(\mathcal{F})$ is #P-hard unless the arity of any non-degenerate signature in $\mathcal{F}$ is no more than 2 or $\mathcal{F}$ is good, in which case it is computable in polynomial time.*

Proof Outline: If the arity of any non-degenerate signature in $\mathcal{F}$ is no more than 2, then $\operatorname{Holant}(\mathcal{F})$ is obviously tractable. The tractability of good $\mathcal{F}$ follows directly from the tractability of #CSP($\mathscr{A}$) and #CSP($\mathscr{P}$) after applying transformation under T. Therefore, we only need to prove the hardness part and we can assume that $\mathcal{F}$ contains a non-degenerate signature whose arity is at least 3.

Our starting point is Theorem IV.7, which states that the dichotomy holds if $\mathcal{F}$ contains a non-degenerate ternary function. To prove this, we use the relationship between Holant problems and #CSP problems. In some cases, we need a dichotomy for special #CSP problems where variables appear a multiple of 3 times. A general dichotomy for such #CSP is proved in Section IV.

The idea then is to realize a non-degenerate ternary function. In the previous dichotomy for Holant* or Holant^c problems, this step is easy because the freely available functions such as Is-Zero and Is-One enable us to realize sub-signatures with small arity. In our case, however, there is no longer freely available unary signature. We can only use signatures from the given set. Probably the simplest gadget one could construct is by adding self loops. For a signature with arity k, we can construct a signature with arity $k-2$ by adding a self loop. If the new signature is degenerate, then it has some very special structure and we can deal with that separately. Otherwise, we have constructed a smaller signature which is still non-degenerate. Repeat this process of adding self loop, and we will finally have a non-degenerate signature of arity 3 or 4, depending on the parity of k. The ternary case is proved in Theorem IV.7. It is not directly applicable for arity 4 case since we would not be able to construct any signature of odd arity from signatures of arity 4. We handle this in Theorem VI.3.

The idea of proving Theorem VI.3 is to realize degenerate binary signatures. A degenerate binary signature can be viewed as two unary signatures and in this sense we can realize a ternary function with the help of this unary signature. As stated in Lemma VI.1, we can show that the dichotomy holds if we have a non-degenerate 4-ary signature and one non-zero unary function. Similar to the ternary case, the proof makes use of the relation between Holant and #CSP.

The main remaining work is to realize a non-zero degenerate binary signature. We generalize the polynomial interpolation technique to achieve this. However, there are cases when this method fails, and for those cases, we use our new reduction tool of approximating. This is done in Section V. There is one exceptional case, namely $[1,0,1/3,0,1]$. By holographic reduction, this problem is equivalent to the Counting-Eulerian-Orientation in 4-regular graphs, which can be proved to be #P-hard. We sketch the proof of its hardness at the end of Section V.

Remark: We note that our main dichotomy for Holant is only for real valued functions. However, the dichotomy for the #CSP where variables appear a multiple of d times is for complex numbers. This is necessary to make it useful in the proof of our main dichotomy. Even starting from real Holant problem, we may come to the field of complex number after some holographic transformation. There are several places in the proofs in Section VI where we use some special properties of real numbers. We believe that the most essential part is polynomial interpolation. To make the interpolation work, we need that the ratio of the two eigenvalues of a 2×2 symmetric matrix is not a root of unity. For a real symmetric matrix, its two eigenvalues are also real. Since the only real roots of unity are ± 1, we can handle these two exceptions by a careful case-by-case analysis. For complex matrices, there are infinitely many exceptions. Overcoming this difficulty and extending our result to complex field is an interesting open question.

IV. #CSP WHERE VARIABLES APPEAR A MULTIPLE OF d TIMES

In this section, we consider a special family of #CSP problem, where the number of occurrences of each variable must be a multiple of d times (d is a given constant). For example, we can consider all the #CSP instance where each variable appears even number of times. We use #CSP$^d(\mathcal{F})$ to denote this problem. Clearly, if #CSP$(\mathcal{F})$ is polynomial time computable, then so is #CSP$^d(\mathcal{F})$. However, the reverse is not necessarily true. We use $\mathcal{T}_d$ to denote set $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & \omega \end{bmatrix} \mid \omega^d = 1 \right\}$. Then applying any $T \in \mathcal{T}_d$ to $\mathcal{F}$ will not change the value of a #CSP$^d(\mathcal{F})$ instance and as a result will not change the complexity of #CSP$^d(\mathcal{F})$. For example, #CSP$^3([1, \omega_3, -\omega_3^2])$, where ω_3 is the primitive third root of unity, is computable in polynomial time since #CSP$^3([1, 1, -1])$ is. On the other hand, note that #CSP$([1, \omega_3, -\omega_3^2])$ is #P-hard without the additional constraints on the number of occurrences of variables. For symmetric function set $\mathcal{F}$, we prove a dichotomy for #CSP$^d(\mathcal{F})$ which shows that these are essentially the only new tractable cases.

Theorem IV.1. *Let $d \geq 1$ be an integer and $\mathcal{F}$ be a set of symmetric functions taking complex values. Then $\#\mathrm{CSP}^d(\mathcal{F})$ is #P-hard unless there exist $T \in \mathcal{T}_{4d}$ such that $(T\mathcal{F}) \subset \mathscr{P}$ or $(T\mathcal{F}) \subset \mathscr{A}$, in which case the problem is in P.*

The following Theorem in [30] gives a reduction between #CSP and Holant, which will be used here as a starting point.

Theorem IV.2. *Consider the bipartite Holant instance $\mathrm{Holant}([1, 0, 0, 1] \cup \mathcal{G}_1 | \mathcal{G}_2)$. We assume that $\mathcal{G}_2$ contains a non-degenerate binary signature $[y_0, y_1, y_2]$. And in the case of $y_0 = y_2 = 0$, we further assume that $\mathcal{G}_2$ contains a unary signature $[a, b]$, where $ab \neq 0$. Then $\mathrm{Holant}([1, 0, 0, 1] \cup \mathcal{G}_1 | \mathcal{G}_2)$ is #P-hard unless there exist a $T \in \mathcal{T}_3$ such that $\mathcal{G}_1 T \cup T^{-1}\mathcal{G}_2 \subset \mathscr{P}$ or $\mathcal{G}_1 T \cup T^{-1}\mathcal{G}_2 \subset \mathscr{A}$, in which cases the problem is in P.*

Before proving Theorem IV.1, we prove in Lemma IV.5 that the conclusion holds if we have IS-ZERO ($[1, 0]$) and IS-ONE ($[0, 1]$) in addition. For general #CSP, one can assume freely available $[1, 0]$ and $[0, 1]$ by the nice pinning lemma from [5]. It is not obvious that this holds for #CSP^d. We start with the following special pinning lemma for #CSP^d.

