THE COMPLEXITY OF SYMMETRIC BOOLEAN PARITY HOLANT PROBLEMS*

HENG GUO[†], PINYAN LU[‡], AND LESLIE G. VALIANT[§]

Abstract. For certain subclasses of NP, \oplus P, or #P characterized by local constraints, it is known that if there exist any problems within that subclass that are not polynomial time computable, then all the problems in the subclass are NP-complete, \oplus P-complete, or #P-complete. Such dichotomy results have been proved for characterizations such as constraint satisfaction problems and directed and undirected graph homomorphism problems, often with additional restrictions. Here we give a dichotomy result for the more expressive framework of Holant problems. For example, these additionally allow for the expression of matching problems, which have had pivotal roles in the development of complexity theory. As our main result we prove the dichotomy theorem that, for the class \oplus P, every set of symmetric Holant signatures of any arities that is not polynomial time computable is \oplus P-complete. The result exploits some special properties of the class \oplus P and characterizes four distinct tractable subclasses within \oplus P. It leaves open the corresponding questions for NP, #P, and $\#_k$ P for $k \neq 2$.

Key words. complexity, parity, dichotomy theorem, parity Holant problems

AMS subject classifications. 68Q25, 68Q17

DOI. 10.1137/100815530

SIAM J. COMPUT.

Vol. 42, No. 1, pp. 324–356

1. Introduction. The complexity class $\oplus P$ is the class of languages L such that there is a polynomial time nondeterministic Turing machine that on input $x \in L$ has an odd number of accepting computations, and on input $x \notin L$ has an even number of accepting computations [37, 34]. It is known that $\oplus P$ is at least as powerful as NP, since NP is reducible to $\oplus P$ via (one-sided) randomized reduction [44]. Also, the polynomial hierarchy is reducible to $\oplus P$ via two-sided randomized reduction [36]. There exist decision problems, such as graph isomorphism, that are not known to be in P but are known to be in $\oplus P$ [1]. The class $\oplus P$ has also been related to other complexity classes via relativization [2]. Further, while the class $\oplus P$ lies between NP and #P, it is known that there are several natural problems, such as 2SAT, that are $\oplus P$ -complete, where the corresponding existence problem is in P [41], and a range of others, including graph matchings and some coloring problems, for which the parity problem is in P but exact counting is #P-complete [43].

As with the classes NP and #P it is an open question whether $\oplus P$ strictly extends P. For certain restrictions of these classes, however, dichotomy theorems are known. For NP a dichotomy theorem would state that any problem in the restricted subclass of NP either is in P or is NP-complete (or both, in the eventuality that NP equals P). Ladner [32] proved that without any restrictions this situation does not hold: if $P \neq NP$, then there is an infinite hierarchy of intermediate problems that are not

^{*}Received by the editors November 19, 2010; accepted for publication (in revised form) November 1, 2012; published electronically February 14, 2013. A preliminary version of this work appeared as [29].

http://www.siam.org/journals/sicomp/42-1/81553.html

[†]Department of Computer Science, University of Wisconsin-Madison, Madison, WI 53706 (hguo@cs.wisc.edu).

[‡]Microsoft Research Asia, Haidian District, Beijing 100080, China (pinyanl@microsoft.com).

[§]School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138 (valiant@seas.harvard.edu). This author's research was supported in part by grants NSF-CCF-04-27129 and NSF-CCF-0964401.

polynomial time interreducible.

The restrictions for which dichotomy theorems are known can be framed in terms of local constraints, most importantly, constraint satisfaction problems (CSP) [35, 20, 4, 5, 6, 23, 26, 21, 25] and graph homomorphism problems [24, 27, 7]. Explicit dichotomy results, where available, manifest a total understanding of the class of computations in question, to within polynomial time reduction, and modulo the collapse of the class.

In this paper we consider dichotomies in a framework for characterizing local properties that is more general than those mentioned in the previous paragraph and is called the Holant framework [13, 15]. A particular problem in this framework is characterized by a set of constraint functions (also called signatures) as defined in the theory of holographic algorithms [42, 41]. The CSP framework can be viewed as the special case of the Holant framework in which equality relations of any arity are always available [15]. The addition of equality relations in CSP makes many sets of constraints complete that would not be otherwise.

A brief description of the Holant framework is as follows. A signature grid $\Omega = (G, \mathcal{F}, \pi)$ is a triple, where G = (V, E) is an undirected graph, \mathcal{F} is a set of functions on variables from a domain D to a range R (possibly different from D), and π labels each $v \in V$ with a function $f_v \in \mathcal{F}$ of arity equal to the degree of that vertex v and associates each argument of f_v with an edge incident to v. An assignment σ maps each edge $e \in E$ to an element of D and determines a value $\prod_{v \in V} f_v(\sigma \mid_{E(v)})$, where E(v) denotes the incident edges of v, and $\sigma \mid_{E(v)}$ denotes the restriction of σ to E(v). The counting problem on the instance Ω is the problem of computing the following sum over all possible assignments σ :

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

For example, consider the PERFECT MATCHING problem on G. This corresponds to $D = \{0, 1\}$ and f_v , which is the EXACTLY ONE function at every vertex of G. Then σ corresponds to a subset of the edges, and Holant_{Ω} counts the number of perfect matchings in G. If we use the AT-MOST-ONE function at every vertex, then we count all (not necessarily perfect) matchings. We use the notation $\text{Holant}(\mathcal{F})$ to denote the class of Holant problems where the functions f_v are chosen from the set \mathcal{F} .

In this paper we consider symmetric Boolean parity Holant problems \oplus Holant(\mathcal{F}); that is, in the definition of Holant_{Ω}, $D = R = \{0, 1\}$, \mathcal{F} is a set of symmetric functions with variables in D and range in R, and summation is modulo two. A function or signature is called symmetric if its output depends only on the Hamming weight of the input. We often denote a symmetric function by the list of its outputs sorted by input Hamming weights in ascending order. For example, [0, 1, 1] is the binary OR function. The output is 1 if the input is 01, 10, or 11, and 0 otherwise. More examples are given in section 2.5.

1.1. Our contribution. The main dichotomy result exhibits four classes of signature sets that are polynomial time computable. The first is the class of affine signatures, \mathscr{A} , which express linear equations and are solvable by Gaussian elimination. The second, \mathscr{M} , corresponds to problems that can be reduced to counting the number of (perfect and general) matchings. The counting counterparts of these problems are #P-hard in general. The third, \mathscr{F} , corresponds to Fibonacci signatures [13] with the addition of the binary Boolean inversion signature [0, 1, 0]. The fourth, \mathscr{O} , is what we call *vanishing signature sets*, for which the Holant value is always even.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

We show that for any other set of symmetric signatures, the \oplus Holant problem is \oplus P-complete. In particular, subsets of the union of these four classes that are not subsets of any one are \oplus P-complete.

THEOREM 1.1. Let \mathcal{F} be a set of symmetric signatures. If $\mathcal{F} \subseteq \mathscr{A}$, $\mathcal{F} \subseteq \mathscr{M}$, $\mathcal{F} \subseteq \mathscr{F} \cup \{[0,1,0]\}$, or $\mathcal{F} \in \mathscr{O}$, then the parity problem \oplus Holant(\mathcal{F}) is computable in polynomial time. Otherwise it is $\oplus P$ -complete.

In this paper, along the way to proving our main result, we prove dichotomies (Theorem 5.3) for both the planar and the general case of \oplus Holant^c. \oplus Holant^c is concerned with signature sets that contain both of the unary signatures [0, 1] and [1, 0] (which, like equivalence relations in #CSP problems, often make sets complete that would not be otherwise).

We also prove a dichotomy result (Theorem 5.4) for the symmetric Boolean parity problem for 2-3 regular bipartite graphs in the case that the signature set consists of one signature of arity two and one of arity three, which is the simplest nontrivial setting (and previously investigated in the Holant framework [13, 14, 31, 10] for #P). Note that such a dichotomy is not subsumed by our Theorem 1.1 since bipartite graphs have a more restricted structure than general graphs.

1.2. Related works. Our main theorem, Theorem 1.1, is the first general dichotomy result for the Holant framework. No dichotomy theorem is known for comparable restrictions of #P, NP, or $\#_kP$ for $k \neq 2$. For #P, dichotomy results are known only for Holant^c problems, where Holant^c denotes that the unary constant signatures [0, 1] and [1, 0] are assumed to be available. The known results for Holant^c are for the symmetric case over the real numbers [15], over the complex numbers [10], and for planar graphs in the former case [16]. For NP, Cook and Bruck [18] gave a dichotomy theorem for singleton sets of constraints of arity up to three in the general nonsymmetric case.

Analogous dichotomy results have been obtained for the #CSP problem modulo k. In Faben's dichotomy theorem for Boolean #CSP modulo k [25], the affine signatures form the only polynomially computable class for general k. For our case of k = 2there is the second class of those that vanish for the simple reason that the solution sets are are always closed under Boolean complementation and therefore always even in number. A dichotomy result is also known for the more general setting of weighted Boolean #CSP modulo k [28], but the tractable classes there appear to have no immediate counterpart in the parity setting where there are no weights.

Finding analogues of our main result for NP, #P, or $\#_kP$ for $k \neq 2$ remains a challenge for the future, as is also the same question for $\oplus P$ for nonsymmetric signatures. For NP, the analogue of symmetric signatures consists of H-factors [33]. Cornuéjols [19] has given a polynomial time algorithm for a certain class of such signatures.

Remark. After this paper was completed, some progress was made toward #P dichotomies regarding the Holant framework. Huang and Lu [30] proved the case of real weighted symmetric Boolean signatures, and Cai, Guo, and Williams [9] proved the complex case. Compared to the result in this paper, matching problems, of which the parity version is tractable, are #P-complete in either real or complex setting. If weights are restricted to real values, there is no vanishing signature. However, the capture of vanishing signatures is the key ingredient of the complex dichotomy.

2. Preliminaries.

2.1. Problems and definitions. As previously stated, a signature grid $\Omega = (H, \mathcal{F}, \pi)$ consists of a graph H = (V, E), where each vertex is labeled by a function

 $f_v \in \mathcal{F}$ of arity equal to the degree $\deg(v)$ of v in H. The label π associates with each v the $f_v \in \mathcal{F}$ and also associates each edge incident to v with an argument of f_v . The Holant problem on instance Ω is that of evaluating $\operatorname{Holant}_{\Omega} = \sum_{\sigma} \prod_{v \in V} f_v(\sigma \mid_{E(v)})$, a sum over all edge assignments $\sigma : E \to D$.

The framework of Holant problems for #P is usually defined for functions f_v mapping $D^{deg(v)} \to \mathbb{C}$. In this paper, for $\oplus P$ we assume throughout functions $f : \{0,1\}^{deg(v)} \to \{0,1\}$.

A function f_v of k arguments can be represented as a truth table of 2^k entries. We use f^{α} to denote the value $f(\alpha)$, where α is a $\{0,1\}$ string of length k. A function $f \in \mathcal{F}$ is also called a *signature*. A *symmetric* function f on k Boolean variables can be expressed as $[f_0, f_1, \ldots, f_k]$, where f_i is the value of f on inputs of Hamming weight i. In this paper we will consider only symmetric signatures. As a signature of arity k must be placed on a vertex of degree k, we can also view the signature as a vertex with k dangling edges.

In this paper, where it is not otherwise stated, we view any entry of a signature as an element of the field $\mathbb{Z}_2 = \{0, 1\}$. The operations and relations on them are then also viewed as being in this field.

A Holant problem is parameterized by a set of signatures.

DEFINITION 2.1. Given a set of signatures \mathcal{F} , we define the following counting problem as Holant(\mathcal{F}):

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

Output: $\operatorname{Holant}_{\Omega}$.

The following family Holant^c of Holant problems is important [15, 10, 16]. This is the class of all Holant problems (on Boolean variables) where the graph has some dangling edges, i.e., edges that have a node of H as one endpoint and an external input as another, and these can be forced to have value 0 or 1. In other words, the unary constant signature [1,0] (for the constant function 0) and the unary constant signature [0,1] (for the constant function 1) are always available for use.

DEFINITION 2.2. Given a set of signatures \mathcal{F} , $\operatorname{Holant}^{c}(\mathcal{F})$ denotes $\operatorname{Holant}(\mathcal{F} \cup \{[1,0],[0,1]\})$.

In this paper, we consider the parity version of Holant problems.

DEFINITION 2.3. Given a set of signatures \mathcal{F} , where each signature in \mathcal{F} takes values from $\mathbb{Z}_2 = \{0, 1\}$, we define the parity problem \oplus Holant(\mathcal{F}) as follows:

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

Output: Holant $_{\Omega}$ mod 2.

We also define \oplus Holant^c problems analogously.

DEFINITION 2.4. Given a set of signatures \mathcal{F} , we use \oplus Holant^c(\mathcal{F}) to denote \oplus Holant($\mathcal{F} \cup \{[0,1], [1,0]\}$).

Planar (parity) Holant problems are (parity) Holant problems on planar graphs.

2.2. Holographic reduction. The notions and reductions described in this subsection and the next are valid for general Holant problems regardless of the base field. In this paper we consider only the case in \mathbb{Z}_2 .

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 preserving the Holant value at the expense of a larger signature set. In particular, we replace each edge in the graph by a path of length 2 and assign to the new vertex the binary EqualITY signature $(=_2) = [1, 0, 1]$.

We use $\operatorname{Holant}(\mathcal{R} \mid \mathcal{G})$ to denote the Holant problem on bipartite graphs H = (U, V, E), where each signature for a vertex in U or V is from \mathcal{R} or \mathcal{G} , respectively.

An input instance for the bipartite Holant problem is a bipartite signature grid and is denoted as $\Omega = (H; \mathcal{R} | \mathcal{G}; \pi)$. Signatures in \mathcal{R} are considered as row vectors (or covariant tensors); signatures in \mathcal{G} are considered as column vectors (or contravariant tensors) [22].

Holographic transformations are expressed as linear operations on signatures, where for an *n*-argument function f the signature is a vector of length 2^n that expresses the value of f at all 2^n Boolean inputs. (The symmetric signature notation $[f_0, \ldots, f_n]$ is merely a length n + 1 abbreviation that is useful for discussing signatures in the symmetric case.)

For an *m*-by-*m* matrix *T*, the 2^m -by- 2^m matrix that is the tensor product of *T m* times is denoted by $T^{\otimes n}$. For a 2-by-2 matrix *T* and a signature set \mathcal{F} , define $T\mathcal{F} = \{g \mid \exists f \in \mathcal{F} \text{ of arity } n, g = T^{\otimes n}f\}$, and define $\mathcal{F}T$ similarly. Whenever we write $T^{\otimes n}f$ or $T\mathcal{F}$, we view the signatures as column vectors; similarly, for $fT^{\otimes n}$ or $\mathcal{F}T$ we view them as row vectors.

