A Holant Dichotomy: Is the FKT Algorithm Universal?

Jin-Yi Cai*, Zhiguo Fu†, Heng Guo*, and Tyson Williams*
* *Computer Sciences Department, University of Wisconsin–Madison, Madison, WI, USA*
{jyc,hguo,tdw}@cs.wisc.edu
† *School of Mathematics, Jilin University, Changchun, Jilin, China*
fuzg@jlu.edu.cn

Abstract

We prove a complexity dichotomy for complex-weighted Holant problems with an arbitrary set of symmetric constraint functions on Boolean variables.

In the study of counting complexity, such as #CSP, there are problems which are #P-hard over general graphs but P-time solvable over planar graphs. A recurring theme has been that a holographic reduction [36] to FKT precisely captures these problems. This dichotomy answers the question: Is this a *universal* strategy? Surprisingly, we discover new planar tractable problems in the Holant framework (which generalizes #CSP) that are not expressible by a holographic reduction to FKT. In particular, the putative form of a dichotomy for planar Holant problems is false. Nevertheless, we prove a dichotomy for $\#\text{CSP}^2$, a variant of #CSP where every variable appears even times, that the presumed *universality* holds for $\#\text{CSP}^2$. This becomes an important tool in the proof of the full dichotomy, which refutes this universality in general. The full dichotomy says that the new P-time algorithms and the strategy of holographic reductions to FKT together are *universal* for these locally defined counting problems.

As a special case of our new planar tractable problems, counting perfect matchings (#PM) over k-uniform hypergraphs is P-time computable when the incidence graph is planar and $k \geq 5$. The same problem is #P-hard when $k = 3$ or $k = 4$, also a consequence of the dichotomy. More generally, over hypergraphs with specified hyperedge sizes and the same planarity assumption, #PM is P-time computable if the greatest common divisor (gcd) of all hyperedge sizes is at least 5.

Keywords

Computational Complexity; Counting Problems; Dichotomy Theorem; Holographic Algorithms; Holant Problems;

I. Introduction

The Fisher-Kasteleyn-Temperley (FKT) algorithm [32], [21], [22] is a classical gem that counts perfect matchings over planar graphs in polynomial time. This was an important milestone in a decades-long research program by physicists in statistical mechanics to determine what is known as Exactly Solved Models [1], [20], [30], [41], [42], [25], [32], [21], [22], [26], [27], [40].

For four decades, the FKT algorithm stood as *the* polynomial-time algorithm for any counting problem over planar graphs that is #P-hard over general graphs. Then Valiant introduced *matchgates* [34], [33] and *holographic* reductions to the FKT algorithm [36], [35]. These reductions differ from classical ones by introducing quantum-like superpositions. This novel technique extended the reach of the FKT algorithm and produced polynomial-time algorithms for a number of problems for which only exponential-time algorithms were previously known.

Since the new polynomial-time algorithms appear so exotic and unexpected, and the problems appear so close to being #P-hard, they challenge our faith in the well-accepted conjecture that P $\neq$ NP. Quoting Valiant [35]: “The objects enumerated are sets of polynomial systems such that the solvability of any one member would give a polynomial time algorithm for a specific problem. . . . the situation with the P

0272-5428/15 $31.00 © 2015 IEEE
DOI 10.1109/FOCS.2015.81

= NP question is not dissimilar to that of other unresolved enumerative conjectures in mathematics. The possibility that accidental or freak objects in the enumeration exist cannot be discounted if the objects in the enumeration have not been studied systematically." Indeed, if any "freak" object exists in this framework, it would collapse #P to P.

Therefore, over the past 10 to 15 years, this technique has been intensely studied in order to gain a systematic understanding to the limit of the trio of holographic reductions, matchgates, and the FKT algorithm [33], [3], [4], [10], [37], [11], [24], [28], [29]. Without settling the P versus #P question, the best hope is to achieve a complexity classification. This program finds its sharpest expression in a complexity dichotomy theorem, which classifies *every* problem expressible in a framework as either solvable in P or #P-hard, with nothing in between.

Out of this work, a strong theme has emerged. For a wide variety of problems, such as those expressible as a #CSP, holographic reductions to the FKT algorithm is a *universal* technique for turning problems that are #P-hard in general to P-time solvable over planar graphs. In fact, a preponderance of evidence suggests the following putative classification of all counting problems defined by local constraints into *exactly* three categories: (1) those that are P-time solvable over general graphs; (2) those that are P-time solvable over planar graphs but #P-hard over general graphs; and (3) those that remain #P-hard over planar graphs. Moreover, category (2) consists precisely of those problems that are holographically reducible to the FKT algorithm. This theme is so strong that it has become an intuitive and trusty guide for us when we investigate unknown problems and plan proof strategies. In fact, many of the results in the present paper were proved in this way. However, one is still left wondering whether a holographic reduction to the FKT algorithm is a *universal* strategy for all such counting problems that are planar tractable but not in general.

We list some of the supporting evidence for this putative classification. These date back to the classification of the complexity of the Tutte polynomial [39], [38]. It has also been an unfailing theme in the classification of spin systems and #CSP [23], [12], [9], [18]. However, these frameworks do not capture all locally specified counting problems. Some natural problems, such as counting perfect matchings (#PM), are not expressible as a point on the Tutte polynomial, and #PM is provably not expressible as a partition function of spin systems (vertex assignment models) [16], [15], [31]. However, this is the problem for which FKT was designed, and is the basis of Valiant's matchgates and holographic reductions.

A refined framework, called Holant problems [13], was proposed to address this issue. It is an edge assignment model. It naturally encodes and expresses #PM as well as Valiant's matchgates and holographic reductions. Thus, Holant is the proper framework in which to study the power of holographic algorithms. It is also more general than #CSP in the sense that a complete complexity classification for Holant problems implies one for #CSP.

In this paper, we classify for the first time the complexity of Holant problems over planar graphs. Our result generalizes both the dichotomy for Holant [19], [6] and the dichotomy for planar #CSP [12], [18]. Although the #CSP dichotomy does not resolve the complexity of #PM, planar tractable classes of #CSP are tractable due to holographic algorithms with matchgates, which essentially relies on counting (weighted) perfect matchings by FKT. On the other hand, #PM, even for d-regular graphs, is shown to be #P-hard under the Holant framework [19], yet its planar tractability is not addressed in either [19] or [6] until the current work.

Surprisingly, we discover new planar tractable problems that are not expressible by a holographic reduction to matchgates and FKT. To the best of our knowledge, this is the first primitive extension since FKT to a counting problem solvable in P over planar instances but #P-hard in general. We consider this a primitive extension because it is provably not based on a (holographic) transformation to the

FKT algorithm. Furthermore, our dichotomy theorem says that this completes the picture: there are no more undiscovered extensions for problems expressible in this framework, unless #P collapses to P. In particular, the putative form of the planar Holant dichotomy is *false*.

Before stating our main theorem, we give a brief description of the Holant framework [13]. Fix a set of local constraint functions $\mathcal{F}$. A *signature grid* $\Omega = (G, \pi)$ is a tuple, where $G = (V, E)$ is a graph, π labels each $v \in V$ with a function $f_v \in \mathcal{F}$ with input variables from the incident edges $E(v)$ at v. Each f_v maps $\{0,1\}^{\deg(v)}$ to $\mathbb{C}$. We consider all 0-1 edge assignments. 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 restriction of σ to $E(v)$. The counting problem on the instance Ω is to compute

$$\text{Holant}_\Omega = \sum_{\sigma: E \to \{0,1\}} \prod_{v \in V} f_v \left(\sigma \mid_{E(v)}\right). \tag{I.1}$$

For example, #PM, the problem of counting perfect matchings in G, corresponds to assigning the EXACTONE function at every vertex of G. The Holant problem parameterized by the set $\mathcal{F}$ is denoted by $\text{Holant}(\mathcal{F})$.

At a high level, we can state our main theorem as follows.

Theorem I.1. *Let $\mathcal{F}$ be a set of complex-valued, symmetric functions on Boolean variables. Then there is an effective classification for all possible $\mathcal{F}$, according to which,* $\text{Holant}(\mathcal{F})$ *is either (1) P-time computable over general graphs, or (2) P-time computable over planar graphs but #P-hard over general graphs, or (3) #P-hard over planar graphs.*

Note that here we restrict our focus on symmetric functions, which are invariant under permutation of arguments. Most natural combinatorial problems like counting vertex covers or perfect matchings can be encoded by symmetric functions.

The complete statement is given in Theorem III.1. The classification is explicit. The tractability criterion is decidable in polynomial time due to [11], [7]. Tractable problems over general graphs have been previously studied in [6]. The planar tractable class includes both those solvable by holographic reductions to FKT and those newly discovered. Explicit criteria for these are also proved in this paper.