Lemma IV.3.

$$\#\mathrm{CSP}^d(\mathcal{F}) \equiv_T \#\mathrm{CSP}^d(\mathcal{F} \cup \{[1, 0]^{\otimes d}, [0, 1]^{\otimes d}\}).$$

The proof is exactly the same as [5] so we omit here. The only thing one need to notice is that when adding auxiliary variables, it is important that they appear a multiple of d times, and in our case this is guaranteed by $[1, 0]^{\otimes d}$ and $[0, 1]^{\otimes d}$.

Now we proceed to show that we can still effectively realize the idea of pinning by a similar idea used in [41].

Lemma IV.4. *$\#\mathrm{CSP}^d(\mathcal{F})$ is #P-hard (or in P) iff $\#\mathrm{CSP}^d(\mathcal{F} \cup \{[1, 0], [0, 1]\})$ is #P-Hard (or in P).*

Proof: Obviously the first one can be reduced to the second one. Hence if the second problem is in P, so is the first. We have already proved a dichotomy theorem for the second one in Lemma IV.5. So now we may assume the second problem is #P-hard, and show that the first problem is also #P-hard.

We observe that in all the proofs in this paper and [42], when we prove the second problem to be #P-hard for any signature set, we reduce one of the following three problems to it by a chain of reductions: (a) Holant$([1, 0, 0, 1] | [1, 1, 0])$, (b) Holant$([1, 1, 0, 0])$, or (c) Holant$([0, 1, 0, 0])$ (VERTEX COVER or MATCHING or PERFECT MATCHING for 3-regular graph). There are only three reduction methods in this reduction chain, direct gadget construction, polynomial interpolation, and holographic reduction.

Given an instance G of Holant$([1, 0, 0, 1] | [1, 1, 0])$, Holant$([1, 1, 0, 0])$, or Holant$([0, 1, 0, 0])$, we consider the graph $G^{\otimes d}$, which denotes the disjoint union of d copies of G.

Notice that the value of Holant$([1, 0, 0, 1] | [1, 1, 0])$, Holant$([1, 1, 0, 0])$, or Holant$([0, 1, 0, 0])$ on the instance G is a non-negative integer, and the value on $G^{\otimes d}$ is its d-th power. So we can compute the value on G uniquely from its d-th power. Suppose the reduction chain on the instance G produced instances $G_1, G_2, \ldots, G_m$ of the second problem. The same reduction applied to $G^{\otimes d}$ produces instances of the form $G_1^{\otimes d}, G_2^{\otimes d}, \ldots, G_{m'}^{\otimes d}$. (We note that the reduction on $G^{\otimes d}$ may produce polynomially more instances than on G because of polynomial interpolation.)

For each $G_i^{\otimes d}$ as an instance of $\#\mathrm{CSP}^d(\mathcal{F} \cup \{[1, 0], [0, 1]\})$, the number of occurrences of $[0, 1]$ or $[1, 0]$ is a multiple of d. Hence, we can view it as an instance of $\#\mathrm{CSP}^d(\mathcal{F} \cup \{[1, 0]^{\otimes d}, [0, 1]^{\otimes d}\})$. By the assumption that $\#\mathrm{CSP}^d(\mathcal{F} \cup \{[1, 0], [0, 1]\})$ is hard, we conclude that $\#\mathrm{CSP}^d(\mathcal{F} \cup \{[1, 0]^{\otimes d}, [0, 1]^{\otimes d}\})$ is #P-hard. By Lemma IV.3, we have that $\#\mathrm{CSP}^d(\mathcal{F})$ is #P-hard. ■

Theorem IV.1 then follows directly from the following lemma.

Lemma IV.5. *Let $d \geq 1$ be an integer and $\mathcal{F}$ be a set of symmetric functions taking complex values. Then $\#\mathrm{CSP}^d(\mathcal{F} \cup \{[1, 0], [0, 1]\})$ is #P-hard unless there exist $T \in \mathcal{T}_{4d}$ such that $(T\mathcal{F}) \subset \mathscr{P}$ or $(T\mathcal{F}) \subset \mathscr{A}$, in which case the problem is in P.*

We give a sketch of proof here, some details are omitted due to space limitation. The full proof would appear in the full version of this paper.

Proof Sketch: We use the following bipartite Holant problem to express #CSP$^d(\mathcal{F} \cup \{[1,0],[0,1]\})$

$$\mathrm{Holant}(\{=_d, =_{2d}, \dots\}|\mathcal{F} \cup \{[1,0],[0,1]\}).$$

We first show the tractability part. Let $T \in \mathcal{T}_{4d}$ be the matrix such that $(T\mathcal{F}) \subset \mathscr{P}$ or $(T\mathcal{F}) \subset \mathscr{A}$. Applying a holographic reduction on the above problem under basis T^{-1}, we have

$$\begin{aligned} &\mathrm{Holant}(\{=_d, =_{2d}, \dots\}|\mathcal{F} \cup \{[1,0],[0,1]\}) \\ \equiv_T &\mathrm{Holant}(\{=_d, =_{2d}, \dots\}T^{-1}|(T\mathcal{F}) \cup \{[1,0],[0,1]\}). \end{aligned}$$

Since $\{=_d, =_{2d}, \dots\}T^{-1} \subset \mathscr{P} \cap \mathscr{A}$, we have that either all the signatures involved in the above Holant problem are in $\mathscr{P}$ or all the signatures involved in the above Holant problem are in $\mathscr{A}$. The polynomial time algorithm follows directly from that.

Now we prove the hardness part. We realize $\{[1,0],[0,1]\}$ on the LHS with the equality signatures on the LHS and the $[1,0]$'s and $[0,1]$'s on the RHS. We can then realize any sub-signature on the RHS. The overall idea now is to analyze the substructure of functions.