Let T be an invertible 2-by-2 matrix. The holographic transformation by T is the following operation: given a signature grid $\Omega = (H; \mathcal{R} \mid \mathcal{G}; \pi)$, for the same graph H, we get a new grid $\Omega' = (H; \mathcal{R}T \mid T^{-1}\mathcal{G}; \pi)$ by replacing each signature in \mathcal{R} or \mathcal{G} with the corresponding signature in $\mathcal{R}T$ or $T^{-1}\mathcal{G}$. This leads to the following result of Valiant.

THEOREM 2.5 (Holant theorem [40]). If there is a holographic transformation mapping signature grid Ω to Ω' , then $\operatorname{Holant}_{\Omega} = \operatorname{Holant}_{\Omega'}$.

Therefore, an invertible holographic transformation does not change the complexity of the Holant problem in the bipartite setting. Furthermore, there is a special kind of holographic transformation, the orthogonal transformation, that preserves binary equality and thus can be used freely in the general nonbipartite setting.

PROPOSITION 2.6 (Theorem 2.2 in [15]). Suppose T is a 2-by-2 orthogonal matrix $(TT^T = I_2)$, and let $\Omega = (H, \mathcal{F}, \pi)$ be a signature grid. Under a holographic transformation by T, we get a new grid $\Omega' = (H, T\mathcal{F}, \pi)$ and $\operatorname{Holant}_{\Omega} = \operatorname{Holant}_{\Omega'}$.

However, the only orthogonal matrices in $SL(2,\mathbb{Z}_2)$ are $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. These two matrices can provide little help because one is the identity and the other transforms signatures to their inverses. This proposition is more useful in the general setting, and we do need it when moving out of \mathbb{Z}_2 to prove the tractability for a special set of signatures.

2.3. Realization. One basic notion throughout the whole paper is realization. We say a signature f is realizable or constructable from a signature set \mathcal{F} if there is a gadget with some dangling edges such that each vertex is assigned a signature from \mathcal{F} , and the resulting graph, viewed as a black box signature with inputs on dangling edges, is exactly f. If f is realizable from \mathcal{F} , then we can freely add f into \mathcal{F} preserving the complexity.

Formally, such a notion is defined as an \mathcal{F} -gate [15, 16]. An \mathcal{F} -gate is a tuple (H, \mathcal{F}, π) , where H = (V, E, D) is a graph where the edge set consists of regular edges E and dangling edges D. The labeling π assigns a function from \mathcal{F} to each internal node. The dangling edges define external variables for the \mathcal{F} -gate. (See Figure 1 for an example.) We denote the regular edges in E by $1, 2, \ldots, m$ and denote the dangling edges in D by $m + 1, \ldots, m + n$. Then we can define a function Γ for this \mathcal{F} -gate as follows:

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



FIG. 1. An \mathcal{F} -gate with 5 dangling edges.

where $(y_1, y_2, \ldots, y_n) \in \{0, 1\}^n$ denotes an assignment on the dangling edges and $H(x_1, x_2, \ldots, x_m, y_1, y_2, \ldots, y_n)$ denotes the value of the signature grid on an assignment of all edges, which is the product of evaluations at all internal nodes. We will also call this function the signature Γ of the \mathcal{F} -gate. An \mathcal{F} -gate can be used in a signature grid as if it is just a single node with the particular signature.

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

We note that even for a very simple signature set \mathcal{F} , the signatures for all \mathcal{F} -gates could be quite complicated and expressive.

2.4. Some useful constructions and observations. First we mention two very simple kinds of gadget construction. The first is to connect two signatures via several edges. Let us say the two signatures are f and g. The connection edges are I_f from f and I_g from g. Rewrite the signature f as a matrix F where rows are indexed by input values from $\overline{I_f}$ and columns I_f , and rewrite g as a matrix G where rows are indexed by input values from I_g and columns $\overline{I_g}$. Notice that the index rules of the two matrices are opposites. The resulting signature in the matrix form is just FG.

Another simple gadget construction is to just put several signatures together and view them as a new signature, where all dangling edges are inputs edges of the new signature. The resulting signature is the tensor product of all component signatures. On the other hand, if a signature can be written as a tensor product of several signatures, then it can be decomposed into several smaller signatures. This observation leads to the notion of degeneracy. For symmetric signatures, the only possible degenerate case is that all component signatures are (the same) unary signatures.

DEFINITION 2.7. A signature is degenerate iff it is a tensor product of unary signatures.

For a signature $f = [f_0, f_1, \ldots, f_k]$ and any $0 \le l < h \le k$, we call $[f_l, f_{l+1}, \ldots, f_h]$ a subsignature of f. Note that with the help of the two unary signatures [0, 1] and [1, 0], any subsignature of a given signature is realizable. To see this, if we connect [0, 1] to f, then it is like forcing the value on the connecting edge to be 1, because otherwise the unary signature will contribute a multiplicative factor of 0. Hence, the resulting signature consists of the last k entries $[f_1, \ldots, f_k]$ of the original signature. Similarly, connecting [1, 0] would give us the first k entries. Repeatedly using these two unary signatures, we can get any successive entries we want from f, i.e., a subsignature. The crucial advantage of considering \oplus Holant^c problems is that we can freely use a subsignature of any signature from the set. This simplifies the needed case analysis.

All the signatures we consider here are in the Boolean domain. If we flip the 0 and 1 in the domain, a symmetric signature will be changed into its reverse, and the Holant values are the same. That is, the complexity of Holant problems for a set of signatures is the same as the complexity of Holant problems for the set composed by those signatures reversed. Another way to view this is by transforming the signature under the orthogonal matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. In this paper such an operation will be performed repeatedly.

2.5. Examples. Here we give several examples to help the reader understand Holant notation and the results of this paper.

Parity Matching Problem.

Input: A graph G = (V, E) and $V_0 \subseteq V$.

Output: Parity of the number of (partial) matchings that saturate all the vertices in V_0 .

Remark. In section 3.3, we show a polynomial algorithm for this problem. In particular, the parity of perfect matchings is a special case of this problem. Note that the corresponding counting problem is #P-hard [37]. In the signature language, this problem is \oplus Holant(\mathcal{F}), where all signatures in \mathcal{F} are either perfect matching signatures $[0, 1, 0, \ldots, 0]$ or partial matching signatures $[1, 1, 0, \ldots, 0]$ of arbitrary arities.

The algorithm for the parity matching problem can be generalized to any signature that is realizable by matchgates. This category of problems is one of the newly found tractable classes in this paper. The detailed characterization is in section 3.3.

In an instance of an unweighted Holant problem, the signature on each vertex can be viewed as a degree constraint. A certain degree is allowed if the corresponding entry of the signature is 1 and not allowed otherwise. The output is the total number of subsets of edges satisfying all the degree constraints. When a Holant problem involves only one signature, the input graph has to be regular, and the constraint is the same for all vertices.

 \oplus Holant({[1, 0, 0, 1, 0]}).

Input: A 4-regular graph G = (V, E).

Output: The parity of the number of subgraphs where every vertex is either isolated or of degree 3.

Remark. By Lemmas 4.6 and 7.2, this problem is \oplus P-complete. The corresponding counting problem is also #P-hard. On the other hand, if we do not allow isolated vertices in the object to count, the problem becomes \oplus Holant({[0,0,0,1,0]}), which is equivalent to the parity perfect matching problem on 4-regular graphs and is tractable.

The next problem is also defined by a single signature, but it is tractable because the signature is what we call a vanishing signature.

 \oplus Holant({[0, 0, 1, 1, 0]}).

Input: A 4-regular graph G = (V, E).

Output: The parity of the number of subgraphs where all vertices are of degree 2 or 3.

Remark. In section 6.3, we will show that this problem is tractable. The reason is that there is always an even number of subgraphs satisfying the degree constraint. Therefore the answer is always 0. The corresponding counting problem is #P-hard.

We call this kind of signature vanishing because the output is always 0. The vanishing signatures form the other newly found tractable problem family. The details are given in section 6.

3. Tractable families. We shall identify three tractable families for \oplus Holant^c problems. The first family, *affine signatures*, is adopted directly from the corresponding family for #CSP, where it is the only tractable class [20, 21]. The second family we derive from the *Fibonacci signatures*. For general counting problems, we also have a tractable family of Fibonacci signatures, but for parity problems, as we shall show, the family remains tractable even with the addition of the inversion signature [0, 1, 0]. This addition for general counting problems would give rise to #P-hardness. The third tractable family, *matchgate signatures*, is special to parity problems.

3.1. Affine signatures. Affine signatures correspond to simultaneous linear equations over \mathbb{Z}_2 and are defined as follows.

DEFINITION 3.1. A signature is affine iff its support is an affine space. We denote the set of all affine signatures by \mathscr{A} .

By definition, an affine signature can be viewed as a constraint defined by a set of linear equations. Viewing the edges as variables in \mathbb{Z}_2 , every assignment which contributes 1 in the summation corresponds to a solution which satisfies all the linear equations. Then the Holant value is exactly the number of solutions of the linear system, which can be computed in polynomial time.

THEOREM 3.2. If $\mathcal{F} \subseteq \mathscr{A}$, \oplus Holant^c(\mathcal{F}) is polynomial time computable.

Here we explicitly list all nondegenerate symmetric affine signatures.

LEMMA 3.3. Every nondegenerate symmetric affine signature is of one of the following forms:

- Equality signatures: [1, 0, 0, ..., 0, 1].
- Parity signatures: [1, 0, 1, 0, ..., 0/1] or [0, 1, 0, 1, ..., 0/1], where the last entry depends on whether the arity is odd or even.

3.2. Fibonacci signatures and [0, 1, 0]. The family of Fibonacci signatures was introduced in [13] to characterize a new family of holographic algorithms. It has played an important role in some previous dichotomy theorems [13, 15]. Formally, we have the following definition.

DEFINITION 3.4. A symmetric signature $[f_0, f_1, \ldots, f_n]$ is called a Fibonacci signature iff for $1 \le k \le n-2$ it is the case that $f_k + f_{k+1} = f_{k+2}$. We denote the set of all Fibonacci signatures by \mathscr{F} .

The Holant of a grid composed of Fibonacci signatures can be computed in polynomial time [13]. Its parity version is therefore also tractable. But here we shall show that the tractability still holds even if we extend the set with a signature [0, 1, 0], which is not a Fibonacci signature. This proof of tractability is based on the properties of Fibonacci signatures and a new observation on [0, 1, 0] as a parity signature.

Since we only care about the parity of the solutions, [0, 1, 0] can be replaced by the unsymmetrical signature (0, 1, -1, 0) in \mathbb{R} . (Note that here (0, 1, -1, 0) is not a symmetric signature. In fact, it is in the vector form rather than the abbreviated form of symmetric signatures.) This (0, 1, -1, 0) is a so-called 2-*realizable signature*, which has a special invariant property under holographic transformations [41, 11, 12]. This property plays an important role in the proof.

THEOREM 3.5. If $\mathcal{F} \subseteq \mathscr{F} \cup \{[0,1,0]\}, \oplus \text{Holant}^{c}(\mathcal{F}) \text{ is polynomial time com$ $putable.}$

In the proof of this theorem, we need to use the field of real numbers. This is the only place in this paper where a signature entry is viewed as a real number rather than as an entry in the field $\mathbb{Z}_2 = \{0, 1\}$.

Proof. As stated above, we replace [0,1,0] by the unsymmetrical signature (0,1,-1,0). We also replace the Fibonacci signatures in the field \mathbb{Z}_2 by real Fibonacci

signatures. For example, [1, 1, 0, 1] is replaced by [1, 1, 2, 3]. After the replacement, the parity of the Holant value does not change. For simplicity, we also denote the set of real Fibonacci signatures by \mathscr{F} . Next we show that $\operatorname{Holant}^{c}(\mathscr{F} \cup (0, 1, -1, 0))$ is computable in polynomial time.

For a Fibonacci signature $f = [f_0, f_1, \ldots, f_n]$ over the real numbers, we have $f_{k+2} = f_{k+1} + f_k$ for all $k = 0, 1, \ldots, n-2$. This is a second-order homogeneous linear recurrence relation. Thus we have $f_i = A\lambda_1^i + B\lambda_2^i$ for $i = 0, 1, \ldots, n$, where $\lambda_1 = (1 - \sqrt{5})/2$, $\lambda_2 = (1 + \sqrt{5})/2$ are the two roots of its characteristic polynomial $x^2 = x+1$, and A, B are two real numbers dependent on f_0 and f_1 . In tensor notation, we have $f = A(\frac{1}{\lambda_1})^{\otimes n} + B(\frac{1}{\lambda_2})^{\otimes n}$. One crucial point is that λ_1 and λ_2 are the same for all Fibonacci signatures (while A and B can vary for different signatures). Therefore we can do a holographic reduction as in Proposition 2.6 under the orthogonal matrix

$$T = \begin{pmatrix} \frac{1}{\sqrt{\lambda_1^2 + 1}} & \frac{\lambda_1}{\sqrt{\lambda_1^2 + 1}} \\ \frac{1}{\sqrt{\lambda_2^2 + 1}} & \frac{\lambda_2}{\sqrt{\lambda_2^2 + 1}} \end{pmatrix}$$

(We note that this is an orthogonal matrix because $\lambda_1 \lambda_2 = -1$.) This does not change the Holant value by Proposition 2.6. But all the Fibonacci signatures have a nicer format since

$$T^{\otimes n}f = T^{\otimes n}\left(A\begin{pmatrix}1\\\lambda_1\end{pmatrix}^{\otimes n} + B\begin{pmatrix}1\\\lambda_2\end{pmatrix}^{\otimes n}\right)$$
$$= AT^{\otimes n}\begin{pmatrix}1\\\lambda_1\end{pmatrix}^{\otimes n} + BT^{\otimes n}\begin{pmatrix}1\\\lambda_2\end{pmatrix}^{\otimes n}$$
$$= A\left(T\begin{pmatrix}1\\\lambda_1\end{pmatrix}\right)^{\otimes n} + B\left(T\begin{pmatrix}1\\\lambda_2\end{pmatrix}\right)^{\otimes n}$$
$$= A\left(1\begin{pmatrix}0\end{pmatrix}^{\otimes n} + B\begin{pmatrix}0\\1\end{pmatrix}^{\otimes n}$$
$$= [A, 0, \dots, 0, B].$$

For the signature (0, 1, -1, 0), it is easy to verify that $T^{\otimes 2}(0, 1, -1, 0)$ is $det(T) \cdot (0, 1, -1, 0)$ and det(T) = -1. By Proposition 2.6, the Holant value does not change after the orthogonal transformation. Afterward, all signatures are of the form $[A, 0, \ldots, 0, B]$ or (0, 1, -1, 0). For a signature of the form $[A, 0, \ldots, 0, B]$, any valuation which contributes nonzero to the Holant must have the same value on all its edges. For (0, 1, -1, 0) any such valuation must have the opposite values on its two edges. Thus, if the value of one edge in the graph is chosen, the nonzero contribution of its connected component is totally determined. Hence we can compute the Holant value of one component as a sum of at most two values. The Holant of the whole grid can be computed by multiplying the values of all connected components. \Box

Next, we shall list explicitly all the nondegenerate binary Fibonacci signatures. A Fibonacci signature is determined by its first two bits, by definition. These will be 00,01,10, or 11. However, the 00 case leads to the trivial signature, which is degenerate. Hence we have the following lemma.