Let us meet some new tractable problems. They can be described as orientation problems, which are Holant problems after a complex-valued holographic transformation. Given a planar (multi)graph, we allow two kinds of vertices. The first kind can be either a sink or a source while the second kind only allow one incoming edge. The goal is to compute the number orientations satisfying these constraints. This problem can be expressed in the Holant framework under a Z-transformation.[1] It can be shown that this is equivalent to the Holant problem on the edge-vertex incidence graph where we assign the DISEQUALITY function to every edge, and to each vertex, we assign either the EQUALITY function or the EXACTONE function. Suppose vertices assigned EQUALITY functions all have degree k. If $k = 2$, then this problem can be solved by FKT. We show that this problem is #P-hard if $k = 3$ or $k = 4$, but is tractable again if $k \geq 5$. The algorithm involves a recursive procedure that simplifies the instance until it can be solved by known algorithms, including FKT. This simplilfication process pins edges to fixed values, yet the final answer is still possible to be non-trivial as the pinning will end when the instance is solvable by known algorithms. The algorithm crucially uses global topological properties of a planar graph, in particular Euler's characteristic formula. If the graph is not planar, then this algorithm does not work, and indeed the problem is #P-hard over general graphs.

[1]This transformation is $Z = \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$. It is common that one problem can be transformed to another over $\mathbb{C}$ while one or both problems are specified by real-valued constraint functions, and provably no transformation exists over $\mathbb{R}$. Thus to study the classification question over complex-valued constraint functions is natural and proper. For example, the integer-valued orientation problem studied here, if expressed as Holant directly, is complex weighted.

More generally, we allow vertices of arbitrary degrees to be assigned EQUALITY. If all the degrees are at most 2, then the problem is tractable by the FKT algorithm. Otherwise, the complexity depends on the greatest common divisor (gcd) of the degrees. The problem is tractable if $\gcd \geq 5$ and #P-hard if $\gcd \leq 4$. It is worth noting that the criterion for tractability is not a degree lower bound. Moreover, the planarity assumption and the degree rigidity pose a formidable challenge in the hardness proofs for $\gcd \leq 4$. We note that these degree restrictions and planarity will not make new tractable instances vacuous, since there are two types of vertices and we do not lower bound degrees of those assigned EXACTONE. In addition, as common in the study of #CSP, we allow multigraphs as valid instances.

If the graph is bipartite with EQUALITY functions assigned on one side and EXACTONE functions on the other, then this is the problem of #PM over hypergraphs with planar incidence graphs. Our results imply that the complexity of this problem depends on the gcd of the hyperedge sizes. The problem is computable in polynomial time when $\gcd \geq 5$ and is #P-hard when $\gcd \leq 4$ (assuming there are hyperedges of size at least 3).

Most reductions in previous Holant dichotomy theorems [19], [6] do not hold for planar graphs, so we are forced to develop new techniques. In particular, an important ingredient in previous proofs is the #CSP^d dichotomy by Huang and Lu [19]. Here #CSP^d denotes #CSP where every variable appears a multiple of d times. The very first step in the #CSP^d dichotomy proof uses the pinning technique. Multiple copies of an instance are created and vertices are connected across different copies. This construction violates planarity. Moreover, this violation is unavoidable, a consequence of the new dichotomy. Due to our newly discovered tractable problems, the putative form of a planar #CSP^d dichotomy is *false* when $d \geq 5$. Nevertheless, we prove a dichotomy for planar #CSP^2 for which the putative form is, luckily for us, true (but not obvious in hindsight). A dichotomy for planar #CSP^2 is essential because it captures a significant fraction of planar Holant problems either directly or through reductions. We manage to prove the planar Holant dichotomy without appealing to planar #CSP^d for $d \geq 3$.

The proof of the planar #CSP^2 dichotomy comprises the entire Part II of the full version [8] starting on page 63. A brief outline of the proof is given in Section IV. Among the concepts and techniques introduced are some special tractable families of constraint functions specific to the #CSP^2 framework. We also introduce a *derivative* ∂ and its inverse operator *integral* $\int$ to streamline the proof argument. There is also an application of the theory of *cyclotomic fields*.

We began this project expecting to prove the putative form of the planar Holant dichotomy. It was determined that a planar #CSP^d dichotomy would be both a more modest and attainable intermediate step as well as a good launch station for the final goal. However after some attempts, even the planar #CSP^d dichotomy appeared too difficult to achieve, and so we scaled back the ambition to solve just $d = 2$. Luckily, a successful #CSP^2 dichotomy can carry most of the weight of a full #CSP^d dichotomy, *and*, as it turned out, the putative form of the planar #CSP^2 dichotomy is *true* while that for planar #CSP^d is not. Ironically, many steps of our proof in this paper were guided by the putative form of the complexity classification. The discovery of the new tractable problems changed the original plan, but also helped complete the picture.

Coming back to the challenge of the P vs. NP question posed by Valiant's holographic algorithms, we venture the opinion that the dichotomy theorem provides a satisfactory answer. Indeed, it would be difficult to conceive a world where #P is in fact equal to P, and yet all this algebraic theory can somehow maintain a consistent, sharp but faux division where there is none. (Consider the following Gedankenexperiment: #P is really equal to P, but the Supreme Fascist keeps scores on how much of #P we have learned to be in P. For every problem in this broad class that is yet unknown to be in P the SF lets we prove it #P-hard—a superfluous notion really. Nevertheless for every problem in the class known to be in P, the SF makes sure our proof method for #P-hardness on that problem fails, thus preventing one from making the ultimate discovery.)

II. Preliminaries

Fix a set of local constraint functions $\mathcal{F}$. A *signature grid* $\Omega = (G, \pi)$ is a tuple, where $G = (V, E)$ is a graph, π labels each $v \in V$ with a function $f_v \in \mathcal{F}$ with input variables from the incident edges $E(v)$ at v. Each f_v maps $\{0,1\}^{\deg(v)}$ to $\mathbb{C}$. We consider all 0-1 edge assignments. 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 restriction of σ to $E(v)$. The counting problem on the instance Ω is to compute $\text{Holant}_\Omega = \sum_{\sigma: E \to \{0,1\}} \prod_{v \in V} f_v \left(\sigma \mid_{E(v)}\right)$. The Holant problem parameterized by the set $\mathcal{F}$ is denoted by $\text{Holant}(\mathcal{F})$ and Pl-$\text{Holant}(\mathcal{F})$ is defined similarly using a signature grid with a planar graph.

A function f_v can be represented by listing its values in lexicographical order as in a truth table, which is a vector in $\mathbb{C}^{2^{\deg(v)}}$, or as a tensor in $(\mathbb{C}^2)^{\otimes \deg(v)}$. A symmetric function f on k Boolean variables can be expressed as $[f_0, f_1, \ldots, f_k]$, where f_w is the value of f on inputs of Hamming weight w. This is called the *signature* of f, and we may use the terms "signature" and "function" interchangeably below. For example, we use $=_k$ to denote the EQUALITY signature $[1, 0, \ldots, 0, 1]$ of arity k.

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

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

For a 2-by-2 matrix T and a signature set $\mathcal{F}$, define $T\mathcal{F} = \{g \mid \exists f \in \mathcal{F} \text{ of arity } n,\ g = T^{\otimes n} f\}$, similarly for $\mathcal{F}T$. Whenever we write $T^{\otimes n} f$ or $T\mathcal{F}$, we view the signatures as column vectors; similarly for $fT^{\otimes n}$ or $\mathcal{F}T$ as row vectors. We use $\text{Holant}\left(\mathcal{R} \mid \mathcal{G}\right)$ to denote the Holant problem on bipartite graphs $H = (U, V, E)$, where each vertex in U or V is assigned a signature in $\mathcal{R}$ or $\mathcal{G}$, respectively. Let T be an invertible 2-by-2 matrix. The holographic transformation defined by T is the following operation: given a signature grid Ω with underlying graph H, we create a new grid Ω' such that the graph is still H, and any functions f on the left (or g on the right) is replaced by $fT^{\otimes n}$ (or $\left(T^{-1}\right)^{\otimes m} g$) where n and m are arities of f and g. We frequently apply a holographic transformation defined by the matrix $Z = \left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$.

Theorem II.1 (Valiant's Holant Theorem [36]). *If there is a holographic transformation mapping signature grid Ω to Ω', then* $\text{Holant}_\Omega = \text{Holant}_{\Omega'}$.

In order to do holographic transformations on a general graph, we can always modify it into an equivalent bipartite graph preserving the Holant value as follows. For each edge in the graph, we replace it by a path of length two. (This operation is called the *2-stretch* of the graph and yields the edge-vertex incidence graph.) Each new vertex is then assigned the binary EQUALITY signature $(=_2) = [1, 0, 1]$.