If all the binary sub-signatures of signatures in $\mathcal{F}$ are degenerate, then $\mathcal{F} \subset \mathscr{P}$ and we are done. Now we assume that we can realize a non-degenerate binary $[y_0, y_1, y_2]$ on the RHS.

Let $f := [f_0, f_1, \cdots, f_r]$ be a function in $\mathcal{F}$. If there exists some $i \in \{0, \cdots, r-1\}$ such that $f_i f_{i+1} \neq 0$, then we realize subsignature $[f_i, f_{i+1}]$. We then construct $[*,0,0,*]$ and transform it to $[1,0,0,1]$ under some diagonal matrix M and apply Theorem IV.2 to complete this case.

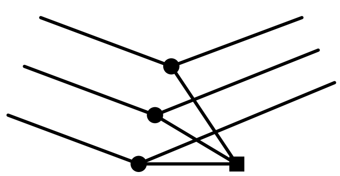

Figure 1. The circle vertices has signature $[0,1,0,0]$, and the square vertex is an equality function (we use $=_3$ as example here)

Now we assume that for every $f := [f_0, f_1, \cdots, f_r] \in \mathcal{F}$ and every $i \in \{0, \cdots, r-1\}$ we have $f_i f_{i+1} = 0$. Without loss of generality, we assume that we can construct a non-degenerate ternary signature f of form $[0, f_1, 0, f_3]$ where $f_1 \neq 0$, or $[0,1,0,a]$ after scale. If $a = 0$, we prove #P-hardness by using the construction in Figure 1 to reduce the #P-hard problem #CSP$([1,1,0])$ to it. If $a \neq 0$, then we can realize $[1,0,a]$ on the RHS. If d is odd, we connect $\frac{3d-3}{2}$ copies of $[1,0,a]$ to $=_{3d}$ to realize $[1,0,0,a^{\frac{3d-3}{2}}]$ on the LHS. We apply a suitable holographic transformation to make it into $[1,0,0,1]$ and apply Theorem IV.2. We can complete the proof similarly if d is even and there exists a non-degenerate signature of form $[*,0,0,0,\dots,0,*]$ with odd arity on the RHS.

Now we assume that d is even, and that all non-degenerate signatures of form $[*,0,\cdots,0,*]$ on the RHS have even arity. By grouping d copies of $[1,0,a,0]$, we can reduce the problem #CSP$([0,1,0,a^d])$ to #CSP$^d([0,1,0,a])$. Since $[0,1,0,a^d] \notin \mathscr{P}$, we have that #CSP$^d([0,1,0,a])$ is #P-hard unless $[0,1,0,a^d] \in \mathscr{A}$, which implies that $a^d = \pm 1$. We apply a holographic reduction under some suitable basis $T \in \mathcal{T}_{4d}$ to transform $[0,1,0,a]$ on the RHS to $[0,1,0,1]$. Note that $T \in \mathcal{T}_{4d}$ since $(a^{\frac{d}{2}})^{4d} = (\pm 1)^2 = 1$, so all the $=_{4kd}$ on the LHS will remain unchanged. We realize $[1,0,1]$ from $[0,1,0,1]$, and then realize all of $\{=_2, =_4, =_6 \dots, \}$ on the LHS with it. Since we have $=_2$ on both sides now, we reduce the following non-bipartite Holant problem to the original problem.

$$\mathrm{Holant}(\{=_2, =_4, =_6 \dots, \} \cup T\mathcal{F} \cup \{[1,0],[0,1],[0,1,0,1]\}).$$

So it is enough to show that this problem is #P-hard unless $(T\mathcal{F}) \subset \mathscr{A}$. To this end, we first show that if we can find in $T\mathcal{F}$ a signature of form $[*,0,0,0,\dots,0,*]$ of even arity but is not in $\mathscr{A}$, then the problem is #P-hard. We realize $[1,0,b]$ by connecting it to a suitable equality function. Similar to the above, we group signatures together to achieve reduction from general #CSP. The difference is that here we use one copy of $[1,0,b]$ and one copy of $[1,0,1]$ to form a group and realize a signature $[1,0,b]$ in the grouped instance, and reduce #CSP$([0,1,0,1],[1,0,b])$ to this and conclude that it is #P-hard if $b^4 \neq 1$. Using a same flavor of grouping, we prove a similar result for signatures of form $[*,0,*,0]$ and $[0,*,0,*]$ — they are #P-hard unless they are in $\mathscr{A}$, i.e., the ratio of the two nonzero elements is ± 1. Extending this result to general signatures, we conclude that either the whole signature is in $\mathscr{A}$ or we can construct a longer sub-signature that are multiples of $[1,0,1,0,-1]$ or $[1,0,-1,0,-1]$ after scale. For those two signatures, it is not hard to construct gadgets not in $\mathscr{A}$ and therefore they are #P-hard. ■

Next, we show how to apply Theorem IV.1 to prove dichotomy if we know that $\mathcal{F}$ contains some signature X of certain types.

In [30], we proved the following dichotomy for single ternary signature. Note that we omitted one additional tractable case here since it cannot happen for real signatures.

Theorem IV.6. *Let $[x_0, x_1, x_2, x_3]$ be a real non-degenerate signature. Then $\mathrm{Holant}([x_0, x_1, x_2, x_3])$ is #P-hard unless there exists a 2×2 matrix T such that $[x_0, x_1, x_2, x_3] = T^{\otimes 3}[1,0,0,1]$ and $[1,0,1]T^{\otimes 2}$ is in $\mathscr{A} \cup \mathscr{P}$.*

We now prove that the main dichotomy holds if $\mathcal{F}$ contains a non-degenerate ternary signature.

