From Holant to #CSP and Back: Dichotomy for Holant^c Problems

Jin-Yi Cai[1⋆], Sangxia Huang[2⋆⋆], and Pinyan Lu[3]

[1] University of Wisconsin-Madison and Peking University. jyc@cs.wisc.edu
[2] Kungliga Tekniska högskolan (KTH). sangxia@kth.se
[3] Microsoft Research Asia. pinyanl@microsoft.com

Abstract. We explore the intricate interdependent relationship among counting problems, considered from three frameworks for such problems: Holant Problems, counting CSP and weighted H-colorings. We consider these problems for general complex valued functions that take Boolean inputs. We show that results from one framework can be used to derive results in another, and this happens in both directions. Holographic reductions discover an underlying unity, which is only revealed when these counting problems are investigated in the complex domain $\mathbb{C}$. We prove three complexity dichotomy theorems, leading to a general theorem for Holant^c problems. This is the natural class of Holant problems where one can assign constants 0 or 1. More specifically, given any signature grid on $G = (V, E)$ over a set $\mathscr{F}$ of symmetric functions, we completely classify the complexity to be in P or #P-hard, according to $\mathscr{F}$, of

$$\sum_{\sigma:E\rightarrow\{0,1\}} \prod_{v\in V} f_v(\sigma\mid_{E(v)}),$$

where $f_v \in \mathscr{F} \cup \{\mathbf{0}, \mathbf{1}\}$ ($\mathbf{0}$, $\mathbf{1}$ are the unary constant 0, 1 functions). Not only is holographic reduction the main tool, but also the final dichotomy is naturally stated in the language of holographic transformations. The proof goes through another dichotomy theorem on Boolean complex weighted #CSP.

1 Introduction

In order to study the complexity of counting problems, several interesting frameworks have been proposed. One is called counting Constraint Satisfaction Problems (#CSP) [1–3, 13, 17]. Another well studied framework is called H-coloring or Graph Homomorphism, which can be viewed as a special case of #CSP problems [4, 5, 14–16, 19, 20]. Recently, we proposed a new refined framework called Holant Problems [8, 10] inspired by Valiant's Holographic Algorithms [25, 26]. One reason such frameworks are interesting is because the language is *expressive* enough so that they can express many natural counting problems, while

⋆ Supported by NSF CCF-0830488 and CCF-0511679.

⋆⋆ Work partly done when the author was an undergraduate student at Shanghai Jiao Tong University and an intern student at Microsoft Research Asia.

specific enough so that we can prove *dichotomy theorems* (i.e., every problem in the class is either in P or #P-hard) [11]. By a theorem of Ladner, if P $\neq$ NP, or P $\neq$ $P^{\#P}$, then such a dichotomy for NP, or for #P, is *false*. Many natural counting problems can be expressed in all three frameworks. This includes counting the number of vertex covers, the number of k-colorings in a graph, and many others. However, some natural and important counting problems, such as counting the number of perfect matchings in a graph, *cannot* be expressed as a graph homomorphism function [18], but can be naturally expressed as a Holant Problem. Both #CSP and Graph Homomorphisms can be viewed as special cases of Holant Problems. The Holant framework of counting problems makes a finer complexity classification. A rich mathematical structure is uncovered in the Holant framework regarding the complexity of counting problems, which is sometimes difficult even to state in #CSP. This is particularly true when we apply holographic reductions [25, 26, 8].

We give a brief description of the Holant framework. [4] A *signature grid* $\Omega = (G, \mathscr{F}, \pi)$ is a tuple, where $G = (V, E)$ is a graph, and π labels each $v \in V$ with a function $f_v \in \mathscr{F}$. We consider all edge assignments (in this paper 0-1 assignments). An assignment σ for every $e \in E$ gives an evaluation $\prod_{v \in V} f_v(\sigma \mid_{E(v)})$, where $E(v)$ denotes the incident edges of v, and $\sigma \mid_{E(v)}$ denotes the restriction of σ to $E(v)$. The counting problem on the instance Ω is to compute

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

For example, consider the Perfect Matching problem on G. This problem corresponds to attaching the Exact-One function at every vertex of G, and then consider all 0-1 edge assignments. In this case, Holant_Ω counts the number of perfect matchings. If we use the At-Most-One function at every vertex, then we count all (not necessarily perfect) matchings. We use the notation Holant($\mathscr{F}$) to denote the class of Holant problems where all functions are given by $\mathscr{F}$.