We say a signature set $\mathcal{F}$ is $\mathscr{C}$-transformable if there exists a transformation $T \in \mathbf{GL}_2(\mathbb{C})$ such that $[1, 0, 1]T^{\otimes 2} \in \mathscr{C}$ (viewed as row vectors) and $\mathcal{F} \subseteq T\mathscr{C}$ (viewed as column vectors). The importance of this definition is that if Pl-$\text{Holant}(\mathscr{C})$ is tractable, then Pl-$\text{Holant}(\mathcal{F})$ is also tractable for any $\mathscr{C}$-transformable set $\mathcal{F}$.

III. Main Theorem and Proof Outline

In this section, we state the main theorem and give an outline of its proof.

We use $\mathscr{A}$, $\mathscr{P}$, $\mathscr{V}$, and $\mathscr{M}$ to denote four base classes of tractable signatures. The classes $\mathscr{A}$ and $\mathscr{P}$ are identified as tractable for #CSP [14]. Problems defined by $\mathscr{A}$ are tractable essentially by Gauss sums [2].

The signatures in $\mathscr{P}$ are tensor products of signatures whose supports are among two complementery bit vectors. Problems defined by them are tractable by a propagation algorithm. The class $\mathscr{V}$ contains vanishing signatures [17], [6], which means the Holant value is always 0. We split $\mathscr{V}$ into $\mathscr{V}^+$ and $\mathscr{V}^-$. Any subset of $\mathscr{V}^+$ or $\mathscr{V}^-$ vanishes, but mixing these two classes does not necessarily vanish. Valiant [34], [33] introduced matchgates, which we denote by $\mathscr{M}$. They can be locally expressed by weighted perfect matchings, so problems defined by them are tractable by the FKT algorithm over planar graphs. The full version [8] contains complete definitions and characterizations of these four classes. As mentioned at the end of last section, a problem defined by a signature that is transformable to any of these tractable classes is also tractable. In fact, $\mathscr{V}$ is closed under this transformable notion.

We need some more notations. Let $\mathscr{R}_2^{\pm}$ denote the set of all unary signatures plus symmetric signatures $f = [f_0, f_1, \ldots, f_n]$ satisfying $f_i \pm 2i f_{i+1} + f_{i+2} = 0$ for all $0 \le i \le n-2$. For a signature set $\mathcal{F}$, let $\mathcal{F}^*$ denote $\mathcal{F}$ with all degenerate signatures $[a,b]^{\otimes m}$ replaced by unary $[a,b]$. We denote by $\textsc{ExactOne}_d$ the signature $[0,1,0,\ldots,0]$ of arity d. Let $\mathcal{EO} = \{\textsc{ExactOne}_d \mid d \ge 3\}$.

Theorem III.1. *Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures in Boolean variables. Then* $\text{Pl-Holant}(\mathcal{F})$ *is #P-hard unless $\mathcal{F}$ satisfies one of the following conditions:*

1) *All non-degenerate signatures in $\mathcal{F}$ are of arity at most 2;*
2) *$\mathcal{F}$ is $\mathscr{A}$- or $\mathscr{P}$-transformable;*
3) $\mathcal{F} \subseteq \mathscr{V}^\sigma \cup \{f \in \mathscr{R}_2^\sigma \mid \text{arity}(f) = 2\}$ *for some* $\sigma \in \{+,-\}$*;*
4) *All non-degenerate signatures in $\mathcal{F}$ are in $\mathscr{R}_2^\sigma$ for some $\sigma \in \{+,-\}$.*
5) *$\mathcal{F}$ is $\mathscr{M}$-transformable;*
6) $\mathcal{F} \subseteq Z(\mathscr{P} \cup \mathcal{EO})$ *or* $Z\left(\mathscr{P} \cup \left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]\mathcal{EO}\right)$*, and the greatest common divisor of the arities of all signatures in $\mathcal{F}^* \cap \mathscr{P}_2$ is at least* 5.

In each exceptional case, $\text{Pl-Holant}(\mathcal{F})$ *is computable in polynomial time.* $\text{Holant}(\mathcal{F})$ *is computable in polynomial time without planarity if $\mathcal{F}$ satisfies conditions 1, 2, 3, or 4, and is #P-hard otherwise.*

Proof sketch: We first prove a dichotomy theorem when $\mathcal{F}$ contains a single non-degenerate signature f of arity ≥ 3 (cf. Theorem 6.1 in the full version [8]). The proof is by induction on the arity of f. Base cases are when the signature has arity 3 or 4, which have been proved in previous work [12], [18]. The inductive step reduces the arity of f by two each time, and then we apply the induction hypothesis. This essentially yields seven different cases that are $\mathscr{A}$- or $\mathscr{P}$- or $\mathscr{M}$-transformable or in $\mathscr{V}$ (not in 1-to-1 correspondence), plus the possibility of degenerate cases. However, we can roughly split them into two categories: (1) those tractable by orthogonal and related transformations; and (2) those tractable by a $Z = \left[\begin{smallmatrix}1&1\\i&-i\end{smallmatrix}\right]$ transformation. We show that any case in category (1) can be solved by the Pl-#CSP^2 dichotomy via reductions. We handle each case in category (2) separately and show hardness using gadget constructions and polynomial interpolations.

Given the single signature dichotomy, we assume that every nontrivial signature in $\mathcal{F}$ falls into one of the two categories above. Again, if any signature is in category (1), then we can apply the Pl-#CSP^2 dichotomy through reductions. Otherwise, all nontrivial signatures are from category (2). Then we rule out any possible mixing of signatures in $\mathscr{V}$ with other signatures in category (2). This leaves two kinds of signatures in category (2), from $Z\mathscr{P}$ or from $Z\mathcal{EO} \cup Z\left[\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right]\mathcal{EO}$. A putative form of the planar Holant dichotomy would dictate that any mixture from these two sets is intractable. However, we found that there are tractable cases violating the putative dichotomy, which are summarized above as Case 6. Then we finish the proof by showing that there are no other tractable cases. ∎

IV. Dichotomy for Pl-#CSP^2

In this section, we state the dichotomy for Pl-#CSP^2, and provide a sketch of the proof here. Let $\mathcal{T}_k = \left\{\left[\begin{smallmatrix}1 & 0\\ 0 & \omega\end{smallmatrix}\right] \in \mathbb{C}^{2\times 2} \mid \omega^k = 1\right\}$. Let $\widehat{\mathscr{M}} = \left[\begin{smallmatrix}1 & 1\\ 1 & -1\end{smallmatrix}\right] \mathscr{M}$.

Theorem IV.1. *Let $\mathcal{F}$ be a set of symmetric signatures. Then Pl-#CSP$^2(\mathcal{F})$ is #P-hard unless $\mathcal{F}$ satisfies one of the following conditions:*

1) *there exists $T \in \mathcal{T}_8$ such that $\mathcal{F} \subseteq T\mathscr{A}$;*
2) *$\mathcal{F} \subseteq \mathscr{P}$;*
3) *there exists $T \in \mathcal{T}_4$ such that $\mathcal{F} \subseteq T\widehat{\mathscr{M}}$.*

In each exceptional case, Pl-#CSP$^2(\mathcal{F})$ is computable in polynomial time.

Proof Sketch: We first define some tractable families of signatures specific to the Pl-#CSP^2 framework. Let $\widetilde{\mathscr{A}} = \mathscr{A} \cup \left[\begin{smallmatrix}1 & 0\\ 0 & e^{\pi i/4}\end{smallmatrix}\right] \mathscr{A}$ and $\widetilde{\mathscr{M}} = \widehat{\mathscr{M}} \cup \left[\begin{smallmatrix}1 & 0\\ 0 & i\end{smallmatrix}\right] \widehat{\mathscr{M}}$. One can show that $\widetilde{\mathscr{A}}$ covers Case 1 above, and $\widetilde{\mathscr{M}}$ covers Case 3. The proof will revolve around these tractable classes.

The overall plan is to break the proof into two main steps.

The first step is to prove the dichotomy theorem for Pl-#CSP$^2(\mathcal{F})$ when there is at least one nonzero signature of *odd* arity in $\mathcal{F}$. In this case we can make use of a lemma that shows that we can simulate Pl-#CSP$(\mathcal{F})$ by Pl-#CSP$^2(\mathcal{F})$ if $\mathcal{F}$ includes a unary signature $[a, b]$ with $ab \neq 0$. Then we can apply the known planar #CSP dichotomy [18] for Pl-#CSP. However this strategy (provably) *cannot* work in the case when every signature in $\mathcal{F}$ satisfies the *parity* constraint. In that case we employ other means. This first step of the proof is relatively uncomplicated.