Theorem IV.7. *Let $\mathcal{F}$ be a set of real signatures, and X be a real non-degenerate ternary signature.* $\operatorname{Holant}(X, \mathcal{F})$ *is #P-hard unless $\mathcal{F} \cup \{X\}$ is good, for which case there is a polynomial-time algorithm.*

Proof: If $\operatorname{Holant}(X)$ is #P-hard according to Theorem IV.6, then we are done. Otherwise we take T as guaranteed in Theorem IV.6, and we have $\operatorname{Holant}(X, \mathcal{F}) \equiv_T \operatorname{Holant}([1,0,0,1], T^{-1}\mathcal{F} | [1,0,1]T^{\otimes 2})$ by applying holographic reduction. We also have that $[1,0,1]T^{\otimes 2}$ is non-degenerate. If $[1,0,1]T^{\otimes 2}$ is not of form $[0,\lambda,0]$, we are done by Theorem IV.2. Otherwise, we have $[0,1,0]$ on the RHS and can realize all the the equalities of arity $3k$ on the LHS. Then we can view it as a #CSP^3 problem and we are done by Theorem IV.1. ∎

For a signature with arity larger than 3, it is not necessarily true that we can transform it to $[1,0,0,0,1]$ by holographic reduction. But for some special signatures, we can. Formally, we have the following two corollaries.

Corollary IV.8. *Let X be a real non-degenerate generalized Fibonacci signature of arity no less than 3 and $\mathcal{F}$ be a set of symmetric signatures. Then* $\operatorname{Holant}(\mathcal{F}, X)$ *is #P-hard unless $\mathcal{F} \cup \{X\}$ is good, for which case there is a polynomial-time algorithm.*

Proof: For a real non-degenerate generalized Fibonacci signature, there is an orthogonal holographic reduction that transforms it into the form of $[1,0,0,\dots,0,a]$ where $a \neq 0$ after scale. By Theorem II.2, we assume that this is already the case with X. If the arity of X is odd, we realize $[1,0,0,a]$ by adding some self-loops and then apply Theorem IV.7. If the arity is even, we connect two copies of the signature using half of their dangling edges to realize $[1,0,0,\dots,0,a^2]$ with same arity. Keep doing this, and we can realize $[1,0,0,\dots,0,a^t]$ for any t. If a is a p-th root of unity, we get $[1,0,0,\dots,0,1]$ by choosing $t=p$. Otherwise, we get $[1,0,0,\cdots,0,1]$ by interpolation. Having $[1,0,0,\dots,0,1]$, we can realize all equality functions with even arity and the result follows from Theorem IV.1. ∎

Corollary IV.9. *Let $X = [x, y, -x, -y, x, y, -x, -y \dots]$ be a non-degenerate signature of arity $k \geq 3$, and $\mathcal{F}$ be a set of symmetric signatures. Then* $\operatorname{Holant}(\mathcal{F}, X)$ *is #P-hard unless $\mathcal{F} \cup \{X\}$ is good, for which case there is a polynomial-time algorithm.*

Proof: Rewrite $\operatorname{Holant}(\mathcal{F}, X)$ as $\operatorname{Holant}(\mathcal{F}, X | =_2)$. Applying holographic reduction under basis $Z = \begin{bmatrix} A & Bi \\ Ai & B \end{bmatrix}$ with suitable A and B, we can make X into $=_k$ where k is the arity of k and $=_2$ on the RHS into $[0,1,0]$. With $[0,1,0]$ on the RHS, we can realize all the equality function whose arity is a multiple of k on the LHS. Then we can apply Theorem IV.1 for #CSP^k. ∎

V. Realizing a signature by approximating it

In this section, we study $\operatorname{Holant}([1,a,b,-a,1] \cup \mathcal{F})$. We first show that we could always find an orthogonal transformation Q that converts $[1,a,b,-a,1]$ to $[1,0,b',0,1]$ for some b'.

Claim V.1. *There exists a real orthogonal 2×2 matrix Q, such that $[1,a,b,-a,1]Q^{\otimes 4} = [1,0,b',0,1]$ for some b'.*

This is proved by straightforward calculations. Details would appear in the full version of the paper.

Claim V.1 converts $\operatorname{Holant}([1,a,b,-a,1] \cup \mathcal{F})$ to $\operatorname{Holant}([1,0,b',0,1] \cup (\mathcal{F}Q))$. In the following, we simply assume that we are given $\operatorname{Holant}([1,0,b,0,1] \cup \mathcal{F})$. If $b \in \{0,1,-1\}$, we are done by Corollary IV.8 and Corollary IV.9. For $b \notin \{0,1,-1\}$, we will prove that $\operatorname{Holant}([1,0,b,0,1])$ is #P-hard, and these together give the following main lemma of this section.

Lemma V.2. *Let $X = [1,a,b,-a,1]$ be a non-degenerate signature. Then* $\operatorname{Holant}(\mathcal{F}, X)$ *is #P-hard unless $\mathcal{F} \cup \{X\}$ is good, for which case there is a polynomial-time algorithm.*

In the remaining of this section, we prove the hardness of $\operatorname{Holant}([1,0,b,0,1])$.

Lemma V.3. *If $b \notin \{0,1,-1\}$, then Holant$([1,0,b,0,1])$ is #P-hard.*

To prove this lemma, observe that if we can realize $[1,0,0,0,1]$, then we can use it to simulate $CSP^2([1,0,b,0,1])$, which is #P-hard by Theorem IV.1 and the fact $b \notin \{0,1,-1\}$. If we can realize $[1,0,1,0,1]$, then we can apply orthogonal transformation under $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ to convert $[1,0,1,0,1]$ to $[1,0,0,0,1]$ and $[1,0,b,0,1]$ to $[2+6b, 0, 2-2b, 0, 2+6b]$. To see that $[2+6b, 0, 2-2b, 0, 2+6b]$ is among the hard cases in Theorem IV.1, note that $|2+6b| \neq |2-2b|$ if $b \notin \{0,1,-1\}$, so it could not be transformed into $\mathscr{A}$ by $T \in \mathcal{T}_{4d}$. It is also not hard to verify that $T \notin \mathscr{P}$.

In the following, we introduce two new techniques for realizing special signatures $[1,0,0,0,1]$ or $[1,0,1,0,1]$. First, we generalize the widely-used interpolation technique to enable us to interpolate 4-ary signatures instead of unary signatures in the traditional setting. This generalization is already powerful enough for almost all b. The failed b are roots of some integer coefficient polynomials and thus must be algebraic numbers. So we have that $\operatorname{Holant}([1,0,b,0,1])$ is #P-hard if b is a transcendental real number. The idea of the proof is similar to the interpolation of unary signatures. For the cases when b is an algebraic real number, we use our second new technique to realize a signature by approximating it. Here is the formal statement.