To see that Holant is a more expressive framework, we show that #CSP can be simulated by Holant. Represent an instance of a #CSP problem by a bipartite graph where left hand side (LHS) are labeled by variables and right hand side (RHS) are labeled by constraint functions. Now the signature grid Ω on this bipartite graph is as follows: Every variable node on LHS is labeled with an Equality function, every constraint node on RHS is labeled with the given constraint. Then Holant_Ω is exactly the answer to the counting CSP problem. In effect, the Equality function on a node in LHS forces the incident edges to take the same value; this effectively reduces to a vertex assignment on LHS as in #CSP. We can show that #CSP is equivalent to Holant problems where Equality functions of k variables, for arbitrary k (denoted by $=_k$), are freely and implicitly available as constraints. However, this process cannot be reversed in general. While #CSP is the same as adding all $=_k$ to Holant, the effect of

[4] The term Holant was first used by Valiant in [25]. It denotes a sum which is a special case (corresponding to Perfect Matching) of Holant_Ω in the definition here [8, 10]. The term Holant emphasizes its relationship with holographic transformations.

making them freely available is non-trivial. From the lens of holographic transformations, $=_3$ is a full-fledged non-degenerate symmetric function of arity 3.

Starting from the Holant framework, rather than assuming Equality functions are free, one can consider new classes of counting problems which are difficult to express as #CSP problems. One such class, called Holant* Problems [10], is the class of Holant Problems where all unary functions are freely available. If we allow only two special unary functions $\mathbf{0}$ and $\mathbf{1}$ as freely available, then we obtain the family of counting problems called Holant^c Problems, which is even more appealing. This is the class of all Holant Problems (on Boolean variables) where one can set any particular edge (variable) to 0 or 1 in an input graph.

Previously a dichotomy theorem was proved for Holant$^*(\mathscr{F})$, where $\mathscr{F}$ is any set of complex-valued symmetric functions [10]. It was used to prove a dichotomy theorem for #CSP in [10]. For Holant$^c(\mathscr{F})$ we were able to prove a dichotomy theorem valid only real-valued functions [10]. In this paper we manage to traverse in the other direction, going from #CSP to Holant Problems. First we establish a dichotomy theorem for a special Holant class. Second we prove a more general dichotomy for bipartite Holant Problems. Finally by going through #CSP, we prove a dichotomy theorem for complex-valued Holant^c Problems. Now we describe our results in more detail.

A symmetric function $f : \{0,1\}^k \to \mathbb{C}$ will be written as $[f_0, f_1, \ldots, f_k]$, where f_j is the value of f on inputs of Hamming weight j. Our first main result (in Section 3) is a dichotomy theorem for Holant($\mathscr{F}$), where $\mathscr{F}$ contains a single ternary function $[x_0, x_1, x_2, x_3]$. More generally, we can apply holographic reductions to prove a dichotomy theorem for Holant($[y_0, y_1, y_2]|[x_0, x_1, x_2, x_3]$) defined on 2-3 regular bipartite graphs. Here the notation indicates that every vertex of degree 2 on LHS has label $[y_0, y_1, y_2]$ and every vertex of degree 3 on RHS has label $[x_0, x_1, x_2, x_3]$. This is the foundation of the remaining two dichotomy results in this paper. Previously we proved a dichotomy theorem for Holant($[y_0, y_1, y_2]|[x_0, x_1, x_2, x_3]$), when all x_i, y_j take values in $\{0, 1\}$ [8]. Kowalczyk extended this to $\{-1, 0, 1\}$ in [21]. In [9], we gave a dichotomy theorem for Holant($[y_0, y_1, y_2]|[1, 0, 0, 1]$), where y_0, y_1, y_3 take arbitrary real values. Finally this last result was extended to arbitrary complex numbers [22]. Our result here is built upon these results, especially [22].