The second step is to deal with the case when all signatures in $\mathcal{F}$ have even arity. This is where the real difficulty lies. In this case it is impossible to directly construct *any* unary signature. So we cannot use that lemma pertaining to a unary signature. But we prove another lemma which provides a way to simulate Pl-#CSP$(\mathcal{F})$ by Pl-#CSP$^2(\mathcal{F})$ in a *global* fashion, *if* $\mathcal{F}$ includes some tensor power of the form $[a, b]^{\otimes 2}$ where $ab \neq 0$. Moreover, we have a lucky break (for the complexity of the proof) if $\mathcal{F}$ includes a signature that is in $\widehat{\mathscr{M}} \setminus (\mathscr{P} \cup \widetilde{\mathscr{A}})$. In this case, we can construct a special binary signature, and obtain $[1, 1]^{\otimes 2}$ by interpolation. This proof uses the theory of *cyclotomic fields*. This simplifies the proof greatly. For all other cases (when $\mathcal{F}$ has only even arity signatures), the proof gets going in earnest—we will attempt an induction on the arity of signatures.

The lowest arity of this induction will be two. We will try to reduce the arity to two whenever possible; however for many cases an arity reduction to two destroys the #P-hardness at hand. Therefore the true basis of this induction proof of Pl-#CSP^2 starts with arity 4. Consequently we will first prove a dichotomy theorem for Pl-#CSP$^2(f)$, where f is a signature of arity 4. Several tools will be used. These include the rank criterion for redundant signatures, complex weighted k-regular graph homomorphisms [5] for arity two signatures, and a trick we call the *Three Stooges* by domain pairing.

However in the next step we do not attempt a general Pl-#CSP^2 dichotomy for a *single* signature of even arity. This would have been natural at this point, but it would have been too difficult. We will need some additional leverage by proving a conditional No-Mixing Lemma for pairs of signatures of even arity. So, taking a detour, we prove that for two signatures f and g both of even arity, that individually belong to some tractable class, but do not belong to a single tractable class in the conjectured dichotomy (that is yet to be proved), the problem Pl-#CSP$^2(f, g)$ is #P-hard. We prove this No-Mixing Lemma for any pair of signatures f and g both of even arity, not restricted to arity 4. Even though at this point we only have a dichotomy for a single signature of arity 4, we prove this No-Mixing Lemma for higher even arity pairs f and g by simulating two signatures f' and g' of arity 4 that belong to different tractable sets, from that of Pl-#CSP$^2(f, g)$. After this arity reduction (within the No-Mixing Lemma), we prove that Pl-#CSP$^2(f', g')$ is #P-hard by the dichotomy for a *single* signature of arity 4. After this, we prove a No-Mixing Lemma for a *set* of signatures $\mathcal{F}$ of even arities, which states that *if* $\mathcal{F}$ is contained in the union of all tractable classes, then it is still #P-hard unless it is *entirely* contained in a single tractable class. Note that at this point we still only have a *conditional* No-Mixing Lemma in the sense that we have to assume every signature in $\mathcal{F}$ belongs to some tractable set.

We then attempt the proof of a Pl-#CSP^2 dichotomy for a *single* signature of arbitrary even arity. This uses all the previous lemmas, in particular the (conditional) No-Mixing Lemma for a set of signatures. However, after completing the proof of this Pl-#CSP^2 dichotomy for a single signature of even arity, the No-Mixing Lemma becomes absolute.

Finally we extend the dichotomy for a single signature of even arity to a dichotomy theorem for Pl-#CSP$^2(\mathcal{F})$ where all signatures in $\mathcal{F}$ have even arity. Together with the first main step when $\mathcal{F}$ contains some nonzero signature of odd arity, this completes the proof of Theorem IV.1. ■

V. New Tractable Problems and Related Hardness Results

We are not able to include the whole proof of Theorem III.1. In this last section, we highlight a tractable case and include some related hardness results, summarized as follows.

Theorem V.1. $\text{Pl-Holant}\,(\neq_2 \mid =_k, \mathcal{EO})$ *is #P-hard when* $k \in \{3, 4\}$*, and is computable in polynomial time when* $k \in \{1, 2\}$ *or* $k \geq 5$*.*

Under Z, $\text{Pl-Holant}\,(\neq_2 \mid =_k, \mathcal{EO})$ is $\text{Holant}(Z(=_k), Z(\mathcal{EO}))$. When $k \in \{1, 2\}$, the problem is tractable by either Case 4 or Case 5 of Theorem III.1. The interesting tractable case is when $k \geq 5$, belonging to Case 6. The claim about hypergraph #PM in the introduction follows from Theorem V.1, where tractability follows directly and hardness requires a gadget, which we omit here.

A. Tractability when $k \geq 5$

We first prove that $\text{Pl-Holant}\,(\neq_2 \mid =_k, \mathcal{EO})$ is tractable when $k \geq 6$. After that, we consider $k = 5$. A key observation is that a planar (bipartite) graph cannot be simple if its degrees are large enough. The proof is a straightforward application of Euler's characteristic equation for planar graphs.

Lemma V.2. *Let* $G = (L \cup R, E)$ *be a planar bipartite graph with parts* L *and* R*. If every vertex in* L *has degree at least* 6 *and every vertex in* R *has degree at least* 3*, then* G *cannot be simple.*

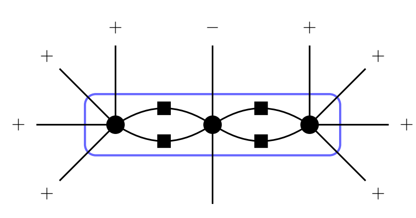


Figure 1: An E_6-block. Circles are $=_6$ and squares are $\neq_2$.

For $\text{Pl-Holant}\,(\neq_2 \mid =_k, \mathcal{EO})$, we may contract edges between $=_k$ and between ExactOne functions. Note that we want to count the number of satisfying assignments as there is no weight. We call an edge *pinned* if it has the same value in all satisfying assignments, if there is any. Any connection among ExactOne's either creates pinned edges, or results in a larger ExactOne. We create components called E_k-blocks composed by $=_k$'s and $\neq_2$'s. An E_k-block is *trivial* if it has no satisfying assignment. A nontrivial E_k-block has exactly two complementary assignments, and we mark edges with signs "+"

and "−" such that edges with the same sign (or distinct signs) take the same value (or distinct values). Figure 1 pictures an example. Parallel edges between an E_k-block and an ExactOne always result in pinned edges. Lemma V.2 does not give us tractability for the case of $k \geq 6$ directly. The reason is that E_k-blocks may have arity less than 6, in which case Lemma V.2 does not apply. However, for $k \geq 6$ and a nontrivial E_k-block of arity n where $n < 6$, we can show that it is either a binary $\neq_2$, or has arity 4, identified in Figure 2a up to a rotation.

(a) Two different arity 4 E_k-blocks.

(b) Replace them by parallel $\neq_2$'s.

Figure 2: Arity 4 E_k-blocks.

In the following lemma, we show how to replace an E_k-block of arity 4 by some other signatures while keeping track of, but not preserving, the Holant value.

Lemma V.3. *For any integer $k \geq 6$,* Pl-Holant $(\neq_2 \mid =_k, \mathcal{EO})$ *is computable in polynomial time.*

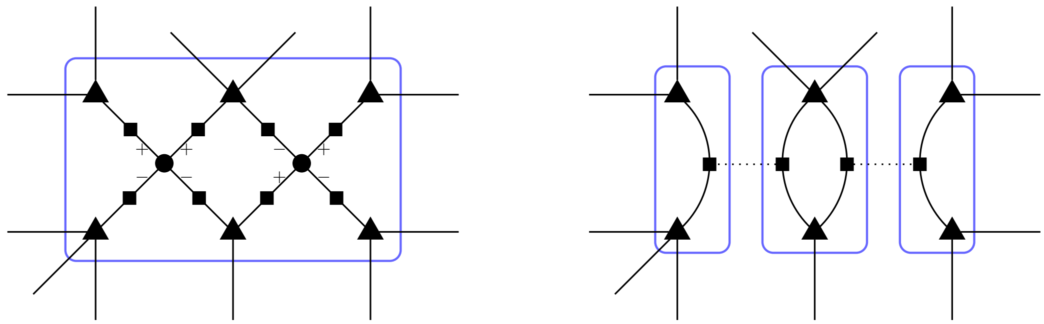
Proof: Let Ω be a connected instance of Pl-Holant $(\neq_2 \mid =_k, \mathcal{EO})$. When an edge is pinned to a known value, we get a smaller instance of Pl-Holant $(\neq_2 \mid =_k, \mathcal{EO})$ without changing the number of satisfying assignments. In our algorithm, we may also find a contradiction and simply return 0.

We claim that there always exists an edge in Ω that is pinned, unless Ω does not contain $=_k$, or does not contain ExactOne$_d$ functions (for some $d \geq 3$), or there is a contradiction. Furthermore if there are $=_k$ or ExactOne$_d$ functions (for some $d \geq 3$), in polynomial time we can find a pinned edge with a known value, or return that there is a contradiction. (If there is a contradiction in Ω, we may still return a purported pinned edge with a known value, which we can apply and simplify Ω. The contradiction will eventually be found.) If Ω does not contain $=_k$, or does not contain ExactOne$_d$ functions (for some $d \geq 3$), then the problem is tractable, since Ω is an instance of $\mathscr{M}$, or an instance of $\mathscr{P}$. The lemma follows from the claim, since we either recurse on a smaller instance or have a tractable instance.