LEMMA 3.6. Every nondegenerate Fibonacci signature modulo 2 is of one of the following forms:

- $[0, 1, 1, 0, 1, 1, \dots, 0/1],$
- $[1, 0, 1, 1, 0, 1, \dots, 0/1],$
- $[1, 1, 0, 1, 1, 0, \dots, 0/1].$

The following lemma regarding the constructability will be useful in the hardness proof later.

LEMMA 3.7. From any nondegenerate ternary Fibonacci signature [0, 1, 1, 0], [1, 0, 1, 1], or [1, 1, 0, 1] and the two unary signatures [1, 0] and [0, 1], one can realize every nondegenerate Fibonacci signature by some gadget.

Proof. By connecting any two Fibonacci signatures, we can get a longer Fibonacci signature [13]. We may connect several copies of the given ternary Fibonacci gates via the pattern in Figure 2 (just to connect them serially). It is easy to verify that the new Fibonacci signature is also nondegenerate. Therefore, we can realize a nondegenerate Fibonacci signature of arbitrary length. All three nondegenerate Fibonacci signature of arity k are subsignatures of any nondegenerate Fibonacci signature of arity k + 2. With the help of unary signatures [1,0] and [0,1], it is easy to get any subsignature of a given gate. This competes the proof.



FIG. 2. The gadget to construct longer Fibonacci signatures.

3.3. Matchgate signatures. Matchgates were introduced to simulate classically certain subclasses of quantum computations [39] and to be the basis of a class of holographic algorithms [42].

DEFINITION 3.8. A signature is called a matchgate signature iff it can be realized by a gadget, where each signature used in the gadget is a perfect matching signature [0, 1, 0, 0, ..., 0] or a partial matching signature [1, 1, 0, 0, ..., 0]. We denote the set of all matchgate signatures by \mathcal{M} .

We remark that the notion of matchgates we use here is in its most general sense: the graph can be either planar or nonplanar, and for each node we can insist on whether or not it has to be saturated by a matching edge. We shall now prove the following.

THEOREM 3.9. If $\mathcal{F} \subseteq \mathscr{M}$, \oplus Holant^c(\mathcal{F}) is polynomial time computable.

Since $\mathcal{F} \subseteq \mathscr{M}$ and we also have $[1,0], [0,1] \in \mathscr{M}$, the problem of $\oplus \text{Holant}^{c}(\mathcal{F})$ is a *parity matching problem* which we define as follows.

Parity Matching Problem.

Input: A graph G = (V, E) and $V_0 \subseteq V$.

Output: Parity of the number of (partial) matchings that saturate all the vertices in V_0 .

This general matching problem can be formalized as a summation of perfect matchings,

$$MatchingS(G) = \sum_{U \supseteq V_0} PM(G(U)),$$

where MatchingS(G) is the value we want to compute, PM(G) is the number of perfect matchings in G, and G(U) is the induced subgraph of G on vertex set U. The

transformation between a parity matching problem and a \oplus Holant^c(\mathscr{M}) problem is simple. The graph is the same. All the vertices in V_0 have perfect matching signatures, and all the other vertices have partial matching signatures.

Before we give the algorithm for the Parity Matching Problem, we need to introduce the definition of the Pfaffian. The Pfaffian of an $n \times n$ skew-symmetric matrix A is defined to be zero if n is odd, one if n = 0, and if n is even with n = 2k and k > 0, then it is defined as

$$Pf(A) = \sum_{\pi} \epsilon_{\pi} A(i_1, i_2) A(i_3, i_4) \cdots A(i_{n-1}, i_n),$$

where

1. $\pi = [i_1, i_2, \dots, i_n]$ is a permutation on $[1, 2, \dots, n]$;

2. summation is over all such permutations π , where further $i_1 < i_2, i_3 < i_4, \ldots, i_{2k-1} < i_{2k}$ and $i_1 < i_3 < \cdots < i_{2k-1}$;

3. $\epsilon_{\pi} \in \{1, -1\}$ is the sign of the permutation.

The following fact, due to Cayley [17] (see also [3] Theorem 9.5.2), relates the Pfaffian to the determinant.

THEOREM 3.10. For any $2k \times 2k$ skew-symmetric matrix A,

$$Det(A) = (Pf(A))^2$$

In the field \mathbb{Z}_2 , we have x = -x, and hence a skew-symmetric matrix is indeed a symmetric matrix. Moreover the sign $\epsilon_{\pi} = 1 = -1$ in \mathbb{Z}_2 can be ignored. Let Abe the adjacency matrix of a graph G, (i.e., the nonzero elements are $A_{i,j} = A_{j,i} = 1$ for $\{i, j\} \in E$). Then each monomial in the Pfaffian corresponds to a distinct perfect matching in G. Therefore, Pf(A) is exactly the parity of the number of perfect matchings in G. We have

(1)
$$\operatorname{PM}(G) = \operatorname{Pf}(A) = (\operatorname{Pf}(A))^2 = \operatorname{Det}(A) \pmod{2},$$

where G has an even number of vertices. So the parity of the number of perfect matchings can be computed in polynomial time. Next we show that this tractability can be extended to partial matchings. We do this through the *Pfaffian sum theorem* [39].

For any $n \times n$ matrix A we call a set $I = \{i_1, i_2, \ldots, i_r\} \subseteq [n]$ an index set. Further we denote by A(I) the $r \times r$ submatrix of A on rows and columns in I.

DEFINITION 3.11. The Pfaffian sum of an $n \times n$ matrix A is a polynomial over indeterminates $\lambda_1, \lambda_2, \ldots, \lambda_n$ such that

$$\operatorname{PfS}(A) = \sum_{I \subseteq [n]} \left(\prod_{i \notin I} \lambda_i \right) \operatorname{Pf}(A(I)).$$

The summation here is over the various principal minors obtained from A by restricting the indices to some subset $I \subseteq [n]$.

In this paper we need only the instances in which each λ_i is either 0 or 1. For a given nonomittable vertex set V_0 , we can define the character vector $\vec{\lambda} = (\lambda_1, \ldots, \lambda_n)$ as follows: for each $i, \lambda_i = 0$ iff $i \in V_0$. Thus, in this case the Pfaffian sum of the character vector $\vec{\lambda}$ is simply the sum of the Pf(A(I)) over those I that contain all the nonomittable indices.

We define the $n \times n$ matrix $\Lambda^{(n)}$ as follows:

$$\Lambda^{(n)}(i,j) = \begin{cases} (-1)^{j-i+1} \lambda_i \lambda_j & \text{if } i < j, \\ (-1)^{i-j} \lambda_i \lambda_j & \text{if } i > j, \\ 0 & \text{if } i = j. \end{cases}$$

Also for an $n \times n$ matrix A we define A^+ to be the $(n+1) \times (n+1)$ matrix of which the first n rows and columns equal A itself, and the (n+1)st row and column entries are all zero.

The following theorem, which relates the Pfaffian sum to a single Pfaffian, was proved in [39].

THEOREM 3.12. For an $n \times n$ skew-symmetric matrix A and indeterminates $\lambda_1, \ldots, \lambda_{n+1}$,

$$PfS(A) = \begin{cases} Pf(A + \Lambda^{(n)}) & \text{if } n \text{ is even,} \\ Pf(A^+ + \Lambda^{(n+1)}) & \text{with } \lambda_{n+1} = 1 \text{ if } n \text{ is odd.} \end{cases}$$

Thus, a Pfaffian sum can be computed in polynomial time. The relation (1) between perfect matchings and Pfaffians can be therefore extended to one between matchings and Pfaffian sums:

(2) MatchingS(G) =
$$\sum_{U \supseteq V_0} PM(G(U)) = \sum_{U \supseteq V_0} Pf(A(U)) = PfS(A)(\vec{\lambda}) \pmod{2}.$$

This relation gives a polynomial time algorithm for the Parity Matching Problem and completes the proof of Theorem 3.9.

Now we go on to list explicitly all the nondegenerate symmetric matchgate signatures. Useful matchgate identities in [8] is an essential tool for characterizing the realizability of matchgates. For completeness we quote the identities as follows.

A pattern α is an *m*-bit string, i.e., $\alpha \in \{0,1\}^m$. A position vector $P = \{p_i\}$, $i \in [l]$, is a subsequence of $\{1, 2, \ldots, m\}$, i.e., $p_i \in [m]$ and $p_1 < p_2 < \cdots < p_l$. It can also be viewed as an *m*-bit string, whose (p_1, p_2, \ldots, p_l) th bits are 1, and the others are 0. Let $e_i \in \{0,1\}^m$ be the pattern with 1 in the *i*th bit and 0 elsewhere. Let $\alpha + \beta$ denote the bitwise XOR of the patterns α and β .

PROPOSITION 3.13 (matchgate identities for signatures). For a signature F realizable by perfect matching gates, for any pattern $\alpha \in \{0,1\}^m$, any l ($0 < l \le m$), and any position vector $P = \{p_i\}, i \in [l]$, the following identity holds:

(3)
$$\sum_{i=1}^{l} (-1)^{i} F(\alpha + e_{p_{i}}) F(\alpha + p + e_{p_{i}}) = 0$$

LEMMA 3.14. Every nondegenerate symmetric matchgate signature is of one of the following forms:

- Perfect matching signatures: $[0, 1, 0, 0, \dots, 0]$ or $[0, 0, \dots, 0, 1, 0]$.
- Partial matching signatures: $[1, 1, 0, 0, \dots, 0]$ or $[0, 0, \dots, 0, 1, 1]$.
- Parity signatures: $[1, 0, 1, 0, \dots, 0/1]$ or $[0, 1, 0, 1, \dots, 0/1]$.

Proof. We first prove that every nondegenerate symmetric signature that is realizable from perfect and partial matching signatures is of one of the forms claimed in the lemma. The definition of matchgate signatures here is slightly different from that in [39, 38, 8] since we do not require the gadget to be planar [42] and do not use modifiers [39]. However, since $1 \equiv -1 \pmod{2}$, our signatures here are equivalent to those with modifiers. Also, matchgate signatures in the field \mathbb{Z}_2 also satisfy the *useful matchgate identities* above if all the signatures involved are perfect matching signatures.

If some of the nodes used in the constructing gadget are omittable (i.e., they have partial matching signatures), we transform the instance into a perfect matchings instance using Theorem 3.12 as follows. We add edges between every pair of partial matching nodes. If there is an odd number of nodes in total, we add an additional node and connect it also to all the partial matching nodes. By Theorem 3.12, after the transformation, we can compute the parity of the number of matchings in the original graph as a parity of the perfect matchings in the modified graph.

To deal with general matchgate signatures, we need the Holant value after the removal of various subsets of the omittable nodes. After removing some nodes, the parity of the number of the remaining nodes may change. As discussed above, it is only when there is an odd number of nodes remaining that we need to add an additional node. Here we slightly change our perspective. We view this additional node also as a special omittable external node. If there is an odd number of nodes remaining, we view this external node as a remaining node as well; otherwise if there is an even number of nodes remaining, we view it as deleted. If f is the signature for the original matchgate and g is the signature for the corresponding matchgate after the transformation, then $g(\alpha, \oplus \alpha) = f(\alpha)$ holds, where $\oplus \alpha$ is 0 or 1 according to the parity $\oplus_{i=1,2,...,m} \alpha_i$. This g has to satisfy the above matchgate identities (3).

In the following, we assume that the original graph has an odd number of nodes. The case of an even number is similar. Let $f = [f_0, f_1, \ldots, f_m]$ be the signature of the original matchgate, and let g be the signature after the transformation and addition of the external node.

For m = 2, all the nondegenerate symmetric signatures are of the forms claimed in the lemma.

We now consider $m \ge 3$ and apply the matchgate identities (3) to the symmetric signature f. Consider the pattern $100\alpha 0$, where α has Hamming weight 2i and $0 \le 2i \le m-3$. Let the position vector be $111000\cdots 01$. Then (3) gives

 $0 = g(000\alpha 0)g(111\alpha 1) - g(110\alpha 0)g(001\alpha 1) + g(101\alpha 0)g(010\alpha 1) - g(100\alpha 1)g(011\alpha 0).$

Translating back to f, we get

$$0 = f_{2i}f_{2i+3} - f_{2i+2}f_{2i+1} + f_{2i+2}f_{2i+1} - f_{2i+1}f_{2i+2} = f_{2i}f_{2i+3} - f_{2i+2}f_{2i+1}$$

For m = 3, this $f_0 f_3 = f_1 f_2$ is the only identity.

For $m \ge 4$, we use the matchgate identities (3) again. Consider the pattern $1000\alpha \oplus \alpha$ where α has Hamming weight i and $0 \le i \le m-4$. Let the position vector be $111100 \cdots 0$. Then (3) gives

$$\begin{split} 0 &= g(0000\alpha, \oplus \alpha)g(1111\alpha, \oplus \alpha) - g(1100\alpha, \oplus \alpha)g(0011\alpha, \oplus \alpha) \\ &+ g(1010\alpha, \oplus \alpha)g(0101\alpha, \oplus \alpha) - g(1001\alpha, \oplus \alpha)g(0110\alpha, \oplus \alpha). \end{split}$$

Translating back to f, we get

$$0 = f_i f_{i+4} - f_{i+2} f_{i+2} + f_{i+2} f_{i+2} - f_{i+2} f_{i+2} = f_i f_{i+4} - f_{i+2} f_{i+2}.$$

Consider the pattern 10^m and the position vector $1^m \oplus (m)$, where $\oplus (m)$ is the parity of m. Then we have

$$0 = g(0^{m+1})g(1^m \oplus (m)) - g(110^{m-1})g(001^{m-2} \oplus (m)) + g(1010^{m-2})g(0101^{m-3} \oplus (m)) - g(10010^{m-3})g(01101^{m-4} \oplus (m)) \pm \cdots$$

The terms cancel except for the first two. Translating into f, we get $f_0 f_m = f_2 f_{m-2}$.

Similarly, consider the pattern 10^m and the position vector $1^{m-1}0 \oplus (m-1)$. Then we have

$$0 = g(0^{m+1})g(1^{m-1}0 \oplus (m-1)) - g(110^{m-1})g(001^{m-3}0 \oplus (m-1)) + g(1010^{m-2})g(0101^{m-4}0 \oplus (m-1)) - g(10010^{m-3})g(01101^{m-5}0 \oplus (m-1)) \pm \cdots$$

The terms cancel except for the first two. Translating into f, we get $f_0 f_{m-1} = f_2 f_{m-3}$. Similarly, we can also get $f_1 f_m = f_3 f_{m-2}$ and $f_1 f_{m-1} = f_3 f_{m-3}$.

These relations imply that the subsequence of the signature for even (or odd) indices is a geometric sequence. In this field of \mathbb{Z}_2 , there are only four types of geometric sequences. They are

1. $[0, 0, \ldots, 0],$

- 2. $[1, 0, 0, \dots, 0],$
- 3. $[0, 0, \ldots, 0, 1],$
- 4. $[1, 1, \ldots, 1]$.