Our second result (Section 4) is a dichotomy theorem, under a mild condition, for bipartite Holant problems Holant($\mathscr{F}_1|\mathscr{F}_2$) (see Sec. 2 for definitions). Under this mild condition, we first use holographic reductions to transform it to Holant($\mathscr{F}_1'|\mathscr{F}_2'$), where we transform some non-degenerate function $[x_0, x_1, x_2, x_3]$ $\in \mathscr{F}_2$ to the Equality function $(=_3) = [1, 0, 0, 1] \in \mathscr{F}_2'$. Then we prove that we can "realize" the binary Equality function $(=_2) = [1, 0, 1]$ in the left side and reduce the problem to #CSP($\mathscr{F}_1' \cup \mathscr{F}_2'$). This is a new proof approach. Previously in [10], we reduced a #CSP problem to a Holant problem and obtained results for #CSP. Here, we go the opposite way, using results for #CSP to prove dichotomy theorems for Holant problems. This is made possible by our complete dichotomy theorem for Boolean complex weighted #CSP [10]. We note that proving this over $\mathbb{C}$ is crucial, as holographic reductions naturally go beyond $\mathbb{R}$.

We also note that our dichotomy theorem here does not require the functions in $\mathscr{F}_1$ or $\mathscr{F}_2$ to be symmetric. This will be useful in the future.

Our third and main result, also the initial motivation of this work, is a dichotomy theorem for symmetric complex Holant^c problems. This improves our previous result in [10]. We made a conjecture in [10] about the dichotomy theorem of Holant^c for symmetric complex functions. It turns out that this conjecture is not correct as stated. For example, Holant$^c([1, 0, i, 0])$ is tractable (according to our new theorem), but not included in the tractable cases by the conjecture. After isolating these new tractable cases we prove *everything else* is #P-hard. Generally speaking, non-trivial and previously unknown tractable cases are what make dichotomy theorems particularly interesting, but at the same time make them more difficult to prove (especially for hardness proofs, which must "carve out" exactly what's left). The proof approach here is also different from that of [10]. In [10], the idea is to interpolate all unary functions and then use the results for Holant* Problems. Here we first prove that we can realize some non-degenerate ternary function, for which we can use the result of our first dichotomy theorem. Then we use our second dichotomy theorem to further reduce the problem to #CSP and obtain a dichotomy theorem for Holant^c.

The study of Holant Problems is strongly influenced by the development of holographic algorithms [25, 26, 7, 8]. Holographic reduction is a primary technique in the proof of these dichotomies, for both the tractability part and the hardness part. More than that—and this seems to be the first instance—holographic reduction even provides the correct language for the *statements* of these dichotomies. Without using holographic reductions, it is not easy to even fully describe what are the tractable cases in the dichotomy theorem. Another interesting observation is that by employing holographic reductions, complex numbers appear naturally and in an essential way. Even if one is only interested in integer or real valued counting problems, in the complex domain $\mathbb{C}$ the picture becomes whole. "It has been written that the shortest and best way between two truths of the real domain often passes through the imaginary one." —*Jacques Hadamard*.

2 Preliminaries

Our functions take values in $\mathbb{C}$ by default. Strictly speaking complexity results should be restricted to computable numbers in the Turing model; but it is more convenient to express this over $\mathbb{C}$. We say a problem is tractable if it is computable in P. The framework of Holant Problems is defined for functions mapping any $[q]^k \to \mathbb{C}$ for a finite q. Our results in this paper are for the Boolean case $q = 2$.