Suppose Ω is an instance where at least one $=_k$ and at least one ExactOne$_d \in \mathcal{EO}$ appear. If a signature ExactOne$_d \in \mathcal{EO}$ is connected to itself by a self-loop through a $\neq_2$, then the remaining $d-2 \geq 1$ edges are pinned to 0 with a factor of 2 to the Holant. Suppose two signatures ExactOne$_d$ and ExactOne$_\ell$ from $\mathcal{EO}$ are connected by some number of $\neq_2$'s. Depending on the number of connecting edges being 1 or 2 or ≥ 3, we replace all three signatures by ExactOne$_{d+\ell-2}$, or find pinned edge, or return 0. We hence assume no connection between any pair of ExactOne's.

Define an E_k-block as a connected component composed of $=_k$ and $\neq_2$. All external connecting edges of each E_k-block are marked with + or − and this can be found by testing bipartiteness of an E_k-block where we treat $\neq_2$'s as edges. If any E_k-block is not bipartite, then it is trivial and we return 0. We contract all E_k-blocks and maintain planarity, one edge at a time, and remove self loops. We may assume all E_k-blocks are nontrivial. If there is a nontrivial E_k-block of arity 2, its signature is $\neq_2$. We replace it with an edge assigned $\neq_2$ to form an instance Ω', maintaining planarity, such that any pinned edge in Ω' corresponds to a pinned edge in Ω. This new edge is between ExactOne signatures and can be dealt with as described earlier. So we may assume the arity of any E_k-block is at least 4. Since $k \geq 6$, the only possible E_k-blocks of arity 4 are those in Figure 2a up to a rotation. Since there is at least one ExactOne$_d$ signature with $d \geq 3$, forming E_k-blocks does not consume all of Ω.

After these steps we may consider Ω a bipartite graph, with one side consisting of E_k-blocks and the other side consisting of EXACTONE signatures. They are now connected by edges assigned $=_2$. It is easy to verify that parallel edges between an E_k-block and an EXACTONE$_d$ signature always lead to some pinned edges. Therefore, we may assume there are no parallel edges between any E_k-block and any EXACTONE signature.



(a) An EO-Eq-4-block. Triangles are assigned EXACTONE signatures, squares are assigned $\neq_2$, and circles are E_k-blocks of arity 4.

(b) Break the EO-Eq-4-block into three components, The one in the middle contains a cycle, and hence is degenerate. The other two are EXACTONE functions.

Figure 3: EO-Eq-4-blocks

Now consider E_k-blocks of arity 4 with EXACTONE signatures together. Call a connected component consisting of E_k-blocks of arity 4 and EXACTONE an EO-Eq-4-block. Figure 3a illustrates an example. Notice that the two possibilities of E_k-blocks of arity 4 can be viewed as two parallel $\neq_2$'s but with some correlation between them, namely their satisfying assignments are paired up in a unique way. This is illustrated in Figure 2b. Note that the two dotted lines in Figure 2b represent different correlations.

At this point, we would like to replace every arity 4 E_k-block by two parallel $\neq_2$'s. However this replacement destroys the equivalence of the Holant values, before and after.

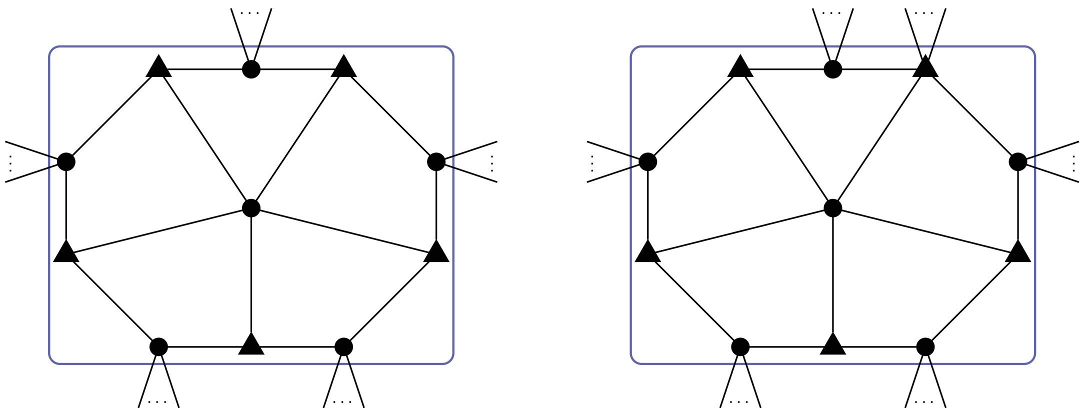
The surprising move of this proof is that we shall do so anyway!

We ignore the correlation for the time being and replace every arity 4 E_k-block by two parallel $\neq_2$'s as in Figure 2b. This replacement produces a *planar* signature grid Ω_1. Every edge in Ω_1 corresponds to a unique edge in Ω. The set of satisfying assignments of Ω_1 is a superset of that of Ω. Moreover, if there is an edge pinned in Ω_1 to a known value, the corresponding edge is also pinned in Ω to the same value. Once we find a pinned edge in Ω_1, we revert back to work in Ω and apply the pinning to the pinned edge.

All that remains to be shown is that pinning always happens in Ω_1. Each EO-Eq-4-block splits into some number of connected components in Ω_1. Figure 3b is an example. We can show that any cycle in such a component creates at least one pinned edge. Hence we may assume there are no cycles in these components, and every such component forms a tree, whose vertices are EXACTONE functions and edges are $\neq_2$'s. Suppose there are $n \geq 2$ vertices and t many leaves in such a tree. One can verify that replacing the whole tree by an EXACTONE$_t$ function of the same arity t maintains the number of satisfying assignments. Since each vertex in the tree has degree at least 3, we have $t \geq 3n - 2(n-1) = n + 2 \geq 4$. We replace these components by EXACTONE$_t$'s.

Thus, each connected component in the graph underlying Ω_1 is a planar bipartite graph with E_k-blocks of arity at least 6 on one side and EXACTONE$_d$ signatures of arity at least 3 on the other. By Lemma V.2, no component is simple, so there are parallel edges between some E_k-block and some

EXACTONE$_d$ signature. Parallel edges between two parts lead to pinned edges, and we can find a pinned edge with a known value in polynomial time. This finishes the proof. ■



(a) Type 1. (b) Type 2.

Figure 4: The wheel structures. Each circle is a E_5-block and triangle a EXACTONE function.

Unlike the situation in Lemma V.2, a planar $(5, 3)$-regular bipartite graph can be simple. However, we show that such graphs have a special structure. We call this structure a "wheel", which is pictured in Figure 4. There is an arity 5 vertex v in the middle, and all faces adjacent to this vertex must be 4-gons (i.e. quadrilaterals). Moreover, at least four neighbors of v are of degree 3. Depending on the degree of the fifth neighbor (whether it is 3 or not), we have two types of wheels, which are pictured in Figure 4a and Figure 4b.

Lemma V.4. *Let $G = (L \cup R, E)$ be a planar bipartite graph with parts L and R. Suppose every vertex in L has degree at least 5 and every vertex in R has degree at least 3. If G is simple, then there exists one of the two wheel structures in Figure 4 in G.*

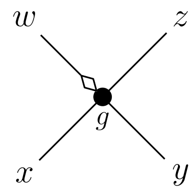
Proof: Let $V = L \cup R$ and F be the set of faces. We assign a "score" s_v on each vertex $v \in V$. We will define s_v so that $\sum_{v \in V} s_v = |V| - |E| + |F| = 2 > 0$. The base score is $+1$ for each vertex, which accounts for $|V|$. For each k-gon face, we assign $\frac{1}{k}$ to each of its vertex. This accounts for $|F|$. Notice that G is bipartite. Hence $k \geq 4$ and a score coming from a face can be at most $\frac{1}{4}$.

For $-|E|$, we need to separate two cases. If one of the two endpoints has degree 3, we give the degree 3 vertex a score of $-\frac{7}{12}$, and the other one $-\frac{5}{12}$. This is well defined because all degree 3 vertices are in R. Otherwise, we give each endpoint $-\frac{1}{2}$. This accounts for $-|E|$.

One can verify that $s_v \leq 0$ unless $v \in L$ has degree 5. Since the total score is positive, there must exist $v \in L$, v has degree 5 and $s_v > 0$. We then claim that there exists such a v so that all its adjacent faces are 4-gons. Suppose otherwise. One can show that a positively scored vertex v is adjacent to exactly one face with more than 4 edges. Call this face F_v.