Theorem V.4. *Let $f = [x_0, \dots, x_k]$ be a symmetric Boolean signature of arity k and $\{g_m\}$ be a sequence of signatures*

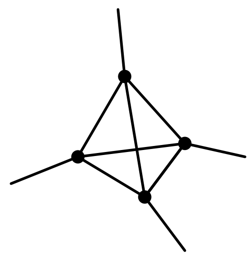

Figure 2. The tetrahedron gadget.

of arity k. We assume that all the signature values are real algebraic numbers, and there exists a constant $C > 1$ such that for all m, we have $|f - g_m|_\infty < C^{-m}$. If we can compute Holant(g_m) in time poly(n, m), where n is the input size, then we can compute Holant(f) in polynomial time.

Proof Sketch: Holant value of any instance of Holant(f) can be written as an integer combination of a fixed number of algebraic numbers. We call the set of these integer combinations S. Using a property of algebraic numbers, we prove that the difference of any two distinct elements in S is at least B^{-n^2}, where $B > 1$ is an absolute constant depending only on f. Now consider replacing f with g_m in a Holant instance. The difference in the final Holant value is at most $D^{n^2}C^{-m}$, where D is an absolute constant only depending on f. Therefore, we can choose $m = En^2$ with sufficient large constant E only depending on D and B (thus f) such that $D^{n^2}C^{-m} < B^{-n^2}$. This Holant(g_m) can be computed in time polynomial of n and m which is a polynomial in n. Using the above property, we know that the true Holant value Holant(f), which is an element in S, is $D^{n^2}C^{-m}$ close to the value returned by Holant(g_m) and in that neighborhood there is no other point from S. So we can use integer programming to recover all the integer coefficients and find out this unique value. Since the number of coefficients (variables for the programming) are constant, we can use the integer programming algorithm from [43], [44] which runs in polynomial time. ■

In the following, we use the above reduction to study the complexity of Holant($[1, 0, b, 0, 1]$). For $[1, 0, b, 0, 1]$, using the Tetrahedron gadget in Figure 2, we realize a new symmetric signature: $[(b + 1)^2(3b^2 - 2b + 1), 0, 2b^2(b + 1)^2, 0, (b + 1)^2(3b^2 - 2b + 1)]$. Assume that $b \neq -1$. Keep doing this recursive construction, we can realize a signature $[1, 0, b_r, 0, 1]$ with $b_r = \frac{2b_{r-1}^2}{3b_{r-1}^2 - 2b_{r-1} + 1}$. The following lemma shows that this recursive construction converges to a fixed point very fast.

Lemma V.5. *Let b be a real algebraic number, $b \neq 0, b \neq \pm 1, b \neq \frac{1}{3}$. Let $[1, 0, b_r, 0, 1]$ be the signature realized by the r-th recursive Tetrahedron gadget starting from $[1, 0, b, 0, 1]$. Let $\beta = 0$ if $b < \frac{1}{3}$, and $\beta = 1$ otherwise. Then $|b_r - \beta| < C^{-2^r}$, where $C < 1$ is some constant. In other words, the recursive construction either converges to $[1, 0, 0, 0, 1]$ or $[1, 0, 1, 0, 1]$, depending on whether b is smaller than $\frac{1}{3}$ or not.*

The r-th gadget contains 4^r nodes. We do the recursive gadget $O(\log n)$ levels so it is still of polynomial size. We note that this is the reason why we cannot use the tetrahedron gadget to interpolate since we would need polynomial many levels. The speed of convergence in Lemma V.5 is so fast that we can approximate the target gadget to within $C^{-poly(n)}$ by a gadget with $O(\log n)$ levels of recursive construction. Then by Theorem V.4 and the above analysis, we get the hardness for Holant($[1, 0, b, 0, 1]$) when $b \notin \{0, 1, -1, \frac{1}{3}\}$.

To complete the proof for Lemma V.3, the only remaining case is $[1, 0, \frac{1}{3}, 0, 1]$. By applying holographic transformation $\begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$, we have that

$$\text{Holant}\left(\left[1, 0, \frac{1}{3}, 0, 1\right]\right) \equiv_T \text{Holant}([0, 0, 1, 0, 0]|[0, 1, 0]).$$

We now show that the right hand side is #P-hard. To this end, we introduce the Counting-Eulerian-Orientation problem. First we define Eulerian orientations.

Definition V.6. *Given a graph $G = (V, E)$ of which all vertices are of even degree. Let σ be an orientation of its edges E. σ is an Eulerian orientation iff for each vertex $v \in V$, the number of incoming edges and outgoing edges of v are the same.*

To prove hardness of counting Eulerian orientations, we show how to use it to calculate a certain hard-to-compute weighted sum of orientations on the medial graph of planar graphs. We recall the definition of medial graphs.

Definition V.7 ([45]). *Let G be a connected planar graph. For simplicity, we assume that every edge of G is contained in exactly two different planes. Define its medial graph $H = (V_H, E_H)$, where V_H consists of the middle points of edges in G, and for each plane in G, connect the middle points on the border of G to get a cycle, and E_H consists of all edges on this cycle.*

Note that medial graphs are 4-regular graphs.

The following theorem shows the relation between Eulerian orientations, medial graphs and Tutte polynomials. For the definition of Tutte polynomial, we refer to [46].

Theorem V.8 ([45]). *Let G be a connected planar graph and let $\mathscr{O}(H)$ be the set of all Eulerian orientations of the medial graph $H = H(G)$. Then*

$$\sum_{O \in \mathscr{O}(H)} 2^{\beta(O)} = 2 \cdot T(G; 3, 3), \tag{1}$$

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

It is known that calculating the right hand side of the above is #P-hard.

Theorem V.9 ([47], [48]). *If we have that* $(x, y) \in \{(1, 1), (-1, -1), (0, 1), (-1, 0)\}$ *or satisfies* $(x-1)(y-1) = 1$*, the Tutte polynomial is computable in polynomial time. Otherwise, it is #P-hard. If the problem is restricted to the class of planar graphs, the points on the hyperbola defined by* $(x-1)(y-1) = 2$ *become polynomial-time computable, but all other points remain #P-hard.*

Before we prove the main theorem of this section, first observe that Holant$([0, 0, 1, 0, 0]|[0, 1, 0])$ is exactly the number of Eulerian orientations in a 4-regular graph.