There are $4 \times 4 = 16$ possible combinations for the even subsequence and the odd subsequence. We use type (i, j) to denote the sequence whose odd subsequence is of type i and even subsequence is of type j. Types (1, 1), (4, 4) are degenerate. Types (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (3, 1), (3, 3), (4, 1) are listed in the lemma. We need only rule out the remaining six types (2, 3), (2, 4), (3, 2), (3, 4), (4, 2), (4, 3). For (2, 4), the first four entries are [1, 1, 0, 1], which does not satisfy $f_{0}f_{3} = f_{1}f_{2}$. This matchgate identity also rules out (4, 2), whose first four entries are [1, 1, 1, 0]. For (3, 4) and (4, 3), their last four entries do not satisfy $f_{m-3}f_m = f_{m-1}f_{m-2}$. For (2, 3), it has form $[1, 0, \ldots, 0, 1]$ or $[1, 0, \ldots, 0, 1, 0]$. It violates either $f_{0}f_m = f_{2}f_{m-2}$ or $f_{0}f_{m-1} = f_{2}f_{m-3}$. For (3, 2), we can similarly argue that it violates either $f_{1}f_m = f_{3}f_{m-2}$ or $f_{1}f_{m-1} = f_{3}f_{m-3}$.

This completes the first part of the proof, namely, that all the realizable matchgate signatures are of one of the claimed forms. The realizability of these signatures as matchgate signatures will follow from Lemma 3.16. \Box

Next we prove some realizability properties regarding symmetric matchgate signatures, which will be used in the proof of the dichotomy for the parity Holant^c problem.

LEMMA 3.15. Every parity signature can be realized by the signatures [0, 1], [1, 0], and [0, 1, 0, 1] (or [1, 0, 1, 0]).

Proof. We may connect instances of the signature [0, 1, 0, 1] (or [1, 0, 1, 0]) to get an arbitrarily long signature using the pattern shown in Figure 3. Note that in general such a gadget will not result in a symmetric signature. However, in the situation here it is not hard to check that the resulting signature is indeed symmetric. In fact it is either $[1, 0, 1, 0, \ldots, 0/1]$ or $[0, 1, 0, 1, \ldots, 0/1]$, depending on the arity as well as on which of [0, 1, 0, 1] or [1, 0, 1, 0] you put in the gadget. Every parity signature is a subsignature of such a gate. \square

LEMMA 3.16. Every parity signature or perfect matching signature can be realized by the signatures [0, 1], [1, 0], and [0, 1, 0, 0] (or [0, 0, 1, 0]).



FIG. 3. The gadget for [0, 1, 0, 1, 0, 1].

FIG. 4. The triangle gadget for [0, 1, 0, 1].

Proof. By symmetry, we need only prove the lemma for [0, 1, 0, 0]. First we observe that if we place [0, 1, 0, 0] at every vertex in the triangle gadget shown in Figure 4, the resulting signature is [0, 1, 0, 1]. Then by Lemma 3.15, every parity signature can be constructed. By connecting one [1,0] to [0, 1, 0, 0], we can get [0, 1, 0]. Then we place [0, 1, 0, 0] at A and B and [0, 1, 0] at C in Figure 5. It is similar to the case of the gadget in Figure 3 that in general the resulting signature is not necessarily symmetric, but for the particular signatures used here in the modulo 2 setting, the resulting signature, [0, 1, 0, 0, 0], is indeed symmetric. Similarly, we may always connect $[0, 1, 0, \ldots, 0]$ of arity k with [0, 1, 0, 0] via [0, 1, 0], and the resulting signature is $[0, 1, 0, \ldots, 0]$ of arity k+1. Therefore we can also construct every perfect matching signature. This completes the proof.



FIG. 5. The gadget for [0, 1, 0, 0, 0] and [1, 1, 0, 0, 0].

LEMMA 3.17. Every matchgate signature can be realized by the signatures [0, 1], [1, 0], and [1, 1, 0, 0] (or [0, 0, 1, 1]).

Proof. By symmetry, we need only prove the lemma for [1,1,0,0]. Note that we can get [1,1] from [1,1,0,0] by connecting it with two unary signatures [1,0]. By connecting [1,1] with [1,1,0,0], we can get [0,1,0]. Then we place [1,1,0,0] at A and B and [0,1,0] at C in Figure 5. As in the proof of Lemma 3.16, we can get a symmetric signature [0,1,0,0,0]. Thus, by Lemma 3.16, any perfect matching signature is constructable from that. Similarly, we may always connect $[0,1,0,\ldots,0]$ of arity k with [1,1,0,0] via [0,1,0], and the resulting signature is $[1,1,0,\ldots,0]$ of arity k + 1. In this way, we can construct every partial matching signature. By definition, we can further construct all matchgate signatures.

Remark. We note that every degenerate signature is a member of each of the above three tractable families. For affine signatures and matchgate signatures, this can be easily verified from their definition. For Fibonacci signatures, we note that our Definition 3.4 as stated is for *nondecomposable* signatures. A signature is *non*-

decomposable iff it cannot be written as a tensor product of two signatures of strictly smaller arities. In this sense $[1, 1, 0, 1] \otimes [0, 1, 1, 0]$ is also considered to be a Fibonacci signature. All the properties stated here are also valid for this extended notion of Fibonacci signatures. In this sense, it is clear that all the degenerate signatures are Fibonacci signatures since they can be decomposed to unary signatures, which are Fibonacci by definition.

4. Some hardness results. In this section, we prove several hardness results regarding \oplus Holant^c problems. These results are preparations for proving dichotomy theorems. Many of them are for low arity signature sets. As we will see, to handle general cases it will come down to lower arity cases.

4.1. An initial hard problem. As the starting point of everything, we first consider the problem of \oplus Pl-Rtw-Mon-3CNF. Pl-Rtw-Mon-3CNF is a special case of the satisfying problem for 3CNF formulae. "PL" means it is restricted to planar graphs. "Rtw" (read twice) means that every variable only appears twice in clauses. "Mon" (monotone) means that for every variable only itself or its negation appears, but not both. Then w.l.o.g. we can assume all variables appear in the positive form. \oplus Pl-Rtw-Mon-3CNF is the parity version of it.

To transform \oplus Pl-Rtw-Mon-3CNF into the Holant setting, we can use vertices to represent all clauses and variables. We draw an edge between a clause vertex and a variable vertex if that variable appears in the clause. Due to the restrictions Rtw and 3CNF, the resulting graph is a 2-3 bipartite graph. Moreover, a variable that appears only positively can be viewed as the signature [1, 0, 1], which means it can be absorbed. The signature on each clause vertex is [0, 1, 1, 1] since it is a clause of a CNF formula. So, in the end, this problem is translated into planar \oplus Holant([0, 1, 1, 1]) in the Holant language.

In [41], Valiant showed that \oplus Pl-Rtw-Mon-3CNF is \oplus P-hard.

THEOREM 4.1. $Planar \oplus Holant([0, 1, 1, 1])$ and equivalently planar $\oplus Holant([1, 1, 1, 0])$ are $\oplus P$ -complete.

Remark. All hardness results in this paper for \oplus Holant^c, but not for \oplus Holant, will hold even if we restrict the input to planar graphs. This is because the above starting point is true for planar graphs, and all the gadgets used in those reductions are also planar.

This \oplus Holant([0, 1, 1, 1]) can also be viewed as \oplus Holant([1, 0, 1]|[0, 1, 1, 1]) in the sense that any instance expressible as one can also be expressed as an instance of the other. In fact, as noted in section 2, one can freely insert a new vertex on any edge by assigning an EQUALITY₂ signature on it, and freely "absorb" any degree 2 vertex with an EQUALITY₂.

We now show that the holographic transformation $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, the Holant value of \oplus Holant([1, 0, 1]|[0, 1, 1, 1]), is the same as that of \oplus Holant ([1, 1, 0]|[1, 0, 0, 1]). Here we provide some details to illustrate how the holographic transformation works. For brevity, in later proofs we will not repeat this. We have G = [1, 0, 1], R = [0, 1, 1, 1], and the transformation matrix $T = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$. The resulting signatures G' and R' should satisfy that $G' = T^{\otimes 2}G$ and $R = R'T^{\otimes 3}$, so $R' = R(T^{-1})^{\otimes 3}$. From the first equation, we have

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 2 \end{pmatrix} \equiv \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \pmod{2}$$

so that G' = [1, 1, 0]. Note that $T^{-1} \equiv T \pmod{2}$. Similarly, from the latter equation, we have

$$\begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} = (7 \quad 4 \quad 4 \quad 2 \quad 4 \quad 2 \quad 2 \quad 1) + (1 + 1) +$$

Notice that the resulting signature $[7, 4, 2, 1] \equiv [1, 0, 0, 1] \pmod{2}$. Thus, the transformed signatures are $\{[1, 1, 0] | [1, 0, 0, 1]\}$.

COROLLARY 4.2. \oplus Holant([1,1,0]|[1,0,0,1]) and, equivalently, \oplus Holant([0,1,1]| [1,0,0,1]), are \oplus P-complete.

Remark. In fact, the result above is also shown in [41]. The problem \oplus Holant ([1,1,0]| [1,0,0,1]) is called \oplus Pl-3/2-Bip-VC and is further equivalent to \oplus Holant ([1,1,0]| [1,1,1,0]) under holographic transformations. This will be illustrated later in section 5.2.

4.2. More hardness results. Next we establish some further hardness results for \oplus Holant^c problems. First we give a quick generalization of Corollary 4.2.

COROLLARY 4.3. \oplus Holant^c([0,1,1], [1,0,...,0,1]) is \oplus *P*-complete, as long as the number of 0's is at least 2.

Proof. Assume we have k 0's in the middle of the signature. If k = 2, by Corollary 4.2, \oplus Holant([0, 1, 1], [1, 0, 0, 1]) is \oplus P-complete. Otherwise we show how to construct [1, 0, 0, 1] from the longer equality signature. We can get the unary signature [1, 1] by connecting [0, 1] with [0, 1, 1]. If we connect [1, 1] with [1, 0, ..., 0, 1], the resulting signature is [1, 0, ..., 0, 1], which has one fewer 0. Thus, we may connect k - 2 copies of [1, 1] with [1, 0, ..., 0, 1] to get [1, 0, 0, 1].

The following results deal with the case when the signature set contains both matchgates and Fibonacci signatures. We first show a base case and then reduce the general case to it.

LEMMA 4.4. \oplus Holant^c([0, 1, 0, 1, 0], [0, 1, 1, 0]) is \oplus *P*-complete.

Proof. First we observe that if we place [0,1,0,0,0] at vertices A, B, C and [0,1,1,0] at vertex O in Figure 6, the resulting signature is [1,1,1,0]. Thus, given a signature grid composed of [1,1,1,0]'s, we can construct a signature grid composed of [0,1,0,0,0]'s and [0,1,1,0]'s having the same value. By Theorem 4.1, \oplus Holant([1,1,1,0]) is \oplus P-complete. Therefore \oplus Holant^c([0,1,0,0,0], [0,1,1,0]) is \oplus P-complete.

Next we show that this can be further simulated by [0, 1, 0, 1, 0] and [0, 1, 1, 0]. We may treat this grid of $\{[0, 1, 0, 0, 0], [0, 1, 1, 0]\}$ as a bipartite graph by replacing every edge with a signature [1, 0, 1], and then performing the holographic transformation $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. [1, 0, 1] will be transformed into [0, 1, 1], and [0, 1, 0, 0, 0] into [0, 1, 0, 1, 0], while [0, 1, 1, 0] remains unchanged. This implies that \oplus Holant^c([0, 1, 0, 1, 0], [0, 1, 1, 0]) is \oplus P-complete, as [0, 1, 1] is easily constructed from [0, 1, 1, 0] and [1, 0].

COROLLARY 4.5. If \mathcal{F} contains a nondegenerate symmetric signature in \mathcal{M} and a nondegenerate Fibonacci signature, both of which have arity at least 3, then \oplus Holant^c(\mathcal{F}) is \oplus P-complete.



FIG. 6. The gadget to construct [1, 1, 1, 0].

FIG. 7. The unsymmetrical gadget.

Proof. By Lemma 3.14, any nondegenerate symmetric signature in \mathcal{M} is a perfect matching signature, a partial matching signature, or a parity signature. Then by Lemmas 3.15, 3.16, and 3.17, we can always construct all the parity signatures, including [0, 1, 0, 1, 0], from a nondegenerate matchgate signature with arity at least 3 and two unary signatures [1, 0] and [0, 1].

By Lemma 3.7, we can construct all the Fibonacci signatures, including [0, 1, 1, 0], from a nondegenerate Fibonacci signature with arity at least 3 and unary signatures [1, 0] and [0, 1].

To sum up, we can construct both [0, 1, 0, 1, 0] and [0, 1, 1, 0] from $\mathcal{F} \cup \{[1, 0], [0, 1]\}$. Thus, by Lemma. 4.4, \oplus Holant^c(\mathcal{F}) is \oplus P-complete.

This result implies that simultaneous occurrence of matchgates and Fibonacci signatures leads to \oplus P-completeness. Similarly, we have the following lemma, which shows that the simultaneous occurrence of matching signatures and equality signatures also leads to \oplus P-completeness.

In the following proof we will use an interesting gadget construction technique, which we call *clustering*. The gadget we use is not symmetric, as in Figure 7. Thus, when plugging in signatures, the resulting signature is not guaranteed to be symmetric. However, if we cluster its edges pairwise (as shown in Figure 7) and guarantee that the values of both edges are forced to be the same, then it will behave like a symmetric signature of half the arity. To force that is equivalent to saying that if the values of the two edges are different, then the total Holant value is 0. To do this, we usually work in a bipartite graph, construct a long equality signature on the other side, and cluster its edges to use it as an equality of half the arity. To get a nonzero value out of an equality gate, all values must be equal. Hence, in particular, every pair of clustered edges has the same value.

LEMMA 4.6. The parity problems \oplus Holant^c([0, 0, 1, 0], [1, 0, 0, ..., 0, 1]), \oplus Holant^c([0, 1, 0, 0], [1, 0, 0, ..., 0, 1]), \oplus Holant^c([0, 0, 1, 1], [1, 0, 0, ..., 0, 1]), and \oplus Holant^c([1, 1, 0, 0], [1, 0, 0, ..., 0, 1]) are all \oplus P-complete if the arity of the equality signature is at least 3.

Proof. By symmetry, we need only prove the lemma for [0, 1, 0, 0] and [1, 1, 0, 0]. By Lemma 3.17, we can construct [0, 1, 0, 0] from [1, 1, 0, 0]. So it is sufficient to prove that \oplus Holant^c($[0, 1, 0, 0], [1, 0, 0, \dots, 0, 1]$) is \oplus P-complete.