Let $\mathscr{F}$ be a set of such functions. A *signature grid* $\Omega = (G, \mathscr{F}, \pi)$ is a tuple, where $G = (V, E)$ is a graph, and $\pi : V \to \mathscr{F}$ labels each $v \in V$ with a function $f_v \in \mathscr{F}$. The Holant problem on instance Ω is to compute $\text{Holant}_\Omega = \sum_{\sigma:E\to\{0,1\}} \prod_{v\in V} f_v(\sigma \mid_{E(v)})$, a sum over all 0-1 edge assignments, of the products of the function evaluations at each vertex. Here $f_v(\sigma \mid_{E(v)})$ denotes the value of f_v evaluated using the restriction of σ to the incident edges $E(v)$ of v. A function f_v can be represented as a truth table. It will be more convenient to denote it as a vector in $\mathbb{C}^{2^{\deg(v)}}$, or a tensor in $(\mathbb{C}^2)^{\otimes \deg(v)}$, when we perform holographic tranformations. We also call it a *signature*. We denote by $=_k$ the EQUALITY signature of arity k. A symmetric function f on k Boolean variables can be expressed by $[f_0, f_1, \ldots, f_k]$, where f_j is the value of f on inputs of Hamming weight j. Thus, for example, $\mathbf{0} = [1, 0]$, $\mathbf{1} = [0, 1]$ and $(=_k) = [1, 0, \ldots, 0, 1]$ (with $(k-1)$ 0's).

Definition 1. *Given a set of signatures $\mathscr{F}$, we define* Holant($\mathscr{F}$)*:*
Input: A signature grid $\Omega = (G, \mathscr{F}, \pi)$;
Output: Holant$_\Omega$.

We would like to characterize the complexity of Holant problems in terms of its signature set $\mathscr{F}$. Some special families of Holant problems have already been widely studied. For example, if $\mathscr{F}$ contains all EQUALITY signatures $\{=_1, =_2, =_3, \ldots\}$, then this is exactly the weighted #CSP problem. In [10], we also introduced the following two special families of Holant problems by assuming some signatures are freely available.

Definition 2. *Let $\mathscr{U}$ denote the set of all unary signatures. Then* Holant$^*(\mathscr{F})$ = Holant($\mathscr{F} \cup \mathscr{U}$).

Definition 3. *For any set of signatures $\mathscr{F}$,* Holant$^c(\mathscr{F})$ = Holant($\mathscr{F} \cup \{\mathbf{0}, \mathbf{1}\}$).

Replacing a signature $f \in \mathscr{F}$ by a constant multiple cf, where $c \neq 0$, does not change the complexity of Holant($\mathscr{F}$). It introduces a global factor to Holant$_\Omega$.

An important property of a signature is whether it is degenerate.

Definition 4. *A signature is degenerate iff it is a tensor product of unary signatures. In particular, a symmetric signature in $\mathscr{F}$ is degenerate iff it can be expressed as $\lambda[x, y]^{\otimes k}$.*

We use $\mathscr{A}$ to denote the set of functions which has the form $\chi_{[AX=0]} \cdot i^{\sum_{j=1}^{n} \langle \alpha_j, X \rangle}$, where $i = \sqrt{-1}$, $X = (x_1, x_2, \ldots, x_k, 1)$, A is matrix over $\mathbb{F}_2$, α_j is a vector over $\mathbb{F}_2$, and χ is a 0-1 indicator function such that $\chi_{[AX=0]}$ is 1 iff $AX = 0$.

We use $\mathscr{P}$ to denote the set of functions which can be expressed as a product of unary functions, binary equality functions ($[1, 0, 1]$) and binary disequality functions ($[0, 1, 0]$).