Now we prove the main theorem. We show how to calculate the LHS in Theorem V.8 given an oracle of COUNTING-EULERIAN-ORIENTATION.

Theorem V.10. COUNTING-EULERIAN-ORIENTATION *is #P-hard for 4-regular graphs.*

Proof Sketch: We reduce calculating the LHS of Equation (1) to Holant$([0, 0, 1, 0, 0]|[0, 1, 0])$. Then since it is known that calculating the Tutte polynomial on graphs at $(3, 3)$ is #P-hard, we conclude that Holant$([0, 0, 1, 0, 0]|[0, 1, 0])$ is #P-hard.

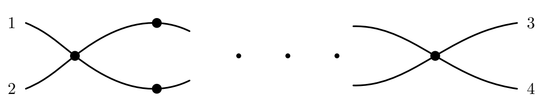

Figure 3. Recursive gadget. 4-ary signatures are $[0, 0, 1, 0, 0]$, and binary ones are $[0, 1, 0]$.

By polynomial interpolation, we can realize the following signature

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

We complete the proof by observing that Holant$_{G_H}(G_x|[0, 1, 0]) = \sum_{O \in \mathscr{O}(H)} 2^{\beta(O)}$, for a suitably constructed bipartite graph G_H. ■

VI. DICHOTOMY FOR REAL HOLANT

In this section we prove our main result. The idea of the proof is to use induction on the arity of the functions. We apply dichotomy theorems for functions with smaller arity for the induction step. The base step would be dichotomy theorems for functions of arity three and four. The ternary case is proved in Theorem IV.7 in Section IV. In this section, we go on to analyze complexity of signatures of arity four. We start with the following lemma in which we have an additional unary signature.

Lemma VI.1. *Let* X *be a non-degenerate real 4-ary signature and* a, b *be two real number which are not both zero. Then Holant*$(\mathcal{F}, X, [a, b])$ *is #P-hard unless* $\mathcal{F} \cup \{X, [a, b]\}$ *is good.*

Proof: Since a, b are not both zero, we could always apply a real orthogonal transformation Q, so that $Q[a, b] = [1, 0]$. Let $Y = Q^{\otimes 4}X = [y_0, y_1, y_2, y_3, y_4]$. Note that Y is still a real signature. Since it always has the same value as H$olant(\mathcal{F}, X, [a, b])$, it is equivalent to consider Holant$(Q\mathcal{F}, Y, [1, 0])$. Since we have $[1, 0]$ in this transformed instance, we could realize $Y' = [y_0, y_1, y_2, y_3]$. If Y' is non-degenerate, we apply Theorem IV.7. Now consider the case that Y' is degenerate. If Y' is an all zero signature, then Y is degenerate, which means that X is degenerate, contradicts to our hypothesis. If $Y' = [1, 0]^{\otimes 3}$, then $Y = [1, 0, 0, 0, *]$ is a non-degenerate generalized Fibonacci signature and we apply Corollary IV.8. If $Y' = [0, 1]^{\otimes 3}$, by adding a self-loop, we can realize $[0, 1]$. Since we have both $[1, 0]$ and $[0, 1]$, we apply the dichotomy theorem for Holant^c. Otherwise, assume that $Y' = [1, t]^{\otimes 3}$, where $t \in \mathbb{R} \setminus \{0\}$, and $Y = [1, t, t^2, t^3, y]$, where $y \neq t^4$. Connecting three copies of $[1, 0]$ to Y, we can realize $[1, t]$. Connecting one copy of $[1, t]$ to Y, we have $Y'' = [1+t^2, t+t^3, t^2+t^4, t^3+yt]$. This is a non-degenerate ternary function for any real t and $y \neq t^4$. We now apply Theorem IV.7 to finish the proof. ■

By a similar argument as in Lemma IV.4, we can replace $[a, b]$ with $[a, b]^{\otimes 2}$.

Lemma VI.2. *Let* X *be a non-degenerate real 4-ary signature and* a, b *be two real number which are not both zero. Then Holant*$(\mathcal{F}, X, [a, b]^{\otimes 2})$ *is #P-hard unless* $\mathcal{F} \cup \{X, [a, b]^{\otimes 2}\}$ *is good.*

We now prove a theorem for Holant problems when we have a non-degenerate 4-ary function.

Theorem VI.3. *Let* X *be a non-degenerate 4-ary signature, and* $\mathcal{F}$ *be a set of signatures. Then Holant*$(X, \mathcal{F})$ *is #P-hard unless* $\mathcal{F} \cup \{X\}$ *is good, for which there is a polynomial-time algorithm.*

Proof: As usual, the tractability part follows from algorithms for #CSP. We prove the hardness part. The main idea is to realize a degenerate binary function and make use of Lemma VI.2.

By adding a self-loop to X, we have $X' = [x_0 + x_2, x_1 + x_3, x_2 + x_4]$. If X' is all zero, then we have $X = [x_0, x_1, -x_0, -x_1, x_0]$ and we apply Corollary IV.9. If $X' = [x_0 + x_2, x_1 + x_3, x_2 + x_4]$ is degenerate and not all zero, then we apply Lemma VI.2 directly.

Now we assume that X' is non-degenerate. We make a polynomial interpolation by a chain of X's. The eigenvalues of $X' = \begin{bmatrix} x_0 + x_2 & x_1 + x_3 \\ x_1 + x_3 & x_2 + x_4 \end{bmatrix}$ are $\lambda_{1,2} = \frac{(x_0+2x_2+x_4) \pm \sqrt{\Delta}}{2}$ where $\Delta = (x_4 - x_0)^2 + 4(x_1 + x_3)^2$. By a chain of X's, we can realize $P \begin{bmatrix} \lambda_1^k & 0 \\ 0 & \lambda_2^k \end{bmatrix} P^{-1}$, where P is the basis formed by its eigenvectors. We already know that $\lambda_1\lambda_2 \neq 0$ since X' is non-degenerate. If we further have that the ratio $\frac{\lambda_1}{\lambda_2}$ is not a root of unity, we can interpolate all the binary signatures expressible as $P\begin{bmatrix}x & 0\\ 0 & y\end{bmatrix}P^{-1}$. In particular, we can interpolate $P\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}P^{-1}$, which is a degenerate non-zero binary signature. We are done by Lemma VI.2.