In F_v, v has two neighbors in R. We match all vertices that have positive scores to their own clockwise next one in F_v. We do this matching in all faces containing at least one positively scored vertex. Suppose a vertex $u \in R$ is matched with ℓ different vertices. This means that u is adjacent to at least ℓ many k-gons with $k \geq 6$. One can verify that $s_u \leq -\frac{\ell}{12}$. It implies that the total score of u and all positively scored vertices matched with u is at most 0. However each positively scored vertex is matched with a vertex in R. Hence the total score cannot be positive. Contradiction.

Therefore there exists $v \in L$ such that $s_v > 0$, $\deg(v) = 5$, and all adjacent faces are 4-gons. We

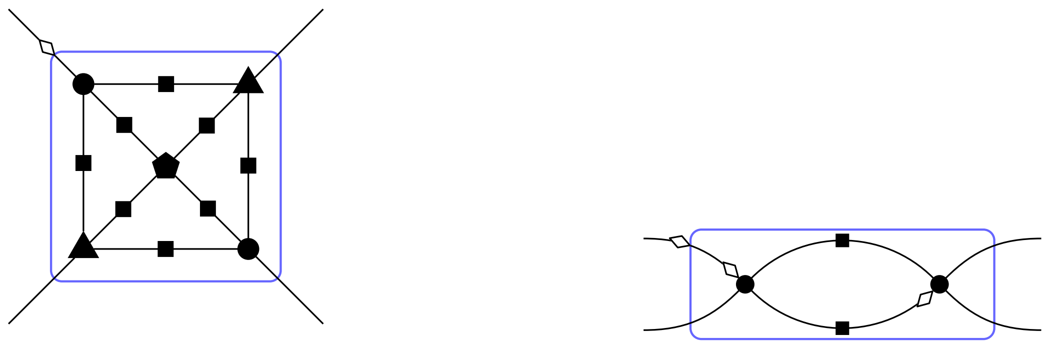


(a)

$$\begin{array}{c|cccc} {}_{wx}\,{}^{zy} & 00 & 01 & 10 & 11 \\ \hline 00 & g^{0000} & g^{0010} & g^{0001} & g^{0011} \\ 01 & g^{0100} & g^{0110} & g^{0101} & g^{0111} \\ 10 & g^{1000} & g^{1010} & g^{1001} & g^{1011} \\ 11 & g^{1100} & g^{1110} & g^{1101} & g^{1111} \end{array}$$

(b)

Figure 5: The quaternary signature g is assigned to the vertex in (a). Its first input corresponds to the edge marked with the diamond, which is w. The order of the remaining inputs is given by traveling counterclockwise. In (b), g^{wxzy} denotes the value $g(w, x, y, z)$.



(a) $(\neq_2 \mid [0, 1, 0, 0, 0], [0, 0, 0, 1, 0], \hat{g})$-gate on right side

(b) Gadget with a symmetric signature matrix

Figure 6: Two gadgets used in Lemma V.6.

further note that at most one neighbor of v is of degree ≥ 4, for otherwise, $s_v \leq 0$. It is type 1 as in Figure 4a if all neighbors of v has degree 3, and is type 2 as in Figure 4b otherwise. ■

Either structure in Figure 4 leads to pinned edges. We get the following lemma, which finishes the tractability of Theorem V.1.

Lemma V.5. Pl-Holant $(\neq_2 \mid =_5, \mathcal{EO})$ *is computable in polynomial time.*

B. Hardness when $k \in \{3, 4\}$

We prove the hardness of Theorem V.1. The proofs differ for $k = 3$ and $k = 4$. For $k = 3$, we use the following technical lemma. This lemma is invoked three times in the full proof of Theorem III.1. The main challenge in these proofs is how to build planar gadgets. As discussed in the previous subsection, parallel edges lead to degeneracy in the instance. A simple calculation based on Euler's characteristic implies that a large number of vertices is necessary to avoid parallel edges.

Lemma V.6. *Let* $\hat{g}$ *be the arity* 4 *signature whose support contains only* 0101 *and* 1010 *(invariant under rotations). Then* Pl-Holant $(\neq_2 \mid [0, 1, 0, 0, 0], [0, 0, 0, 1, 0], \hat{g})$ *is #P-hard.*

Proof: For an arity 4 signature, we can express it as a 4-by-4 matrix, where rows are indexed by the two inputs on the left, and columns by the two inputs on the right in reversed order. This is depicted

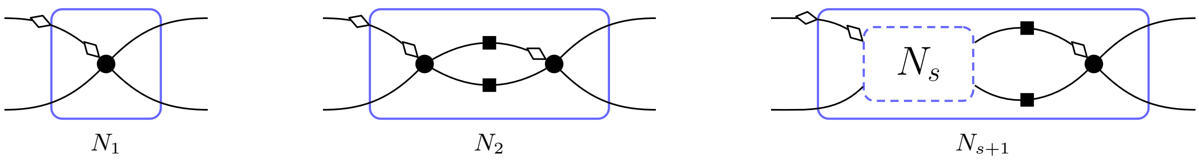
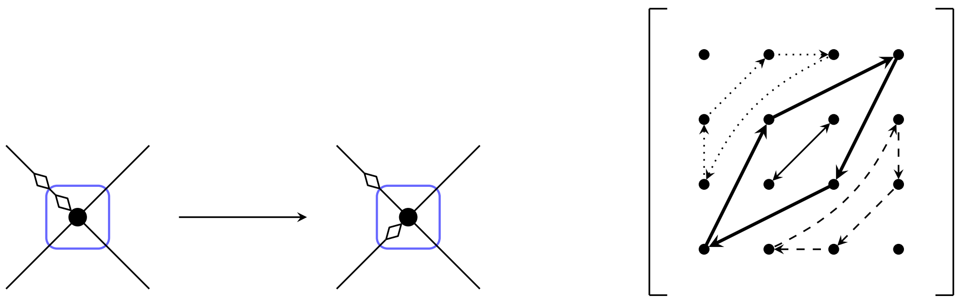


Figure 7: Linear recursive construction used for interpolation in a nonstandard basis.



(a) A counterclockwise rotation

(b) Movement of signature matrix entries

Figure 8: The movement of the entries in the signature matrix of a quaternary signature under a counterclockwise rotation of the input edges. Entires of Hamming weight 1 are in the dotted cycle, entires of Hamming weight 2 are in the two solid cycles (one has length 4 and the other one is a swap), and entries of Hamming weight 3 are in the dashed cycle.

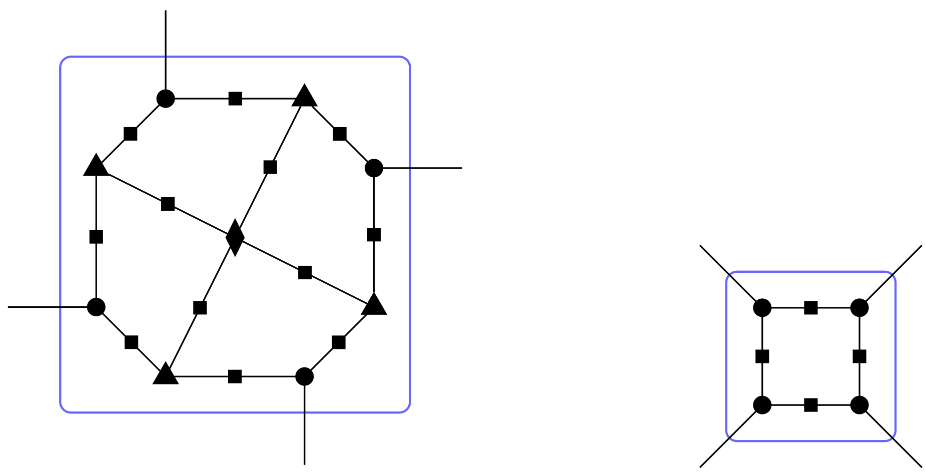
in Figure 5. With this notation, sequential connections correspond to matrix multiplications.

Consider the gadget in Figure 6a. We assign $[0,0,0,1,0]$ to triangles, $[0,1,0,0,0]$ to circles, $\hat{g}$ to the pentagon, and $[0,1,0]$ to squares. The resulting signature is $\hat{h}$ with $M_{\hat{h}} = \left[\begin{smallmatrix} 0&0&0&1\\ 0&1&3&0\\ 0&1&1&0\\ 1&0&0&0 \end{smallmatrix}\right]$. Consider the gadget in Figure 6b. We assign $\hat{h}$ to circles and $[0,1,0]$ to squares. The resulting signature is $\hat{r}$ with $M_{\hat{r}} = \left[\begin{smallmatrix} 0&0&0&1\\ 0&6&4&0\\ 0&4&2&0\\ 1&0&0&0 \end{smallmatrix}\right]$. We use $\hat{r}$ to interpolate a signature $\hat{r}'$ with $M_{\hat{r}'} = \left[\begin{smallmatrix} 0&0&0&1\\ 0&3&1&0\\ 0&1&1&0\\ 1&0&0&0 \end{smallmatrix}\right]$. Consider an instance Ω of $\text{Pl-Holant}\left(\neq_2 \mid \hat{r}'\right)$. Suppose that $\hat{r}'$ appears n times in Ω. We construct from Ω a sequence of instances Ω_s of $\text{Pl-Holant}\left(\neq_2 \mid \hat{r}\right)$ indexed by $s \geq 1$. We obtain Ω_s from Ω by replacing each occurrence of $\hat{r}'$ with the gadget N_s in Figure 7 with $\hat{r}$ assigned to circles and $[0,1,0]$ assigned to squares. In Ω_s, the edge corresponding to the ith significant index bit of N_s connects to the same location as the edge corresponding to the ith bit of $\hat{r}'$ in Ω.