First we reduce the arity of an equality gate of arity at least 5 by connecting an arbitrary pair of its dangling edges. Eventually it will become an equality gate of arity 3 or 4 depending on the parity of the original gate's arity. For either case we can realize an equality gate [1, 0, 0, 0, 0, 0, 1] of arity 6. It will be used as the long equality gate in the clustering as a shorter [1, 0, 0, 1].

On the other hand, we connect one edge of two [0, 1, 0, 0] gates as in Figure 7.

This is the unsymmetric gadget we use. As mentioned above for clustering, under the guarantee this gate behaves like [1,1,0]. With the arity 3 equality gate, we can simulate \oplus Holant([1,1,0]|[1,0,0,1]). Thus, by Corollary 4.2, \oplus Holant^c([0,1,0,0], $[1,0,0,\ldots,0,1]$) is \oplus P-complete. This completes the proof. \square

This lemma implies the following direct corollary for signatures that contain both equality and matching signatures as subsignatures.

COROLLARY 4.7. \oplus Holant^c([1,0,...,0,1,0]) and \oplus Holant^c([1,0,...,0,1,1]) are \oplus *P*-complete, as long as the number of 0's is at least 2.

Finally, there are still two more special cases not covered by the above. They can be treated in more or less the same manner. The clustering is used again.

LEMMA 4.8. \oplus Holant^c([0, 0, 1, 0, 0]) and \oplus Holant^c([0, 0, 1, 0, 1]) are \oplus *P*-complete. *Proof.* We show this claim by the same technique as that employed in the proof of

Lemma 4.6. Note that [0, 0, 1, 0] is a subsignature of both [0, 0, 1, 0, 0] and [0, 0, 1, 0, 1]. We need only construct the arity 6 equality signature [1, 0, 0, 0, 0, 0, 1].

If we place [0,0,1,0,0] at every vertex in the gadget shown in Figure 8, the resulting signature is [1,0,0,0,1]. Connecting one edge of two [1,0,0,0,1] gates, we can get the gate [1,0,0,0,0,0,1].

The case of [0, 0, 1, 0, 1] is more complex. We can get [1, 0, 1, 0, 0] by connecting a [0,1,0] at each edge of [0, 0, 1, 0, 1]. Then we place this [1, 0, 1, 0, 0] at B and [0, 0, 1, 0, 1] at A in the gadget in Figure 8. The combined gadget is depicted in Figure 9, where the signatures of nodes C_1, C_2, C_3 , and C_4 are [0, 1, 0]. The resulting signature is again [1, 0, 0, 0, 1]. The remaining proof is the same as that of [0, 0, 1, 0, 0].

We need to mention that the gadget in Figure 8 does not necessarily simulate a symmetric signature, but it does in the modulo 2 setting with the specific signatures that we put in the proof above.

 C_3

 C_4



FIG. 8. The gadget to construct [1, 0, 0, 0, 1] from [0, 0, 1, 0, 0].

FIG. 9. The gadget to construct [1, 0, 0, 0, 1] from [0, 0, 1, 0, 1].

 C_2

 C_1

A

5. Two dichotomy theorems. Based on the algorithms in section 3 and the hardness results in section 4, we can show the dichotomy theorem for \oplus Holant^c problems, which is a stepping stone toward the final dichotomy.

As an aside, we also show a dichotomy theorem for the 2-3 bipartite regular graphs with a single signature on either side. This case is not covered by the general dichotomy because the underlying graph has a more restricted structure.

5.1. Dichotomy for \oplus **Holant**^{*c*} **problems.** At first we mention some normalization of the signature set \mathcal{F} . Any symmetric degenerate signature is of the form $[x, y]^{\otimes k}$. It can be replaced by the corresponding unary signature [x, y] without changing the complexity of the problem. Hence, we always assume that every signature in \mathcal{F} of arity greater than 1 is nondegenerate.

Also as we mentioned before, for a signature f and its inverse f^{-1} , their Holant values for any graph are the same. Thus, in the following proofs we will often ignore the reversal case.

If there is a single signature that does not belong to the three tractable sets, then the \oplus Holant problem is hard. Formally, we have the following lemma.

LEMMA 5.1. For a signature set \mathcal{F} , if there exists a single signature $f \in \mathcal{F}$ such that $f \notin \mathscr{A} \cup \mathscr{M} \cup (\mathscr{F} \cup \{[0,1,0]\})$, then the parity problem $\oplus \operatorname{Holant}^{c}(\mathcal{F})$ is $\oplus P$ -complete.

The proof is a case-by-case analysis. Basically our discussion is in terms of the maximum number of consecutive "0" bits and then that of "1" bits in its symmetric form. But before that we will take some precautions.

Proof. First we notice that \mathcal{M} contains all signatures with arity less than or equal to two. Thus, the arity of f is at least three.

Then we rule out some patterns that will later appear more than once. Assume fcontains [0, 1, 1, 0], [1, 0, 1, 1], or [1, 1, 0, 1] as a subsignature. Because $f \notin \mathscr{F}$, it must extend that subsignature in either one or both of the directions. Thus f must contain [0, 1, 1, 0, 0], [1, 0, 1, 1, 1], [1, 1, 0, 1, 0] or their reversals as a subsignature.

In fact, any of them would lead to \oplus P-completeness as follows:

- For [0, 1, 1, 0, 0], \oplus Holant^c(\mathcal{F}) is \oplus P-complete by Corollary 4.5 since it contains both the Fibonacci signature [0, 1, 1, 0] and the matchgate signature [1, 1, 0, 0] as subsignatures.
- For [1,0,1,1,1], \oplus Holant^c(\mathcal{F}) is \oplus P-complete by Theorem 4.1 since it contains [0, 1, 1, 1] as its subsignature.
- For [1, 1, 0, 1, 0], \oplus Holant^c(\mathcal{F}) is \oplus P-complete by Corollary 4.5 since it contains both the Fibonacci signature [1, 1, 0, 1] and the matchgate signature [1, 0, 1, 0] as its subsignatures.

Next we consider the maximum number of consecutive 0 bits of f in its symmetric form. First we assume f contains at least two consecutive 0's. Then consider a sequence of consecutive 0's of the maximum length k_0 in f. If both ends of this sequence are 1, f must contain a subsignature of the form $[1, 0, \ldots, 0, 1, 0]$, $[1, 0, \ldots, 0, 1, 1]$ or their reversals, because otherwise f is an equality signature $[1, 0, \ldots, 0, 1] \in \mathscr{A}$. Then by Corollary 4.7, \oplus Holant^c(\mathcal{F}) is \oplus P-complete.

Otherwise, w.l.o.g., we may assume that the first k_0 bits of f are 0. Then we consider the number of subsequent 1's after these to be 0. It must be one of the following three cases:

- If there are more than three 1's, then we have [0, 1, 1, 1] as its subsignature, and by Theorem 4.1 we are done.
- If there are just two 1's, f cannot end here because the partial matching gate $[0, \ldots, 0, 1, 1]$ is in \mathcal{M} . Then we have [0, 0, 1, 1, 0] as a subsignature. It contains [0, 1, 1, 0], and this has been discussed above.
- If there is only one 1, f cannot end because $[0, \ldots, 0, 1]$ is degenerate. Also because $[0,\ldots,0,1,0]$ is in \mathcal{M} , f must be of the form $[0,\ldots,0,1,0,0]$ or $[0, \ldots, 0, 1, 0, 1]$. By Lemma 4.8 both cases are \oplus P-complete.

Now we deal with the case that f contains no two consecutive 0's. Consider the maximum number k_1 of consecutive 1's in f. If $k_1 \ge 3$, f must contain [0, 1, 1, 1] or its reversal, and we get \oplus P-completeness by Theorem 4.1. If $k_1 = 1$, f must be a parity signature which is impossible. So we have $k_1 = 2$. But in that case f must contain a Fibonacci signature [0, 1, 1, 0], [1, 0, 1, 1], or [1, 1, 0, 1] as its subsignature, which is already shown to imply $\oplus P$ -completeness. П

In order to show our dichotomy, the case left is that $\mathcal{F} \in \mathcal{A} \cup \mathcal{M} \cup (\mathcal{F} \cup \{[0, 1, 0]\})$, but \mathcal{F} is not a subset of any of them. The next lemma shows that this case also implies $\oplus P$ -completeness.

LEMMA 5.2. For a signature set \mathcal{F} , if $\mathcal{F} \subseteq \mathscr{A} \cup \mathscr{M} \cup (\mathscr{F} \cup \{[0,1,0]\})$, but $\mathcal{F} \not\subseteq \mathscr{A}$, $\mathcal{F} \not\subseteq \mathscr{M}$, and $\mathcal{F} \not\subseteq \mathscr{F} \cup \{[0,1,0]\}$, then the parity problem \oplus Holant^c(\mathcal{F}) is $\oplus P$ -complete.

Proof. Since $\mathcal{F} \not\subseteq \mathcal{M}$ and every signature with arity at most 2 is a matchgate signature, there must exist a signature $f \in \mathcal{F}$ of arity at least 3, which is not a matchgate signature. Therefore f is an equality signature or a Fibonacci signature of arity at least 3.

If f is an equality signature [1, 0, ..., 0, 1], and because \mathcal{F} is not a subset of \mathscr{A} , there must be a signature in \mathcal{F} that contains a subsignature [0, 1, 1] or [0, 1, 0, 0]. According to Corollary or Lemma 4.6, \oplus Holant^c(\mathcal{F}) is \oplus P-complete.

If f is a Fibonacci signature with arity at least 3, and because \mathcal{F} is not a subset of $\mathscr{F} \cup \{[0,1,0]\}$, there must be a ternary matchgate or an equality signature in \mathcal{F} since all binary signatures are also in $\mathscr{F} \cup \{[0,1,0]\}$. Depending on which signature is present, we apply either Corollary 4.3 or Corollary 4.5 to conclude \oplus Holant^c(\mathcal{F}) is always \oplus P-complete. \Box

In section 3 we have shown that if $\mathcal{F} \subseteq \mathscr{A}$, $\mathcal{F} \subseteq \mathscr{M}$, or $\mathcal{F} \subseteq \mathscr{F} \cup \{[0,1,0]\}$, then \oplus Holant^c(\mathcal{F}) is computable in polynomial time. Together with the above two lemmas, we have the dichotomy of \oplus Holant^c problems.

THEOREM 5.3. If $\mathcal{F} \subseteq \mathscr{A}$, $\mathcal{F} \subseteq \mathscr{M}$, or $\mathcal{F} \subseteq \mathscr{F} \cup \{[0, 1, 0]\}$, then the parity problem \oplus Holant^c(\mathcal{F}) is computable in polynomial time. Otherwise it is \oplus P-complete. The same statement also holds for planar graphs.

5.2. Dichotomy for 2-3 regular graphs. The following dichotomy for 2-3 regular graphs is a side product along the way to proving the \oplus Holant dichotomy. It is not subsumed into the general dichotomy because the underlying graph has a better structure. This case is of independent interest because many problems investigated in [41] correspond to certain cases here. Also, the same setting was previously studied in the Holant framework [13, 14, 31, 10] for #P.

THEOREM 5.4. Let $f_2 = [y_0, y_1, y_2]$ and $g_3 = [x_0, x_1, x_2, x_3]$ be two Boolean symmetric signatures. If this signature pair is any of the six [1, 0, 1]|[0, 1, 1, 1], [1, 0, 1]|[1, 1, 1, 0], [1, 1, 0]|[1, 0, 0, 1], [0, 1, 1]|[1, 0, 0, 1], [0, 1, 1]|[0, 1, 1, 1], and <math>[1, 1, 0]|[1, 1, 1, 0],then \oplus Holant($[y_0, y_1, y_2]|[x_0, x_1, x_2, x_3]$) is \oplus P-complete. In all other cases it is polynomial time computable. The same statement also holds for planar graphs.

Proof. The first four hardness results are exactly the starting points given in Theorem 4.1 and Corollary 4.2. The hardness for [0, 1, 1]|[0, 1, 1, 1] (and symmetrically [1, 1, 0]|[1, 1, 1, 0]) is proved by Valiant in [41] as \oplus Pl-3/2Bip-Mon-2CNF. This can be proved also by holographic transformation from \oplus Holant([0, 1, 1]|[1, 0, 0, 1]) as follows. We prove it for [0, 1, 1]|[0, 1, 1, 1], and the proof of [1, 1, 0]|[1, 1, 1, 0] is similar. Under the basis $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ (notice that $T^{-1} = T$ under module 2), we have $T^{\otimes 3}[0, 1, 1, 1] =$ [1, 0, 0, 1] and $[0, 1, 1]T^{\otimes 2} = [0, 1, 1]$. Therefore, \oplus Holant([0, 1, 1]|[0, 1, 1, 1]) is polynomially equivalent to \oplus Holant([0, 1, 1]|[1, 0, 0, 1]), which is \oplus P-complete.

Now, we consider the tractable cases. Most of them belong to one of the identified tractable families directly, as summarized in the following table (we omit all the degenerate cases and some symmetric cases).

| $g_3 \mid f_2$ | [0, 1, 0] | [1, 0, 1] | [1, 1, 0] |
|----------------|---------------------------------|---------------------------------|----------------------|
| [0, 0, 1, 0] | М | М | М |
| [0, 0, 1, 1] | М | М | М |
| [0, 1, 0, 0] | М | М | М |
| [0, 1, 0, 1] | \mathscr{A} and \mathscr{M} | \mathscr{A} and \mathscr{M} | М |
| [0, 1, 1, 0] | $\mathscr{F} \cup \{[0,1,0]\}$ | Ŧ | Ŧ |
| [0, 1, 1, 1] | Tractable | $\oplus P$ -complete | Tractable |
| [1, 0, 0, 1] | A | A | $\oplus P$ -complete |
| [1, 0, 1, 0] | \mathscr{A} and \mathscr{M} | \mathscr{A} and \mathscr{M} | М |
| [1, 0, 1, 1] | $\mathscr{F} \cup \{[0,1,0]\}$ | Ŧ | Ŧ |
| [1, 1, 0, 0] | М | М | М |
| [1, 1, 0, 1] | $\mathscr{F} \cup \{[0,1,0]\}$ | Ŧ | Ŧ |
| [1, 1, 1, 0] | Tractable | $\oplus P$ -complete | $\oplus P$ -complete |

There are three entries marked "Tractable." They do not belong to any tractable family directly, but we show that after a holographic reduction they do. Under the basis $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, we have that \oplus Holant([0, 1, 0] | [0, 1, 1, 1]) is equivalent to \oplus Holant([0, 1, 0] | [1, 0, 0, 1]), for which we have a polynomial algorithm since both signatures are affine. The problem of \oplus Holant([0, 1, 0] | [1, 1, 1, 0]) is symmetric with \oplus Holant([0, 1, 0] | [0, 1, 1, 1]), so it is also in P. Under the same basis $T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, we have that \oplus Holant([1, 1, 0] | [0, 1, 1, 1]) is polynomially equivalent to \oplus Holant([1, 0, 0, 1] | [1, 0, 0, 1]), for which we have a polynomial algorithm since both signatures are affine. This completes the proof of the dichotomy theorem.