The exceptional case is that the ratio $\frac{\lambda_1}{\lambda_2}$ is a root of unity. Since X' is a real symmetric function, both λ_1 and λ_2 are real. So the only possible roots of unity are ± 1. We have that $\lambda_1 = \lambda_2$ iff $\Delta = 0$ iff $x_4 = x_0$ and $x_1 = -x_3$. Also, $\lambda_1 = -\lambda_2$ iff $(x_0 + x_2) = -(x_2 + x_4)$. We deal with these exceptional cases separately as follows.

Case 1: $x_4 = x_0 \neq 0$ and $x_1 = -x_3$. This is of form $[1, a, b, -a, 1]$, and we apply Lemma V.2.

Case 2: $x_4 = x_0 = 0$ and $x_1 = -x_3$. If we further have $x_2 = 0$, then this is a signature of form $[x, y, -x, -y, x]$ and we apply Corollary IV.9. Otherwise, it is of form $[0, a, 1, -a, 0]$. By the Tetrahedron gadget, we can realize a signature of $[6a^2 + 3, a, 2a^2 + 2, -a, 6a^2 + 3]$. Since $6a^2 + 3 \neq 0$, this case is proved in Case 1.

Case 3: $(x_0+x_2) = -(x_2+x_4)$. Using Tetrahedron gadget, we can realize signature $[y_0, y_1, y_2, y_3, y_4]$ such that $(y_0 + y_2) \neq -(y_2+y_4)$. We can verify that the problem is reduced to one of the cases proved above. ∎

Now we are ready to prove our main result.

Proof of Theorem III.2: As stated in the outline in Section III, we prove this theorem by showing that for any non-degenerate signature X with arity at least 3, $\mathrm{Holant}(X, \mathcal{F})$ is tractable iff there exists a 2×2 matrix satisfying the conditions. We prove by induction on the arity k of X.

The cases of $k = 3$ and $k = 4$ are proved in Theorem IV.7 and Theorem VI.3.

Suppose for arity $k < n$, we have proved our claim. Now we have a signature X of arity n. We obtain an $(n-2)$-ary signature X' by adding a self-loop to X. If X' is non-degenerate, then we are done by induction hypothesis. If X' is all zero, then X is of form $[x, y, -x, -y, x, y, -x, -y \ldots]$ and we apply Corollary IV.9. The only remaining case is that X' is degenerate but not all zero, and without loss of generality, we assume that $X' = [a, b]^{\otimes(n-2)}$. By applying an appropriate *real* orthogonal transformation, we could transform X' into $[1, 0]^{\otimes(n-2)}$, and X into $Y = XQ^{\otimes n} \triangleq [y_0, y_1, \ldots, y_n]$. By Theorem II.2, we may just assume that we actually have Y in the place of X. The fact that X' is transformed into $Y' = [1, 0]^{\otimes(n-2)}$ implies that $Y = [y_0, y_1, y_2, -y_1, -y_2, \ldots]$. After adding enough self-loops to Y' we can either get $[1, 0]$ or $[1, 0, 0]$ depending on the parity of n. Then connecting some copies of $[1, 0]$ or $[1, 0, 0]$ to Y, we can either get $Y'' = [y_0, y_1, y_2, -y_1]$ or $Y'' = [y_0, y_1, y_2, -y_1, -y_2]$. If Y'' is degenerate, then the ratio must be $\pm i$, and $y_0 = -y_2$. This would imply that Y' is an all zero signature, a contradiction. Now we know that such a Y'' is not degenerate, and we can complete the proof by Theorem IV.7 or Theorem VI.3. ∎

ACKNOWLEDGEMENT

The first author is supported by ERC Advanced investigator grant 226203. Work partly done when he was an intern student at Microsoft Research Asia and an undergraduate student at Shanghai Jiao Tong University.

REFERENCES

[1] A. A. Bulatov, “A dichotomy theorem for constraint satisfaction problems on a 3-element set,” *J. ACM*, vol. 53, no. 1, pp. 66–120, 2006.

[2] ——, “The Complexity of the Counting Constraint Satisfaction Problem,” in *ICALP (1)*, 2008, pp. 646–661.

[3] A. Bulatov, M. Dyer, L. A. Goldberg, M. Jalsenius, and D. Richerby, “The complexity of weighted boolean #csp with mixed signs,” *Theor. Comput. Sci.*, vol. 410, pp. 3949–3961, September 2009.

[4] A. A. Bulatov and V. Dalmau, “Towards a dichotomy theorem for the counting constraint satisfaction problem,” *Inf. Comput.*, vol. 205, no. 5, pp. 651–678, 2007.

[5] M. E. Dyer, L. A. Goldberg, and M. Jerrum, “The Complexity of Weighted Boolean #CSP,” *SIAM J. Comput.*, vol. 38, no. 5, pp. 1970–1986, 2009.

[6] T. Feder and M. Y. Vardi, “The Computational Structure of Monotone Monadic SNP and Constraint Satisfaction: A Study through Datalog and Group Theory,” *SIAM J. Comput.*, vol. 28, no. 1, pp. 57–104, 1998.

[7] M. E. Dyer and D. Richerby, “On the complexity of #CSP,” in *STOC*, 2010, pp. 725–734.

[8] ——, “The #CSP Dichotomy is Decidable,” in *STACS*, 2011, pp. 261–272.

[9] J.-Y. Cai, X. Chen, and P. Lu, “Non-negatively weighted #CSP: An effective complexity dichotomy,” in *IEEE Conference on Computational Complexity*. IEEE Computer Society, 2011, pp. 45–54.

[10] A. A. Bulatov and M. Grohe, “The Complexity of Partition Functions,” in *ICALP*, 2004, pp. 294–306.

[11] ——, “The complexity of partition functions,” *Theor. Comput. Sci.*, vol. 348, no. 2-3, pp. 148–186, 2005.

[12] J.-Y. Cai, X. Chen, and P. Lu, “Graph Homomorphisms with Complex Values: A Dichotomy Theorem,” in *ICALP (1)*, 2010, pp. 275–286.

[13] M. E. Dyer, L. A. Goldberg, and M. Paterson, “On counting homomorphisms to directed acyclic graphs,” *J. ACM*, vol. 54, no. 6, 2007.