Theorem 1. *[10] Suppose $\mathscr{F}$ is a set of functions mapping Boolean inputs to complex numbers. If $\mathscr{F} \subseteq \mathscr{A}$ or $\mathscr{F} \subseteq \mathscr{P}$, then #CSP($\mathscr{F}$) is computable in polynomial time. Otherwise, #CSP($\mathscr{F}$) is* #P*-hard.*

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

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

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

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

3 Dichotomy Theorem for Ternary Signatures

In this section, we consider the complexity of $\text{Holant}([x_0, x_1, x_2, x_3])$. It is trivially tractable if $[x_0, x_1, x_2, x_3]$ is degenerate, so in the following we always assume that it is non-degenerate. Similar to that in [10], we classify the sequence $[x_0, x_1, x_2, x_3]$ into one of the following three categories (with the convention that $\alpha^0 = 1$, and $k\alpha^{k-1} = 0$ if $k = 0$, even when $\alpha = 0$): (1) $x_k = \alpha_1^{3-k}\alpha_2^k + \beta_1^{3-k}\beta_2^k$, where $\det \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix} \neq 0$; (2) $x_k = Ak\alpha^{k-1} + B\alpha^k$, where $A \neq 0$; (3) $x_k = A(3-k)\alpha^{2-k} + B\alpha^{3-k}$, where $A \neq 0$. We call the first category as the *generic* case, the second and third one as the *double-root* case.

For the *generic* case, we can apply a holographic reduction using $T = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{bmatrix}$, and have $\text{Holant}([x_0, x_1, x_2, x_3]) \equiv_{\text{T}} \text{Holant}([y_0, y_1, y_2]|[1, 0, 0, 1])$, where $[y_0, y_1, y_2] = [1, 0, 1]T^{\otimes 2}$. (We note that $[x_0, x_1, x_2, x_3] = T^{\otimes 3}[1, 0, 0, 1]$.) Therefore we only need to give a dichotomy for $\text{Holant}([y_0, y_1, y_2]|[1, 0, 0, 1])$, which has been proved in [22]; we quote the theorem here.

Theorem 2. ([22]) *The problem Holant*$([y_0, y_1, y_2]|[1, 0, 0, 1])$ *is #P-hard for all* $y_0, y_1, y_2 \in \mathbb{C}$ *except in the following cases, for which the problem is in P: (1)* $y_1^2 = y_0 y_2$*; (2)* $y_0^{12} = y_1^{12}$ *and* $y_0 y_2 = -y_1^2$ *(* $y_1 \neq 0$*); (3)* $y_1 = 0$*; and (4)* $y_0 = y_2 = 0$.

For the *double-root* case, we have the following lemma.

Lemma 1. *Let* $x_k = Ak\alpha^{k-1} + B\alpha^k$*, where* $A \neq 0$ *and* $k = 0, 1, 2, 3$*. Unless* $\alpha^2 = -1$*,* $\mathrm{Holant}([x_0, x_1, x_2, x_3])$ *is #P-hard. On the other hand, if* $\alpha = \pm i$*, then the problem is in P.*

Proof. If $\alpha = \pm i$, the signature $[x_0, x_1, x_2, x_3]$ satisfies the recurrence relation $x_{k+2} = \alpha x_{k+1} + x_k$, where $k = 0, 1$. This is a generalized Fibonacci signature (see [8]). Thus it is in P by holographic algorithms [8] using Fibonacci gates.

Now we assume that $\alpha \neq \pm i$. We first apply an *orthogonal* holographic transformation. The crucial observation is that we can view $\mathrm{Holant}([x_0, x_1, x_2, x_3])$ as the bipartite $\mathrm{Holant}([1, 0, 1]|[x_0, x_1, x_2, x_3])$ and an orthogonal transformation $T \in \mathbf{O}_2(\mathbb{C})$ keeps $(=_2) = [1, 0, 1]$ invariant: $[1, 0, 1]T^{\otimes 2} = [1, 0, 1]$. By a suitable orthogonal transformation T, we can transform $[x_0, x_1, x_2, x_3]$ to $[v, 1, 0, 0]$ for some $v \in \mathbb{C}$, up to a scalar. (Details are in the full paper [6].) So the complexity of $\mathrm{Holant}([x_0, x_1, x_2, x_3])$ is the same as $\mathrm{Holant}([v, 1, 0, 0])$.