6. Vanishing signature sets. In the remaining two sections we extend our results to obtain the dichotomy result for \oplus Holant without any assumptions. In order to formulate the dichotomy we shall need a fourth family of tractable signature sets, which we call *vanishing signature sets*.

DEFINITION 6.1. A set of signatures \mathcal{F} is called vanishing iff the value of \oplus Holant_{Ω}(\mathcal{F}) is zero for every Ω . We denote the class of all vanishing signature sets by \mathcal{O} .

First we show some general properties of vanishing signature sets. For two signatures f and g of the same arity, f + g denotes the bitwise addition in \mathbb{Z}_2 , i.e., $[f_0 + g_0, f_1 + g_1, \ldots]$.

LEMMA 6.2. Let \mathcal{F} be a vanishing signature set. If a signature f can be realized by a gadget using signatures in \mathcal{F} , then $\mathcal{F} \cup \{f\} \in \mathcal{O}$. If g_0 and g_1 are two signatures in \mathcal{F} with the same arity, then $\mathcal{F} \cup \{g_0 + g_1\} \in \mathcal{O}$.

Proof. The first statement is immediate. Now we prove the second, which says that a vanishing signature set is closed under linear combination.

Let $\Omega = (H, \mathcal{F} \cup \{g_0 + g_1\}, \pi)$ be an instance of \oplus Holant $(\mathcal{F} \cup \{g_0 + g_1\})$. We want to show that Holant $\Omega = 0$. If the signature $g_0 + g_1$ does not appear in H, then Holant Ω is zero since $\mathcal{F} \in \mathcal{O}$. Otherwise, we denote by U the set of vertices having the signature $g_0 + g_1$. Then

$$\begin{aligned} \operatorname{Holant}_{\Omega} &= \sum_{\sigma} \prod_{v \in V} f_v(\sigma \mid_{E(v)}) \\ &= \sum_{\sigma} \left(\prod_{v \notin U} f_v(\sigma \mid_{E(v)}) \prod_{v \in U} (g_0(\sigma \mid_{E(v)}) + g_1(\sigma \mid_{E(v)})) \right) \\ &= \sum_{\sigma} \left(\prod_{v \notin U} f_v(\sigma \mid_{E(v)}) \left(\sum_{i_v \in \{0,1\} \text{ (for all } v \in U)} \prod_{v \in U} g_{i_v}(\sigma \mid_{E(v)}) \right) \right) \\ &= \sum_{i_v \in \{0,1\} \text{ (for all } v \in U)} \sum_{\sigma} \left(\prod_{v \notin U} f_v(\sigma \mid_{E(v)}) \prod_{v \in U} g_{i_v}(\sigma \mid_{E(v)}) \right). \end{aligned}$$

Every term in the outer summation is a Holant value on the same graph obtained by replacing the signature in v by g_{i_v} for every vertex $v \in U$. These are all instances of \oplus Holant(\mathcal{F}), and therefore their values are all zero since $\mathcal{F} \in \mathcal{O}$. As a summation Holant_{Ω} is also zero. This completes the proof. \square

In the following two subsections, we mention some simple vanishing families of signature sets. They are not really used in the dichotomy proof, but they may give some intuition, and some interesting phenomena can be noted from them. In the last part of this section, we will introduce the self-vanishable signatures, which is crucial in the proof of the general dichotomy.

6.1. Complement invariant signatures. A symmetric signature $[f_0, f_1, \ldots, f_n]$ is complement invariant iff $f_k = f_{n-k}$ for all k.

DEFINITION 6.3. A signature f is called complement invariant iff for any input $\alpha \in \{0,1\}^n$ we have $f(\alpha) = f(\overline{\alpha})$.

If all the signatures involved in a Holant instance are complement invariant, then any assignment of edges and its complement have the same value. The value at a vertex for an assignment of edges is the same as for its complement assignment. Hence the Holant value is always zero modulo 2.

PROPOSITION 6.4. Let \mathcal{F} be a set of complement invariant signatures. Then \mathcal{F} is vanishing.

As a side note, this family of signature sets corresponds to the additional tractable case in Faben's work [25] regarding parity dichotomy of the CSP framework.

6.2. Matching-based vanishing signature sets. Here we describe another family of vanishing signature sets. In a graph where all nodes have even degree, the parity of the number of perfect matchings is even. This can be easily shown using the relation (1). The parity of perfect matchings is equal to that of the determinant of its adjacency matrix. Adding up all rows of the adjacency matrix, we get a vector composed of even numbers. Thus this matrix must be singular in the field \mathbb{Z}_2 , and its determinant is zero.

Furthermore, by using relation (2) we can also deduce the same result for graphs composed of perfect matching nodes of even arity and partial matching nodes of odd arity. According to the relation (2), the parity of the number of general matchings equals $Pf(A + \Lambda^{(n)})$ if *n*, the number of nodes, is even, or $Pf(A^+ + \Lambda^{(n+1)})$ if *n* is odd. Noticing that the number of vertices of odd arities must be even, it is easy to verify the summation of all rows in $A + \Lambda^{(n)}$ for even *n*, or the first *n* rows in $A^+ + \Lambda^{(n+1)}$ for odd *n* are zero vectors in \mathbb{Z}_2 . Hence, the Pfaffian, which equals the determinant, is zero modulo 2.

By Lemma 6.2, the linear combination of these matching signatures, or signatures that can be realized from them, also belongs to this vanishing signature family. Using this fact, one can see that some sets of signatures, for which the Holant^c problem is \oplus P-complete, are actually vanishing, and thus the corresponding Holant problem is tractable. For example, Holant^c({[1,0,1,1,1]}) is \oplus P-complete, but the signature [1,0,1,1,1] alone is in this vanishing family because [1,0,1,1,1] = [0,0,0,1,0] + [1,0,1,0,1], where [0,0,0,1,0] is a perfect matching signature and [1,0,1,0,1] can be realized via [0,0,0,1,0].

PROPOSITION 6.5. If a signature set \mathcal{F} is composed of perfect matching signatures of even arities, partial matching signatures of odd arities, signatures realizable from them, and linear combinations of all of the above, then \mathcal{F} is a vanishing set.

6.3. Self-vanishable signatures. In this section, we introduce a new concept called *self-vanishable signatures* which plays an important role in the proof of the general dichotomy. First, we introduce an extended version of the inner product for two signatures of not necessarily the same arity.

DEFINITION 6.6. Let f and g be two signatures with arities n and m $(n \ge m)$, respectively. Their inner product $h = \langle f, g \rangle$ is a signature with arity n - m defined as follows:

$$h^{\alpha} = \sum_{\beta \in \{0,1\}^m} f^{\beta,\alpha} g^{\beta},$$

where $\alpha \in \{0, 1\}^{n-m}$.

If f is symmetric, the final $h = \langle f, g \rangle$ is also symmetric. If both f and g are symmetric, their inner product $h = [h_0, h_1, \ldots, h_{n-m}]$ has the following form: $h_i = \sum_{j=0}^m {m \choose j} f_{j+i} g_j$ for $0 \le i \le n-m$. Hence, in \mathbb{Z}_2 ,

(4)
$$\langle f, [1,1]^{\otimes 2} \rangle = \langle f, [1,1,1] \rangle = \langle f, [1,0,1] \rangle,$$

since $\binom{2}{1}$ is 2. We will use this simple fact in the future.

We can also view this inner product in a combinatorial way. Given two gates with signatures f and g, connecting m dangling edges of f to the edges of g (see Figure 10) results in a gadget with signature $\langle f, g \rangle$.



FIG. 10. The extended inner product.

DEFINITION 6.7. A signature f is called self-vanishable of degree k iff there exists a unique k such that $\langle f, [1,1]^{\otimes k} \rangle = \mathbf{0}$ and $\langle f, [1,1]^{\otimes k-1} \rangle \neq \mathbf{0}$. We denote this by v(f) = k. If such a k does not exist, the signature f is not self-vanishable.

We note that for the trivial signature **0** we have $v(\mathbf{0}) = 0$. Also, f = [1, 1] is self-vanishable with v(f) = 1 since $\langle [1, 1], [1, 1] \rangle = 0$.

To be self-vanishable is a necessary condition for a signature to be a member of a vanishing signature set, as shown in the following lemma. It also partly explains the intuition for why we define this notion of self-vanishable and why we define it in this way. In fact, for even k, $\langle f, [1,1]^{\otimes k} \rangle = \langle f, [1,1,1]^{\otimes \frac{k}{2}} \rangle = \langle f, [1,0,1]^{\otimes \frac{k}{2}} \rangle$, by relation (4), with the last signature being equivalent to connecting together pairwise the k inputs of the original signature.

LEMMA 6.8. If \mathcal{F} contains a signature f which is not self-vanishable, then \mathcal{F} is not a vanishing set.

Proof. Let n be the arity of f. If n is even, we can connect all its inputs pairwise by n/2 edges. The resulting signature is of arity 0, which means it is a single value, and the value is $\langle f, [1,1]^{\otimes n} \rangle$ by the argument above. It is not zero since f is not

self-vanishable. Therefore \mathcal{F} is not vanishing since we can construct an instance of it whose value is not zero.

If n is odd, we can connect all its edges but one pairwise. This is a gadget with one dangling edge, and similarly its signature is $\langle f, [1,0,1]^{\otimes \frac{n-1}{2}} \rangle = \langle f, [1,1]^{\otimes n-1} \rangle$. This cannot be [0,0] or [1,1] since f is not self-vanishable. So it must be [0,1] or [1,0]. In both cases, after connecting the dangling edges of two copies of such a gadget, we get a graph whose Holant value is 1. So \mathcal{F} is not vanishing. This completes the proof. \Box

The following lemma is immediate.

LEMMA 6.9. Let f be self-vanishable of degree $k \ge r > 0$. Then $v(\langle f, [1, 1]^{\otimes r} \rangle) = k - r$.

For a symmetric signature $f = [f_0, f_1, \ldots, f_n]$, we call f_0 the first entry of f and call $f_0, f_1, \ldots, f_{k-1}$ the first k entries of f. It follows from the definition that for a symmetric signature $f = [f_0, f_1, \ldots, f_n]$, we have

$$\langle f, [1,1] \rangle = [f_0 + f_1, f_1 + f_2, \dots, f_{n-1} + f_n].$$

Hence the only symmetric signature of arity n with v(f) = 1 is $[1,1]^{\otimes n}$. There are two symmetric signatures of arity $n \geq 3$ with v(f) = 2, which are the parity signatures $[1,0,1,0,\ldots,0/1]$ and $[0,1,0,1,\ldots,0/1]$. Using this fact and Lemma 6.9, it is easy to verify that there are only four symmetric signatures of arity $n \geq 3$ with v(f) = 3. The first two entries are arbitrary and determine all remaining entries. They are therefore $[0,0,1,1,0,0,1,1,\ldots]$, $[0,1,1,0,0,1,1,0,\ldots]$, $[1,0,0,1,1,0,0,1,\ldots]$, and $[1,1,0,0,1,1,0,0,\ldots]$. One may notice that such a signature is always periodic. In general, we have the following lemma.

LEMMA 6.10. For any $k \ge 2$, there are 2^{k-1} symmetric signatures of arity $n \ge k$ with v(f) = k, of which the first k-1 entries are arbitrary and the remaining entries are determined by them.

It follows from this lemma that any two symmetric signatures with the same arity that are self-vanishable of degree k are identical if they are identical on the first k-1 entries. As a matter of fact, any self-vanishable signature is periodic (in its symmetric form), and the period is determined only by the degree k up to a shift. Moreover, since the first k-1 entries are free, any possible sequence of length k-1 should appear as a subsequence of the period (possibly after a shift), though we will not really need this fact in the future.

Call symmetric self-vanishable signatures of degree k with first k-1 entries all 0's the canonical form of vanishable signatures of degree k. Later we will show that any self-vanishable signature of degree k is a linear combination of canonical self-vanishable signatures of degree at most k.

Regarding this canonical form, first we have the following lemma.

LEMMA 6.11. If f has the canonical form for symmetric self-vanishable signatures of degree k, then $\langle f, [1,1] \rangle$ has the canonical form for symmetric self-vanishable signatures of degree k - 1.

Proof. This lemma is immediate since the first k - 2 entries of $\langle f, [1, 1] \rangle$ are all zero and $v(\langle f, [1, 1] \rangle) = k - 1$.

Let v^k be the canonical symmetric self-vanishable signature of degree k with arity n. Then, $\langle v^k, [1,1]^{\otimes k-1} \rangle$ is self-vanishable with degree 1, which is of form $(1,1,\ldots,1)$. On the other hand, the first bit of $\langle v^k, [1,1]^{\otimes k-1} \rangle$ is v^k_{k-1} since $v^k_i = 0$ for all $i \leq k-2$. From these two facts, we can conclude that $v^k_{k-1} = 1$. Using this, we can show the purpose of defining such a canonical form.

LEMMA 6.12. Every symmetric self-vanishable signature of degree k can be expressed as a sum of several symmetric self-vanishable signatures in canonical form, whose degrees are all less than or equal to k.

Proof. Let v^i be the canonical symmetric self-vanishable signature of degree i with arity n. Let f be a symmetric self-vanishable signature of degree k, whose first k-1 bits are $f_0, f_1, \ldots, f_{k-2}$. Define f' as follows:

$$f' = v^k + \sum_{i=1}^{k-1} x_i v^i,$$

where x_i satisfies $\sum_{i=1}^{k-1} x_i v_j^i = f_j$ for $1 \le i \le k-1$. View this relation as a linear equation system. The solution $\{x_i\}$ always exists because the coefficient matrix is

| /1 | 1 | 1 |
|--|---|---------|
| 0 | 1 | 0/1 |
| : | : | : |
| $\begin{pmatrix} \cdot \\ 0 \end{pmatrix}$ | 0 | 1 |

of full rank. The diagonals are all 1's because of the fact mentioned above. Since $\sum_{i=1}^{k-1} x_i v_j^i = f_j$ and $f' = v^k + \sum_{i=1}^{k-1} x_i v^i$, it is clear that the first k-1 entries of f' are exactly $f_0, f_1, \ldots, f_{k-2}$.

Then by Lemma 6.10, it is sufficient to prove that f' is a self-vanishable signature of degree k. This can be verified as follows:

$$\langle f', [1,1]^{\otimes k} \rangle = \langle v^k, [1,1]^{\otimes k} \rangle + \sum_{i=1}^{k-1} x_i \langle v^i, [1,1]^{\otimes k} \rangle = \mathbf{0}$$

and

$$\langle f', [1,1]^{\otimes k-1} \rangle = \langle v^k, [1,1]^{\otimes k-1} \rangle + \sum_{i=1}^{k-1} x_i \langle v^i, [1,1]^{\otimes k-1} \rangle = \langle v^k, [1,1]^{\otimes k-1} \rangle \neq \mathbf{0}.$$

This completes the proof. \Box

LEMMA 6.13. The canonical symmetric self-vanishable signature of degree $k \ge 1$ can be expressed as follows:

$$v^k = \sum_{S \subseteq [n], |S|=k-1} \bigotimes_{i=1}^n u_{[i \in S]}$$

where $[i \in S] = 1$ iff $i \in S$, $u_0 = [1, 1]$, and $u_1 = [0, 1]$.