[14] M. E. Dyer and C. S. Greenhill, “The complexity of counting graph homomorphisms (extended abstract),” in *SODA*, 2000, pp. 246–255.

[15] ——, “Corrigendum: The complexity of counting graph homomorphisms,” *Random Struct. Algorithms*, vol. 25, no. 3, pp. 346–352, 2004.

[16] L. A. Goldberg, M. Grohe, M. Jerrum, and M. Thurley, “A complexity dichotomy for partition functions with mixed signs,” *CoRR*, vol. abs/0804.1932, 2008.

[17] P. Hell and J. Nesetril, “On the complexity of -coloring,” *J. Comb. Theory, Ser. B*, vol. 48, no. 1, pp. 92–110, 1990.

[18] L. G. Valiant, “Holographic Algorithms,” *SIAM J. Comput.*, vol. 37, no. 5, pp. 1565–1594, 2008.

[19] ——, “Accidental Algorthims,” in *FOCS*, 2006, pp. 509–517.

[20] J.-Y. Cai, P. Lu, and M. Xia, “Holographic Algorithms by Fibonacci Gates and Holographic Reductions for Hardness,” in *FOCS*, 2008, pp. 644–653.

[21] ——, “Holant problems and counting CSP,” in *STOC ’09: Proceedings of the 41st annual ACM symposium on Theory of computing*. New York, NY, USA: ACM, 2009, pp. 715–724.

[22] N. Creignou, S. Khanna, and M. Sudan, *Complexity classifications of boolean constraint satisfaction problems*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2001.

[23] R. E. Ladner, “On the Structure of Polynomial Time Reducibility,” *J. ACM*, vol. 22, no. 1, pp. 155–171, 1975.

[24] L. G. Valiant, “Some observations on holographic algorithms,” in *LATIN*, 2010, pp. 577–590.

[25] ——, “The Complexity of Computing the Permanent,” *Theor. Comput. Sci.*, vol. 8, pp. 189–201, 1979.

[26] C. H. Papadimitriou and S. Zachos, “Two remarks on the power of counting,” in *Proceedings of the 6th GI-Conference on Theoretical Computer Science*, 1982, pp. 269–276.

[27] J. Faben, “The complexity of counting solutions to generalised satisfiability problems modulo k,” *CoRR*, vol. abs/0809.1836, 2008.

[28] H. Guo, S. Huang, P. Lu, and M. Xia, “The complexity of weighted boolean #csp modulo k,” in *STACS*, 2011, pp. 249–260.

[29] E. Ising, “Beitrag zur theorie des ferromagnetismus,” *Zeitschrift für Physik A Hadrons and Nuclei*, vol. 31, no. 1, pp. 253–258, 1925.

[30] J.-Y. Cai, S. Huang, and P. Lu, “From Holant to #CSP and Back: Dichotomy for Holant^c Problems,” in *ISAAC (1)*, 2010, pp. 253–265.

[31] H. Guo, P. Lu, and L. G. Valiant, “The complexity of symmetric boolean parity holant problems - (extended abstract),” in *ICALP (1)*, ser. Lecture Notes in Computer Science, L. Aceto, M. Henzinger, and J. Sgall, Eds., vol. 6755. Springer, 2011, pp. 712–723.

[32] J.-Y. Cai, P. Lu, and M. Xia, “A Computational Proof of Complexity of Some Restricted Counting Problems,” in *TAMC*, 2009, pp. 138–149.

[33] M. Kowalczyk and J.-Y. Cai, “Holant Problems for Regular Graphs with Complex Edge Functions,” in *STACS*, 2010, pp. 525–536.

[34] J.-Y. Cai and M. Kowalczyk, “A Dichotomy for k-Regular Graphs with $\{0, 1\}$-Vertex Assignments and Real Edge Functions,” in *TAMC*, ser. Lecture Notes in Computer Science, J. Kratochvíl, A. Li, J. Fiala, and P. Kolman, Eds., vol. 6108. Springer, 2010, pp. 328–339.

[35] M. Mihail and P. Winkler, “On the Number of Eulerian Orientations of a Graph,” *Algorithmica*, vol. 16, no. 4/5, pp. 402–414, 1996.

[36] G. Brightwell and P. Winkler, “Counting Eulerian Circuits is #P-Complete,” in *ALENEX/ANALCO*, 2005, pp. 259–262.

[37] Q. Ge and D. Stefankovic, “The Complexity of Counting Eulerian Tours in 4-Regular Graphs,” in *LATIN*, 2010, pp. 638–649.

[38] C. T. J. Dodson and T. Poston, *Tensor Geometry*. New York: Springer-Verlag, 1991.

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

[40] L. Blum, F. Cucker, M. Shub, and S. Smale, *Complexity and real computation*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1998.

[41] J.-Y. Cai, P. Lu, and M. Xia, “Holographic Algorithms with Matchgates Capture Precisely Tractable Planar #CSP,” in *FOCS*, 2010, pp. 427–436.

[42] ——, “Holant problems and counting CSP,” in *STOC*, 2009, pp. 715–724.

[43] H. W. J. Lenstra, “Integer Programming with a Fixed Number of Variables,” *Mathematics of Operations Research*, vol. 8, no. 4, November 1983.

[44] V. E. Brimkov and S. S. Dantchev, “On the Complexity of Integer Programming in the Blum-Shub-Smale Computational Model,” in *TCS ’00: Proceedings of the International Conference IFIP on Theoretical Computer Science, Exploring New Frontiers of Theoretical Informatics*. London, UK: Springer-Verlag, 2000, pp. 286–300.

[45] M. L. Vergnas, “On the evaluation at (3, 3) of the Tutte polynomial of a graph,” *J. Combin. Theory Ser. B*, vol. 45, no. 3, pp. 367–372, 1988.

[46] B. Bollobas, *Modern Graph Theory*, ser. Graduate Texts in Mathematics. New York: Springer, 1998, vol. 184.

[47] F. Jaeger, D. L. Vertigan, and D. J. A. Welsh, “On the computational complexity of the Jones and Tutte polynomials,” *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 108, pp. 35–53, 1990.

[48] D. Vertigan, “The Computational Complexity of Tutte Invariants for Planar Graphs,” *SIAM J. Comput.*, vol. 35, no. 3, pp. 690–712, 2005.