The signature matrix of $\hat{r}'$ is $M_{\hat{r}'} = XPDP^{-1}$ where $X = \left[\begin{smallmatrix} 0&0&0&1\\ 0&0&1&0\\ 0&1&0&0\\ 1&0&0&0 \end{smallmatrix}\right]$, $P = \left[\begin{smallmatrix} 1&0&0&0\\ 0&1&1&0\\ 0&\sqrt{3}&-\sqrt{3}&0\\ 0&0&0&1 \end{smallmatrix}\right]$, and D is a diagonal matrix $\operatorname{diag}\left(1, 1+\sqrt{3}, 1-\sqrt{3}, 1\right)$. The signature matrix of N_s is $M_{N_s} = X(XM_{\hat{r}})^s = XPD_1^sP^{-1}$, where $D_1 = \operatorname{diag}\left(1, 4+2\sqrt{3}, 4-2\sqrt{3}, 1\right)$. We can view our construction of Ω_s as first replacing $M_{\hat{r}'}$ with $XPDP^{-1}$ (each matrix corresponds to a vertex), which does not change the Holant value, and then replacing D with D_1^s.

We stratify the assignments in Ω based on the assignments to the n occurrences of the signature corresponding to D. We only need to consider the assignments that assign i many times the bit patterns 0000 or 1111, j many times the bit pattern 0110, and k many times the bit pattern 1001, since any other assignment contributes a factor of 0. Let c_{ijk} be the sum over all such assignments of the products of evaluations of all other signatures (those corresponding to X, P, and P^{-1}) in Ω except for those corresponding to D. Then $\text{Holant}_\Omega = \sum_{i+j+k=n} \left(1+\sqrt{3}\right)^j \left(1-\sqrt{3}\right)^k c_{ijk}$ and the value of the Holant on Ω_s, for $s \geq 1$, is $\text{Holant}_{\Omega_s} = \sum_{i+j+k=n} \left(\left(4+2\sqrt{3}\right)^{j-k} 4^k \right)^s c_{ijk}$. We view c_{ijk} as unknown variables to be solved, and the Vandermonde system given by Holant_{Ω_s} has full rank. To see this, we only need to show that $\left(4+2\sqrt{3}\right)^{j-k} 4^k \neq \left(4+2\sqrt{3}\right)^{j'-k'} 4^{k'}$ unless $(j,k) = (j',k')$. If $\left(4+2\sqrt{3}\right)^{j-k} 4^k = \left(4+2\sqrt{3}\right)^{j'-k'} 4^{k'}$, then we have $\left(4+2\sqrt{3}\right)^{j-k-(j'-k')} 4^{k-k'} = 1$. Since any nonzero integer power of $4+2\sqrt{3}$ is not rational, we have $j-k = j'-k'$, so $k = k'$ and $j = j'$.

Therefore, by polynomially many oracle calls to Holant_{Ω_s}, we can solve for the unknown c_{ijk}'s and obtain Holant_Ω. After a counterclockwise rotation of $\hat{r}'$ (cf. Figure 8), we get a nonsingular redundant matrix. The hardness follows (cf. Corollary 2.31 in the full version [8]). ■



(a) A cycle-like gadget used twice.

(b) A gadget to realize $\hat{g}$.

Figure 9: Two gadgets in the proof of Lemma V.7.

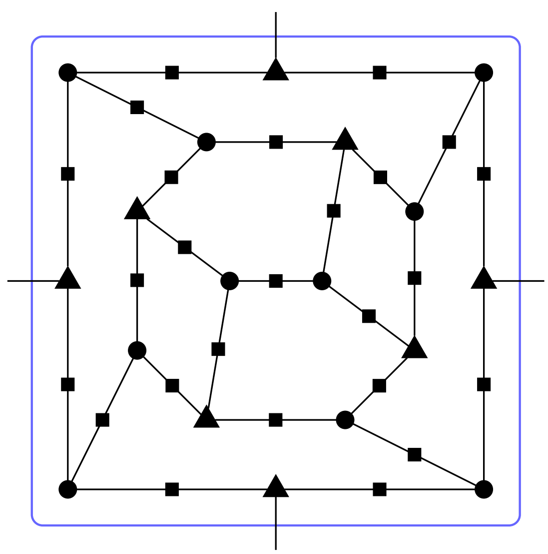


Figure 10: The whole gadget to realize $[0,0,0,1,0]$.

Lemma V.7. $\text{Pl-Holant}\,(\neq_2 \mid =_3, [0,1,0,0])$ *is #P-hard.*

Proof: By connecting two copies of $[0,1,0,0]$ together via $\neq_2$, we have $[0,1,0,0,0]$ on the right. Consider the gadget in Figure 9a. We assign $=_3$ to triangles, $[0,1,0,0]$ to circles, $[0,1,0,0,0]$ to the diamond, and $\neq_2$ to squares. Let f be the signature of this gadget. The support of f is $\{0011, 0110, 1100, 1001\}$. We construct the gadget in Figure 9a again. This time we assign $[0,1,0,0]$ to triangles, $=_3$ to circles, f to the diamond, and $\neq_2$ to squares. The resulting signature has support $\{0111, 1011, 1101, 1110\}$, and therefore is $[0,0,0,1,0]$. The whole gadget is illustrated in Figure 10, where circles are assigned $[0,1,0,0]$, triangles $=_3$, and squares $\neq_2$.

Consider the gadget in Figure 9b. We assign $=_3$ to circles and $\neq_2$ to squares. It follows that the support of the resulting signature is $\{0101, 1010\}$. This is $\hat{g}$ from Lemma V.6. We have constructed $[0,1,0,0,0]$, $[0,0,0,1,0]$, and $\hat{g}$, all on the right, so we are done by Lemma V.6. ■

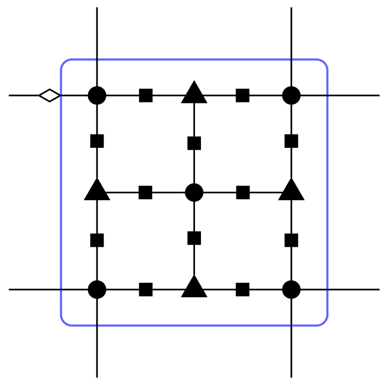


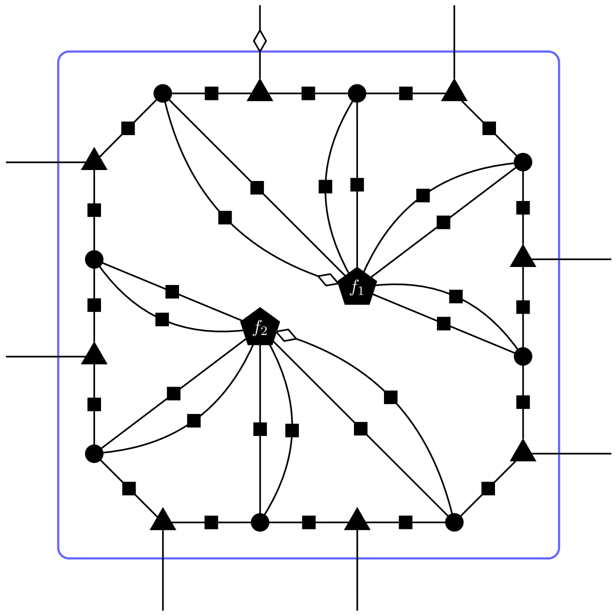
Figure 11: A grid-like gadget used in the proof of Lemma V.8, whose support vectors are 00110011, 11001100, and 11111111.

Lemma V.8. $\text{Pl-Holant}\,(\neq_2 \mid =_4, [0,1,0,0])$ *is #P-hard.*

Proof: Consider the gadget in Figure 11. We assign $\neq_2$ to squares, $=_4$ to circles, and $[0,1,0,0]$ to triangles. The resulting signature has support $\{00110011, 11001100, 11111111\}$, where each vector is the assignment ordered clockwise starting from the diamond. Every two wires at each corner are always of the same value. Further connect each corner to a $=_4$ via two copies of $\neq_2$, resulting in a signature f whose support is $\{11001100, 00110011, 00000000\}$, reversing the original.