Proof. We prove it by induction on k. It is obvious for k = 1. Now we assume

that the lemma holds for k-1. For k>1, let $f=\sum_{S\subset [n],|S|=k-1}\otimes_{i=1}^{n}u_{[i\in S]}$. Then

$$\begin{split} \langle f, [1,1] \rangle &= \left\langle \sum_{S \subseteq [n], |S| = k-1} \bigotimes_{i=1}^{n} u_{[i \in S]}, [1,1] \right\rangle \\ &= \sum_{S \subseteq [n], |S| = k-1} \left\langle \bigotimes_{i=1}^{n} u_{[i \in S]}, [1,1] \right\rangle \\ &= \sum_{S \subseteq [n], |S| = k-1} \langle u_{[n \in S]}, [1,1] \rangle \otimes \bigotimes_{i=1}^{n-1} u_{[i \in S]} \\ &= \sum_{S \subseteq [n], |S| = k-1, n \in S} \bigotimes_{i=1}^{n-1} u_{[i \in S]} \\ &= \sum_{S \subseteq [n-1], |S| = k-2} \bigotimes_{i=1}^{n-1} u_{[i \in S]}. \end{split}$$

By the induction hypothesis, the right-hand side is exactly the canonical symmetric self-vanishable signature of degree k - 1. Hence the first k - 2 entries of $\langle f, [1, 1] \rangle$ are all zero, which means the first k - 1 entries of f are all the same. However, it is easy to check that the first entry is zero since, if k > 1, every term in the summation contributes zero. This completes the proof. \Box

DEFINITION 6.14. Let f be self-vanishable of degree $k \ge 0$ with arity n. It is called strong self-vanishable if $k \le \lfloor \frac{n}{2} \rfloor + 1$ and weak self-vanishable if $\lfloor \frac{n}{2} \rfloor + 2 \le k \le n$.

THEOREM 6.15. Let \mathcal{F} be a set of symmetric strong self-vanishable signatures. Then \mathcal{F} is a vanishing set, i.e., $\mathcal{F} \in \mathcal{O}$.

Proof. By Lemmas 6.2 and 6.12 it is sufficient to prove the theorem for signatures in the canonical form. Each such signature we express is in the form shown in Lemma 6.13. Each term of the decomposition is a degenerate signature, a tensor product of two types of unary signatures [1, 1] and [0, 1]. For a strong self-vanishable signature, we have $k \leq \lfloor \frac{n}{2} \rfloor + 1$, which implies that the number of [1, 1]'s is greater than or equal to the number of [0, 1]'s.

We consider separately the two cases of strict inequality and equality.

First suppose that there is at least one signature used in the input that has the property that $k < \lfloor \frac{n}{2} \rfloor + 1$, which means that there are strictly more [1, 1]'s than [0, 1]'s. In this case, we can further decompose the Holant value as in Lemma 6.2 into a sum of several (possibly exponentially many) Holant values according to the decomposition of canonical signatures in Lemma 6.13. Then in each term, every signature involved is degenerate. A node of degree d can be viewed as d unary signatures ([1, 1] and [0, 1]). Therefore the whole graph is decomposed into isolated edges. For each edge, the signatures on its two ends are either [1, 1] or [0, 1]. The Holant value is the product of every such edge. If both ends of one edge have signature [1, 1], then the value for this edge is zero and so is the Holant value. However, in every Holant, such a cancellation must happen at some edge because there are strictly more [1, 1]'s than [0, 1]'s. Hence, in total, the whole Holant is a sum of (possibly exponentially many) zeros, which is still zero.

If there is a signature of odd arity, even when $k = \lfloor \frac{n}{2} \rfloor + 1$, then there are strictly more [1, 1]'s than [0, 1]'s, and the argument above still works. The remaining case is

that of all signatures having even arity and satisfying $k = \lfloor \frac{n}{2} \rfloor + 1$. In that case we do the same decomposition as in the previous paragraph. The numbers of [1, 1]'s and [0, 1]'s are now exactly equal. There may exist some Holants whose value is one, and hence the argument above does not work. In this case, we need to look further into the structure of the decomposition

$$f = \sum_{S \subseteq [n], |S| = k-1} \bigotimes_{i=1}^{n} v_{[i \in S]} = \sum_{S \subseteq [n], |S| = \frac{n}{2}} \bigotimes_{i=1}^{n} u_{[i \in S]}.$$

Notice that n is even. As in the proof of Lemma 6.2, we have

$$\begin{aligned} \operatorname{Holant}_{\Omega} &= \sum_{\sigma} \prod_{v \in V} f_v(\sigma \mid_{E(v)}) \\ &= \sum_{S_v \subseteq [n_v], |S_v| = \frac{n_v}{2} \text{ (for all } v \in V)} \sum_{\sigma} \prod_{v \in V} \bigotimes_{i=1}^n u_{[i \in S_i]}(\sigma \mid_{E(v)}) \\ &= \sum_{S_v \subseteq [n_v], |S_v| = \frac{n_v}{2} \text{ (for all } v \in V)} \prod_{(i,j) \in E} \left\langle u_{[t^i_{(i,j)} \in S_i]}, u_{[t^j_{(i,j)} \in S_j]} \right\rangle. \end{aligned}$$

Here $t_{(i,j)}^i$ and $t_{(i,j)}^j$ are numbers of the edge (i, j) in the numbering of edges of nodes i and j, respectively. The term indexed by some S_v 's in the summation contributes one iff it satisfies the condition that for all edges $(i, j) \in E$, exactly one of the two $t_{(i,j)}^i \in S_i$ and $t_{(i,j)}^j \in S_j$ is true. Notice that if some S_v 's satisfy the condition, their complement $\overline{S_v}$'s also satisfy it. Hence, if a term indexed by some S_v 's is one, it will be canceled out with the term indexed by the $\overline{S_v}$'s. (Here we use the fact that $|\overline{S_v}| = |S_v| = \frac{n_v}{2}$.) This completes the proof.

As a final remark we note that the family \mathcal{O} of vanishing signature sets has the following difference from the previous tractable families \mathscr{A} , \mathscr{M} , and $\mathscr{F} \cup \{[0, 1, 0]\}$. The union of two sets in \mathcal{O} is not necessarily in \mathcal{O} . For example, the union of the sets $\{[0, 0, 1, 1, 0]\}$, whose member is strong self-vanishable, and $\{[1, 0, 1, 1, 1]\}$, which is matching-based vanishing, is not vanishing.

7. Dichotomy for the whole Holant family. In this final section, we prove our main theorem, the dichotomy for all parity Holant problems with symmetric signatures, without assuming any freely available signatures. This improves on our dichotomy theorem for parity Holant^c problems given in section 5, which we use, however, as our starting point. The main idea is to construct gadgets for the two signatures [0,1] and [1,0]. We will first show that realizing either one of these is enough. When one of these unary signatures is realizable, we reduce the Holant problem to the corresponding Holant^c problem and apply the Holant^c dichotomy result. However, for some signature sets it is impossible to realize [0,1] or [1,0]. We show that those signature sets must be vanishing, in the sense defined in the previous section.

We remark that the gadgets used in the proofs here are not all planar, and hence the dichotomy for planar graphs does not follow.

Therefore we show that it is enough to realize just one of [0,1] and [1,0]. First we show how to use the degenerate signature [0,0,1] (or [1,0,0]) as its base signature [0,1] (or [1,0]).

LEMMA 7.1. For any signature set \mathcal{F} , the complexity of \oplus Holant($\mathcal{F} \cup \{[1,0]\}$) (or \oplus Holant($\mathcal{F} \cup \{[0,1]\}$)) is the same as \oplus Holant($\mathcal{F} \cup \{[1,0,0]\}$) (or \oplus Holant($\mathcal{F} \cup \{[0,0,1]\}$)).



FIG. 11. Simulating [1,0] using [1,0,0] or [1,1,0].

Proof. First we reduce \oplus Holant($\mathcal{F} \cup \{[1,0]\}$) to \oplus Holant($\mathcal{F} \cup \{[1,0,0]\}$). Given a grid composed of signatures from $\mathcal{F} \cup \{[1,0]\}$, we may replicate the grid and replace every pair of corresponding occurrences of [1,0] by [1,0,0], as depicted in Figure 11. View the part of the grid excluding the [1,0] signatures as a signature G. The Holant value of the left grid is $G^{00...0}$, while the Holant value of the right is $(G^{00...0})^2$, which equals $G^{00...0}$ since for any $x, x^2 \equiv x \pmod{2}$.

It is easy to see that \oplus Holant($\mathcal{F} \cup \{[1,0,0]\}$) can be reduced to \oplus Holant($\mathcal{F} \cup \{[1,0]\}$) since [1,0,0] can be realized by two copies of [1,0]. Thus \oplus Holant($\mathcal{F} \cup \{[1,0]\}$) and \oplus Holant($\mathcal{F} \cup \{[1,0,0]\}$) have the same complexity. Similarly, \oplus Holant($\mathcal{F} \cup \{[0,1]\}$) and \oplus Holant($\mathcal{F} \cup \{[0,0,1]\}$) have the same complexity. \square

It may not always be possible that, having only one of [0, 1] and [1, 0] we can construct the other. However, if the signature set is of one of the three tractable families, then we do not need to worry. Otherwise we show that if the other unary signature is not easy to construct, the signature set itself is hard already.

LEMMA 7.2. Let \mathcal{F} be a set of symmetric signatures. If $\mathcal{F} \subseteq \mathscr{A}$, $\mathcal{F} \subseteq \mathscr{M}$, or $\mathcal{F} \subseteq \mathscr{F} \cup \{[0, 1, 0]\}$, then the parity problems \oplus Holant($\mathcal{F} \cup \{[1, 0]\}$), \oplus Holant($\mathcal{F} \cup \{[0, 1]\}$), and \oplus Holant($\mathcal{F} \cup \{[0, 0, 1]\}$) are computable in polynomial time. Otherwise these parity problems are \oplus P-complete.

Proof. By Lemma 7.1 and symmetry, we need only prove the lemma for the case of \oplus Holant($\mathcal{F} \cup \{[1,0]\}$). If \mathcal{F} is a subset of \mathscr{A} , \mathscr{M} , or $\mathscr{F} \cup \{[0,1,0]\}$, then as we have already shown, \oplus Holant($\mathcal{F} \cup \{[1,0]\}$) is computable in polynomial time.

Then we consider whether \mathcal{F} is not in the three tractable families. If we can simulate [0, 1] or [0, 0, 1], then by Theorem 5.3 and Lemma 7.1, \oplus Holant($\mathcal{F} \cup \{[1, 0]\}$) is \oplus P-complete.

Since \mathcal{F} is not a subset of \mathscr{A} , there exists one signature $f \in \mathcal{F}$, which is not degenerate and not in \mathscr{A} . Consider the first bit of f. Assume it is 0. If the next bit is 1, we are done, since with the help of [1,0] we can get any front part of the signature. Otherwise it begins with several successive 0's and one 1. Use [1,0] to get this $[0,0,\ldots,1]$ of arity k. If k = 2, then it is [0,0,1] and we are done. Otherwise connect any two of its dangling edges. The resulting signature is also a $[0,0,\ldots,1]$ but of arity k-2. After repeating this process, it will eventually become [0,1] or [0,0,1].

Next we assume the first bit is 1. If f begins with more than two successive 1's, we may get a signature of the form [1, 1, ..., 1, 0] of arity k.

- If $k \ge 3$, connecting any two of its dangling edges will result in a signature of the form $[0, 0, \dots, 0, 1]$ of arity k 2, and we are done.
- If k = 2, connecting two copies of [1, 1, 0] will give us [0, 1, 1], and we can get [0, 1] out of it by connecting a [1, 0].

- Otherwise, it begins with 1, 0. Now we consider the number of successive 0's here:
 - If there is only one 0 here, f must start with a signature of the form $[1,0,1,0,\ldots,1,0,0]$ or $[1,0,1,0,\ldots,0,1,1]$, since it cannot be a parity signature $[1,0,1,0,\ldots,0/1] \in \mathscr{A}$. In either case, connecting any two of its dangling edges will result in a signature beginning with $[0,0,\ldots,0,1]$.
 - Otherwise f must begin with a signature of the form $[1, 0, \ldots, 0, 1, 0]$ or $[1, 0, \ldots, 0, 1, 1]$, where the number of 0 is at least 2, since f is not degenerate and cannot be an equality signature $[1, 0, \ldots, 0, 1] \in \mathscr{A}$. In either case, we connect two dangling edges to reduce its arity as well as the number of 0's in the middle. Finally, according to the parity of its arity, it will become one of the following four cases: [1, 0, 0, 1, 0], [1, 0, 0, 1, 1], [1, 0, 0, 0, 1, 0], or [1, 0, 0, 0, 1, 1].
 - * For [1, 0, 0, 1, 0], connect one more pair of its dangling edges to get [1, 1, 0], which was discussed before.
 - * For [1, 0, 0, 1, 1], put it in every vertex in the gadget shown in Figure 12, and the resulting signature is [0, 1, 0, 1, 0]. We can get [0, 1] from this by connecting it with three [1, 0]'s.
 - * For [1, 0, 0, 0, 1, 0], put it in every vertex in the gadget shown in Figure 13, and the resulting signature is [0, 0, 1].
 - * For [1,0,0,0,1,1], connect its dangling edges two more times, and we get [0,1].

This completes our proof.



FIG. 12. The tetrahedron gadget.



FIG. 13. The gadget for [1, 0, 0, 0, 1, 0].

Using the idea of replication in Lemma 7.1, [1,1,0] can also be used as [1,0]. Then formally we have the following corollary. Notice that here, to be tractable the signature set \mathcal{F} cannot be a subset of \mathscr{A} , because [1,1,0] together with equality gate $[1,0,\ldots,0,1] \in \mathscr{A}$ would lead to hardness.

COROLLARY 7.3. Let \mathcal{F} be a set of symmetric signatures. If $\mathcal{F} \subseteq \mathcal{M}$ or $\mathcal{F} \subseteq \mathscr{F} \cup \{[0, 1, 0]\}$, then the parity problems \oplus Holant($\mathcal{F} \cup \{[1, 1, 0]\}$) and \oplus Holant($\mathcal{F} \cup \{[0, 1, 1]\}$) are computable in polynomial time. Otherwise they are \oplus P-complete.

Proof. We prove the corollary by reducing \oplus Holant($\mathcal{F} \cup \{[1,0]\}$) to \oplus Holant($\mathcal{F} \cup \{[1,1,0]\}$). We use the same idea as in the reduction of Lemma 7.1. Given a grid composed of signatures from $\mathcal{F} \cup \{[1,0]\}$, we replicate the grid and replace every pair of corresponding occurrences of [1,0] by [1,1,0], as depicted in Figure 11. The Holant value of the right grid is a summation of G^iG^j , where i and j are two index vectors and their bitwise conjunction does not contain any 1. Thus, if $i \neq j$ and G^iG^j contributes one to the summation, G^jG^i also contributes one to it and they vanish modulo 2. For

i = j, there is only one contributing term $(G^{00...0})^2 = G^{00...0}$, which is the Holant value of the left grid.