Next we prove that $\mathrm{Holant}([v, 1, 0, 0])$ is #P-hard for all $v \in \mathbb{C}$. First, for $v = 0$, $\mathrm{Holant}([0, 1, 0, 0])$ is #P-hard, because it is the problem of counting all perfect matchings on 3-regular graphs [12]. Second, let $v \neq 0$. We can realize $[v^3 + 3v, v^2 + 1, v, 1]$ by connecting three $[v, 1, 0, 0]$'s as a triangle, so it is enough to prove that $\mathrm{Holant}([v^3 + 3v, v^2 + 1, v, 1])$ is #P-hard. In tensor product notation this signature is $\frac{1}{2}\left(\begin{bmatrix} v+1 \\ 1 \end{bmatrix}^{\otimes 3} + \begin{bmatrix} v-1 \\ 1 \end{bmatrix}^{\otimes 3}\right)$. Then

$$
\begin{aligned}
\mathrm{Holant}([v^3 + 3v, v^2 + 1, v, 1]) &\equiv_{\mathrm{T}} \mathrm{Holant}([1, 0, 1]|[v^3 + 3v, v^2 + 1, v, 1]) \\
&\equiv_{\mathrm{T}} \mathrm{Holant}([v^2 + 2v + 2, v^2, v^2 - 2v + 2]|[1, 0, 0, 1])
\end{aligned}
$$

where the second step is a holographic reduction using $\begin{bmatrix} v+1 & v-1 \\ 1 & 1 \end{bmatrix}$. We can apply Theorem 2 to $\mathrm{Holant}([v^2 + 2v + 2, v^2, v^2 - 2v + 2]|[1, 0, 0, 1])$. Checking against the four exceptions we find that they are all impossible. Therefore $\mathrm{Holant}([v^3 + 3v, v^2 + 1, v, 1])$ is #P-hard, and so is $\mathrm{Holant}([v, 1, 0, 0])$ for all $v \in \mathbb{C}$.

By Theorem 2 and Lemma 1, we have a complete dichotomy theorem for $\mathrm{Holant}([x_0, x_1, x_2, x_3])$ and for bipartite $\mathrm{Holant}([y_0, y_1, y_2]|[x_0, x_1, x_2, x_3])$.

Theorem 3. $\mathrm{Holant}([x_0, x_1, x_2, x_3])$ *is #P-hard unless* $[x_0, x_1, x_2, x_3]$ *satisfies one of the following conditions, in which case the problem is in P:*

1. $[x_0, x_1, x_2, x_3]$ *is degenerate;*
2. *There is a* 2×2 *matrix* T *such that* $[x_0, x_1, x_2, x_3] = T^{\otimes 3}[1, 0, 0, 1]$ *and* $[1, 0, 1]T^{\otimes 2}$ *is in* $\mathscr{A} \cup \mathscr{P}$*;*
3. *For* $\alpha \in \{2i, -2i\}$*,* $x_2 + \alpha x_1 - x_0 = 0$ *and* $x_3 + \alpha x_2 - x_1 = 0$.

Theorem 4. $\operatorname{Holant}([y_0, y_1, y_2] | [x_0, x_1, x_2, x_3])$ *is #P-hard unless* $[x_0, x_1, x_2, x_3]$ *and* $[y_0, y_1, y_2]$ *satisfy one of the following conditions, in which case the problem is in P:*

1. $[x_0, x_1, x_2, x_3]$ *is degenerate;*
2. *There is a* 2×2 *matrix* T *such that* $[x_0, x_1, x_2, x_3] = T^{\otimes 3}[1, 0, 0, 1]$ *and* $[y_0, y_1, y_2]T^{\otimes 2}$ *is in* $\mathscr{A} \cup \mathscr{P}$*;*
3. *There is a* 2×2 *matrix* T *such that* $[x_0, x_1, x_2, x_3] = T^{\otimes 3}[1, 1, 0, 0]$ *and* $[y_0, y_1, y_2]T^{\otimes 2}$ *is of form* $[0, *, *]$*;*
4. *There is a* 2×2 *matrix* T *such that* $[x_0, x_1, x_2, x_3] = T^{\otimes 3}[0, 0, 1, 1]$ *and* $[y_0, y_1, y_2]T^{\otimes 2}$ *is of form* $[*, *, 0]$.