Consider the gadget in Figure 12a. We assign $\neq_2$ to squares, $=_4$ to circles, $[0,1,0,0]$ to triangles, and f to pentagons. Each pair of parallel edges coming out of f are at the same corner of f. We call the pentagon above f_1, the one below f_2, and the resulting signature g. We order the inputs to f_1, f_2, and g clockwise starting from the diamond-marked edge. With this notation, we get Table 12b listing the support of g.

The support of g is $\{11111111, 00001111, 0001110, 11110000, 00000000, 11100001\}$, and 00000000 has multiplicity 2. We pair adjacent outputs clockwise, starting from the diamond. We treat g as an arity 4 signature, using $=_4$ to do a domain pairing argument. In the paired domain, $=_4$ becomes $=_2$, which lifts the bipartite restriction. Moreover, 0001110 and 11100001 in the support of g are eliminated as they do not agree on adjacent paired outputs. So in the paired domain, the support of g is $\{1111, 0011, 1100, 0000\}$ with multiplicity 2 for 0000. We rotate g so that the support is $\{1111, 0110, 1001, 0000\}$. The arity 4 signature matrix of g is $\text{diag}(2,1,1,1)$. We can show that Pl-#CSP$([2,1,1]) \leq_T$ Pl-Holant(g) (cf. Lemma



(a) The gadget.

f_1	f_2	g
00000000	00000000	11111111
00110011	00000000	00001111
11001100	00000000	00011110
00000000	00110011	11110000
00110011	00110011	00000000
11001100	00110011	-
00000000	11001100	11100001
00110011	11001100	-
11001100	11001100	00000000

(b) The table of supports.

Figure 12: Another gadget used in the proof of Lemma V.8.

7.2 in the full version [8]), $\text{Pl-Holant}(g)$ is #P-hard by the planar #CSP dichotomy (cf. Theorem 2.27 in the full version [8]), so we are done. ■

Acknowledgment

We are thankful to Pinyan Lu who discussed with us in an early stage of this work. We also thank the anonymous referees for their helpful comments. All authors were supported by NSF CCF-1217549. Heng Guo was also supported by a Simons Award for Graduate Students in Theoretical Computer Science from the Simons Foundation. Tyson Williams was also supported by a Cisco Systems Distinguished Graduate Fellowship.

References

[1] Rodney J. Baxter. *Exactly solved models in statistical mechanics*. Academic press London, 1982.

[2] Jin-Yi Cai, Xi Chen, Richard J. Lipton, and Pinyan Lu. On tractable exponential sums. In *FAW*, pages 148–159. Springer Berlin Heidelberg, 2010.

[3] Jin-Yi Cai and Vinay Choudhary. Some results on matchgates and holographic algorithms. *Int. J. Software and Informatics*, 1(1):3–36, 2007.

[4] Jin-Yi Cai, Vinay Choudhary, and Pinyan Lu. On the theory of matchgate computations. *Theory Comput. Syst.*, 45(1):108–132, 2009.

[5] Jin-Yi Cai and Michael Kowalczyk. Spin systems on k-regular graphs with complex edge functions. *Theoretical Computer Science*, 2012.

[6] Jin-Yi Cai, Heng Guo, and Tyson Williams. A complete dichotomy rises from the capture of vanishing signatures (extended abstract). In *STOC*, pages 635–644. ACM, 2013. *CoRR*, abs/1204.6445.

[7] Jin-Yi Cai, Heng Guo, and Tyson Williams. Holographic algorithms beyond matchgates. In *ICALP*, pages 271–282. Springer Berlin Heidelberg, 2014. *CoRR*, abs/1307.7430.

[8] Jin-Yi Cai, Zhiguo Fu, Heng Guo, and Tyson Williams. A Holant Dichotomy: Is the FKT Algorithm Universal? *CoRR*, abs/1505.02993, 2015.

[9] Jin-Yi Cai, Michael Kowalczyk, and Tyson Williams. Gadgets and anti-gadgets leading to a complexity dichotomy. In *ITCS*, pages 452–467. ACM, 2012.

[10] Jin-Yi Cai and Pinyan Lu. On symmetric signatures in holographic algorithms. *Theory Comput. Syst.*, 46(3):398–415, 2010.

[11] Jin-Yi Cai and Pinyan Lu. Holographic algorithms: From art to science. *J. Comput. Syst. Sci.*, 77(1):41–61, 2011.

[12] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Holographic algorithms with matchgates capture precisely tractable planar #CSP. In *FOCS*, pages 427–436. IEEE Computer Society, 2010.

[13] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Computational complexity of Holant problems. *SIAM J. Comput.*, 40(4):1101–1132, 2011.

[14] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. The complexity of complex weighted Boolean #CSP. *J. Comput. System Sci.*, 80(1):217–236, 2014.

[15] Jan Draisma, Dion C. Gijswijt, László Lovász, Guus Regts, and Alexander Schrijver. Characterizing partition functions of the vertex model. *J. Algebra*, 350:197–206, 2012.

[16] Michael Freedman, László Lovász, and Alexander Schrijver. Reflection positivity, rank connectivity, and homomorphism of graphs. *J. Amer. Math. Soc.*, 20(1):37–51, 2007.

[17] Heng Guo, Pinyan Lu, and Leslie G. Valiant. The complexity of symmetric Boolean parity Holant problems. *SIAM J. Comput.*, 42(1):324–356, 2013.

[18] Heng Guo and Tyson Williams. The complexity of planar Boolean #CSP with complex weights. In *ICALP*, pages 516–527. Springer Berlin Heidelberg, 2013. *CoRR*, abs/1212.2284.

[19] Sangxia Huang and Pinyan Lu. A dichotomy for real weighted Holant problems. In *CCC*, pages 96–106. IEEE Computer Society, 2012. Full version available at `http://www.csc.kth.se/~sangxia/papers/2012-ccc.pdf`.

[20] Ernst Ising. Beitrag zür theorie des ferromagnetismus. *Zeitschrift für Physik*, 31(1):253–258, 1925.

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

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

[23] Michael Kowalczyk. *Dichotomy theorems for Holant problems*. PhD thesis, University of Wisconsin—Madison, 2010. `http://cs.nmu.edu/~mkowalcz/research/main.pdf`.

[24] J. M. Landsberg, Jason Morton, and Serguei Norine. Holographic algorithms without matchgates. *Linear Algebra Appl.*, 438(2):782–795, 2013.

[25] T. D. Lee and C. N. Yang. Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model. *Phys. Rev.*, 87(3):410–419, 1952.

[26] Elliott H. Lieb. Residual entropy of square ice. *Phys. Rev.*, 162(1):162–172, 1967.

[27] Elliott H. Lieb and Alan D. Sokal. A general Lee-Yang theorem for one-component and multicomponent ferromagnets. *Comm. Math. Phys.*, 80(2):153–179, 1981.

[28] Jason Morton. Pfaffian circuits. *CoRR*, abs/1101.0129, 2011.

[29] Jason Morton and Susan Margulies. Polynomial-time solvable #CSP problems via algebraic models and Pfaffian circuits. *CoRR*, abs/1311.4066, 2013. To appear in Journal of Symbolic Computation.

[30] Lars Onsager. Crystal statistics. I. A two-dimensional model with an order-disorder transition. *Phys. Rev.*, 65(3-4):117–149, 1944.

[31] Alexander Schrijver. Characterizing partition functions of the spin model by rank growth. *Indag. Math. (N.S.)*, 24(4):1018–1023, 2013.

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

[33] Leslie G. Valiant. Expressiveness of matchgates. *Theor. Comput. Sci.*, 289(1):457–471, 2002.

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

[35] Leslie G. Valiant. Accidental algorthims. In *FOCS*, pages 509–517. IEEE Computer Society, 2006.

[36] Leslie G. Valiant. Holographic algorithms. *SIAM J. Comput.*, 37(5):1565–1594, 2008.

[37] Leslie G. Valiant. Some observations on holographic algorithms. In *LATIN*, pages 577–590. Springer Berlin Heidelberg, 2010.

[38] Dirk Vertigan. The computational complexity of Tutte invariants for planar graphs. *SIAM Journal on Computing*, 35(3):690–712, 2005.

[39] Dirk Llewellyn Vertigan. *On the computational complexity of Tutte, Jones, Homfly and Kauffman invariants*. PhD thesis, University of Oxford, 1991.

[40] Dominic Welsh. *Complexity: Knots, Colourings and Countings*. London Mathematical Society Lecture Note Series. Cambridge University Press, 1993.

[41] C. N. Yang. The spontaneous magnetization of a two-dimensional Ising model. *Phys. Rev.*, 85(5):808–816, 1952.

[42] C. N. Yang and T. D. Lee. Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Phys. Rev.*, 87(3):404–409, 1952.