Then by Lemma 7.2, if none of $(\mathcal{F} \cup \{[1,1,0]\}) \subseteq \mathscr{A}, (\mathcal{F} \cup \{[1,1,0]\}) \subseteq \mathscr{M}, \text{ or } (\mathcal{F} \cup \{[1,1,0]\}) \subseteq \mathscr{F} \cup \{[0,1,0]\} \text{ holds, then } \oplus \text{Holant}(\mathcal{F} \cup \{[1,1,0]\}) \text{ is } \oplus \text{P-complete.}$ Noticing that $[1,1,0] \in \mathscr{M}, [1,1,0] \in \mathscr{F}$ while $[1,1,0] \notin \mathscr{A}$, we see that the above condition simplifies to $\mathcal{F} \subseteq \mathscr{M}$ or $\mathcal{F} \subseteq \mathscr{F} \cup \{[0,1,0]\}$. \Box

Finally we are ready to show our main theorem.

THEOREM 7.4. Let \mathcal{F} be a set of symmetric signatures. If $\mathcal{F} \subseteq \mathscr{A}$, $\mathcal{F} \subseteq \mathscr{M}$, $\mathcal{F} \subseteq \mathscr{F} \cup \{[0,1,0]\}$, or $\mathcal{F} \in \mathscr{O}$, then the parity problem \oplus Holant(\mathcal{F}) is computable in polynomial time. Otherwise it is $\oplus P$ -complete.

Proof. If $\mathcal{F} \in \mathcal{O}$, then \oplus Holant(\mathcal{F}) is trivially computable in polynomial time since we can just return 0 for any input. The other three classes have already been shown to be tractable.

Now we assume that $\mathcal{F} \notin \mathcal{O}$. By definition, there is an instance G of \oplus Holant(\mathcal{F}) whose value is 1. We shall use this instance G as a gadget to establish the reduction. Breaking the graph of G at one arbitrary edge, we get a gadget with two dangling edges. For notational simplicity, we still call this gadget G. Then the value of the original instance is $G^{00} + G^{11}$, which is one. By symmetry, we can assume that $G^{00} = 1$ and $G^{11} = 0$. If $G^{01} = G^{10} = 0$, then we have a gadget with signature [1, 0, 0], and we are done by Lemma 7.2. If $G^{01} = G^{10} = 1$, then we have a gadget with signature [1, 1, 0], and we are done by Corollary 7.3. The remaining cases are $G^{01} = 1$, $G^{10} = 0$ or $G^{01} = 0$, $G^{10} = 1$. These two cases are essentially the same. The only difference is the order of the two dangling edges. So we can assume $G^{01} = 1$, $G^{10} = 0$. This G is $[1, 0] \otimes [1, 1]$. By connecting two copies of this G through their first edge, we get a gadget with signature $[1, 1, 1] = [1, 1] \otimes [1, 1]$. By the same argument as in Lemma 7.1, we can use the signature [1, 1] freely.

Now if all the signatures in \mathcal{F} are strong self-vanishable, then $\mathcal{F} \in \mathcal{O}$ by Theorem 6.15, a contradiction to our assumption. Therefore there exists a signature in \mathcal{F} which is weak self-vanishable or not self-vanishable.

We first assume that there exists a signature f of arity n which is not self-vanishable. If n is odd, by connecting n-1 dangling edges of f in pairs, we get a unary signature $\langle f, [1,0,1]^{\otimes \frac{n-1}{2}} \rangle = \langle f, [1,1]^{\otimes n-1} \rangle$. This is not [0,0] or [1,1] since f is not self-vanishable. So it must be [0,1] or [1,0], and we are done by Lemma 7.2. If n is even, connecting n-2 dangling edges of f in pairs, we get a binary symmetric signature $\langle f, [1,0,1]^{\otimes \frac{n-2}{2}} \rangle = \langle f, [1,1]^{\otimes n-2} \rangle = [a,b,c]$, where $a \neq c$ if f is not self-vanishable. So it must be one of [1,0,0], [0,0,1], [1,1,0], and [0,1,1]. Again we are done by Lemma 7.2 or Corollary 7.3.

Henceforth we may assume that all signatures are self-vanishable and there exists one $f \in \mathcal{F}$, which is weak self-vanishable of degree k. We show that we can therefore construct a gadget with signature [1,0] or [0,1]. This is done by connecting 2(n-k+1)dangling edges of f pairwise with gadget $G = [1,0] \otimes [1,1]$ and 2k - n - 3 dangling edges to a copy of [1,1] each. This gadget is valid because

$$2k - n - 3 \ge 2\left(\left\lfloor \frac{n}{2} \right\rfloor + 2\right) - n - 3 = 2\left\lfloor \frac{n}{2} \right\rfloor - (n - 1) \ge 0,$$

where the first inequality uses the fact that f is weak self-vanishable. Since n - 2(n - k + 1) - (2k - n - 3) = 1, this gadget is of arity 1, and its signature can be calculated

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

as follows:

$$\begin{split} \langle f, G^{\otimes n-k+1} \otimes [1,1]^{\otimes 2k-n-3} \rangle &= \langle f, [1,1]^{\otimes k-2} \otimes [1,0]^{\otimes n-k+1} \rangle \\ &= \langle \langle f, [1,1]^{\otimes k-2} \rangle, [1,0]^{\otimes n-k+1} \rangle. \end{split}$$

The last signature is of arity 1 and hence is composed of the first two entries of $\langle f, [1,1]^{\otimes k-2} \rangle$. By Lemma 6.9, $\langle f, [1,1]^{\otimes k-2} \rangle$ is a self-vanishable signature of degree k - (k-2) = 2. Therefore it must be a parity signature $[1,0,1,0,\ldots,0/1]$ or $[0,1,0,1,\ldots,0/1]$, whose first two bits are either [1,0] or [0,1].

This completes our proof.

REFERENCES

- V. ARVIND AND P. P. KURUR, Graph isomorphism is in spp, Inform. Comput., 204 (2006), pp. 835–852.
- 2] R. BEIGEL, H. BUHRMAN, AND L. FORTNOW, Np might not be as easy as detecting unique solutions, in STOC '98: Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing, ACM, New York, 1998, pp. 203–208.
- [3] R. A. BRUALDI AND H. J. RYSER, Combinatorial Matrix Theory, Cambridge University Press, Cambridge, UK, 1991.
- [4] A. A. BULATOV, A dichotomy theorem for constraint satisfaction problems on a 3-element set, J. ACM, 53 (2006), pp. 66–120.
- [5] A. A. BULATOV, The complexity of the counting constraint satisfaction problem, in ICALP '08: Proceedings of the 35th International Colloquium on Automata, Languages, and Programming, Part I, L. Aceto, I. Damgård, L. A. Goldberg, M. M. Halldórsson, A. Ingólfsdóttir, and I. Walukiewicz, eds., Lect. Notes Comput. Sci. 5125, Springer-Verlag, Berlin, 2008, pp. 646–661.
- [6] A. A. BULATOV AND V. DALMAU, Towards a dichotomy theorem for the counting constraint satisfaction problem, in Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press, Piscataway, NJ, 2003, pp. 562–571.
- [7] J.-Y. CAI, X. CHEN, AND P. LU, Graph homomorphisms with complex values: A dichotomy theorem, in ICALP (1), S. Abramsky, C. Gavoille, C. Kirchner, F. M. auf der Heide, and P. G. Spirakis, eds., Lect. Notes Comput. Sci. 6198, Springer-Verlag, Berlin, 2010, pp. 275– 286.
- [8] J.-Y. CAI, V. CHOUDHARY, AND P. LU, On the theory of matchgate computations, in CCC '07: Proceedings of the Twenty-Second Annual IEEE Conference on Computational Complexity, IEEE Computer Society Press, Washington, DC, 2007, pp. 305–318.
- [9] J.-Y. CAI, H. GUO, AND T. WILLIAMS, A Complete Dichotomy Rises from the Capture of Vanishing Signatures, preprint, arXiv:1204.6445, 2012.
- [10] J.-Y. CAI, S. HUANG, AND P. LU, From Holant to #CSP and Back: Dichotomy for Holant^e Problems, preprint, arXiv:1004.0803, 2010.
- [11] J.-Y. CAI AND P. LU, Holographic algorithms: From art to science, in STOC '07: Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing, ACM, New York, 2007, pp. 401–410.
- [12] J.-Y. CAI AND P. LU, Signature theory in holographic algorithms, in Algorithms and Computation, 19th International Symposium (ISAAC 2008), S.-H. Hong, H. Nagamochi, and T. Fukunaga, eds., Lecture Notes in Comput. Sci. 5369, Springer, New York, 2008, pp. 568– 579.
- [13] J.-Y. CAI, P. LU, AND M. XIA, Holographic algorithms by Fibonacci gates and holographic reductions for hardness, in FOCS '08: Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press, Washington, DC, 2008, pp. 644–653.
- [14] J.-Y. CAI, P. LU, AND M. XIA, A computational proof of complexity of some restricted counting problems, Theoret. Comput. Sci., 412 (2011), pp. 2468–2485.
- [15] J.-Y. CAI, P. LU, AND M. XIA, Holant problems and counting CSP, in Proceedings of the 41st Annual ACM Symposium on Theory of Computing, M. Mitzenmacher, ed., ACM, New York, 2009, pp. 715–724.
- [16] J.-Y. CAI, P. LU, AND M. XIA, Holographic algorithms with matchgates capture precisely tractable planar #CSP, in FOCS '10: Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Science Press, Piscataway, NJ, 2010, pp. 427–436.

- [17] A. CAYLEY, Sur les déterminates gauches, Crelle's J., 1854, pp. 38-93.
- [18] M. COOK AND J. BRUCK, Implementability Among Predicates, Tech. report, California Institute of Technology, Pasadena, CA, 2005.
- [19] G. CORNUÉJOLS, General factors of graphs, J. Combin. Theory Ser. B, 45 (1988), pp. 185–198.
- [20] N. CREIGNOU AND M. HERMANN, Complexity of generalized satisfiability counting problems, Inform. Comput., 125 (1996), pp. 1–12.
- [21] N. CREIGNOU, S. KHANNA, AND M. SUDAN, Complexity Classifications of Boolean Constraint Satisfaction Problems, SIAM Monogr. Discrete Math. Appl. 7, SIAM, Philadelphia, 2001.
- [22] C. T. J. DODSON AND T. POSTON, Tensor Geometry, Grad. Texts in Math. 130, Springer-Verlag, New York, 1991.
- [23] M. DYER, L. A. GOLDBERG, AND M. JERRUM, The complexity of weighted Boolean #CSP, SIAM J. Comput., 38 (2009), pp. 1970–1986.
- [24] M. E. DYER, L. A. GOLDBERG, AND M. PATERSON, On counting homomorphisms to directed acyclic graphs, J. ACM, 54 (2007), 27.
- [25] J. FABEN, The Complexity of Counting Solutions to Generalised Satisfiability Problems Modulo k, preprint, arXiv:0809.1836, 2008.
- [26] 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., 28 (1998), pp. 57–104.
- [27] L. A. GOLDBERG, M. GROHE, M. JERRUM, AND M. THURLEY, A complexity dichotomy for partition functions with mixed signs, SIAM J. Comput., 39 (2010), pp. 3336–3402.
- [28] H. GUO, S. HUANG, P. LU, AND M. XIA, The complexity of weighted boolean #CSP modulo k, in Proceedings of the 28th Annual Symposium on Theoretical Aspects of Computer Science, LIPIcs 9, T. Schwentick and C. Dürr, eds., Schloss Dagstuhl, Leibniz-Zentrum für Informatik, Wadern, Germany, 2011, pp. 249–260.
- [29] H. GUO, P. LU, AND L. G. VALIANT, The complexity of symmetric boolean parity Holant problems (extended abstract), in Proceedings of the 38th International Colloquium Conference on Automata, Languages, and Programming, L. Aceto, M. Henzinger, and J. Sgall, eds., Lect. Notes Comput. Sci. 6755, Springer-Verlag, Berlin, 2011, pp. 712–723.
- [30] S. HUANG AND P. LU, A Dichotomy for Real Weighted Holant Problems, manuscript, 2011.
- [31] M. KOWALCZYK AND J.-Y. CAI, Holant problems for regular graphs with complex edge functions, in Proceedings of the 27th Annual Symposium on Theoretical Aspects of Computer Science, LIPIcs 5, J.-Y. Marion and T. Schwentick, eds., Schloss Dagstuhl, Leibniz-Zentrum für Informatik, Wadern, Germany, 2010, pp. 525–536.
- [32] R. E. LADNER, On the structure of polynomial time reducibility, J. ACM, 22 (1975), pp. 155– 171.
- [33] L. LOVÁSZ, The factorization of graphs. II, Acta Math. Hungar., 23 (1972), pp. 223–246.
- [34] C. H. PAPADIMITRIOU AND S. ZACHOS, Two remarks on the power of counting, in Proceedings of the 6th GI-Conference on Theoretical Computer Science, Springer-Verlag, London, 1982, pp. 269–276.
- [35] T. J. SCHAEFER, The complexity of satisfiability problems, in Proceedings of the Tenth Annual ACM Symposium on Theory of Computing, ACM, New York, 1978, pp. 216–226.
- [36] S. TODA AND M. OGIWARA, Counting classes are at least as hard as the polynomial-time hierarchy, SIAM J. Comput., 21 (1992), pp. 316–328.
- [37] L. G. VALIANT, The complexity of computing the permanent, Theoret. Comput. Sci., 8 (1979), pp. 189–201.
- [38] L. G. VALIANT, Expressiveness of matchgates, Theoret. Comput. Sci., 289 (2002), pp. 457–471.
- [39] L. G. VALIANT, Quantum circuits that can be simulated classically in polynomial time, SIAM J. Comput., 31 (2002), pp. 1229–1254.
- [40] L. G. VALIANT, Holographic algorithms (extended abstract), in FOCS '04: Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press, Washington, DC, 2004, pp. 306–315.
- [41] L. G. VALIANT, Accidental algorithms, in FOCS '06: Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science, IEEE Computer Society Press, Washington, DC, 2006, pp. 509–517.
- [42] L. G. VALIANT, Holographic algorithms, SIAM J. Comput., 37 (2008), pp. 1565–1594.
- [43] L. G. VALIANT, Some observations on holographic algorithms, in Proceedings of the 9th Latin American Conference on Theoretical Informatics, A. López-Ortiz, ed., Lect. Notes Comput. Sci. 6034, Springer-Verlag, Berlin, 2010, pp. 577–590.
- [44] L. G. VALIANT AND V. V. VAZIRANI, NP is as easy as detecting unique solutions, Theoret. Comput. Sci., 47 (1986), pp. 85–93.

Downloaded 04/25/16 to 202.120.2.30. Redistribution subject to SIAM license or copyright; see http://www.siam.org/journals/ojsa.php