4 Reductions between Holant and #CSP

In this section, we extend the dichotomies in Section 3 for a single ternary signature to a set of signatures. We will give a dichotomy for $\operatorname{Holant}([x_0, x_1, x_2, x_3] \cup \mathscr{F})$, or more generally for $\operatorname{Holant}([y_0, y_1, y_2] \cup \mathscr{G}_1 | [x_0, x_1, x_2, x_3] \cup \mathscr{G}_2)$, where $[y_0, y_1, y_2]$ and $[x_0, x_1, x_2, x_3]$ are non-degenerate. In this section, we focus on the generic case of $[x_0, x_1, x_2, x_3]$, and the double root case will be handled in the next section in Lemma 3. For the generic case, we can apply a holographic reduction to transform $[x_0, x_1, x_2, x_3]$ to $[1, 0, 0, 1]$. Therefore we only need to give a dichotomy for Holant problems of the form $\operatorname{Holant}([y_0, y_1, y_2] \cup \mathscr{G}_1 | [1, 0, 0, 1] \cup \mathscr{G}_2)$, where $[y_0, y_1, y_2]$ is non-degenerate. We make one more observation: for any $T \in \mathscr{T}_3 \triangleq \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & \omega \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & \omega^2 \end{bmatrix} \right\}$, where $\omega = \omega_3 = e^{2\pi i/3}$, we have

$$\operatorname{Holant}([y_0, y_1, y_2] | [1, 0, 0, 1] \cup \mathscr{F}) \equiv_{\mathrm{T}} \operatorname{Holant}([y_0, y_1, y_2]T^{\otimes 2} | [1, 0, 0, 1] \cup T^{-1}\mathscr{F}).$$

As a result, we can normalize $[y_0, y_1, y_2]$ by a holographic reduction with any $T \in \mathscr{T}_3$. In particular, we call a symmetric binary signature $[y_0, y_1, y_2]$ *normalized* if $y_0 = 0$ or it is not the case that y_2 is y_0 times a t-th primitive root of unity, and $t = 3t'$ where $\gcd(t', 3) = 1$. We can always normalize $[y_0, y_1, y_2]$ by applying a transformation $\begin{bmatrix} 1 & 0 \\ 0 & \omega^k \end{bmatrix} \in \mathscr{T}_3$. So in the following, we only deal with normalized $[y_0, y_1, y_2]$. In one case, we also need to normalize a unary signature $[x_0, x_1]$, namely $x_0 = 0$ or x_1 is not a multiple of x_0 by a t-th primitive root of unity, and $t = 3t'$ where $\gcd(t', 3) = 1$. Again we can normalize the unary signature by a suitable $T \in \mathscr{T}_3$.

Theorem 5. *Let* $[y_0, y_1, y_2]$ *be a normalized and non-degenerate signature. And in the case of* $y_0 = y_2 = 0$*, we further assume that* $\mathscr{G}_1$ *contains a unary signature* $[a, b]$*, which is normalized and* $ab \neq 0$*. Then*

$$\operatorname{Holant}([y_0, y_1, y_2] \cup \mathscr{G}_1 | [1, 0, 0, 1] \cup \mathscr{G}_2) \equiv_{\mathrm{T}} \#\mathrm{CSP}([y_0, y_1, y_2] \cup \mathscr{G}_1 \cup \mathscr{G}_2).$$

Thus, $\operatorname{Holant}([y_0, y_1, y_2] \cup \mathscr{G}_1 | [1, 0, 0, 1] \cup \mathscr{G}_2)$ *is #P-hard unless* $[y_0, y_1, y_2] \cup \mathscr{G}_1 \cup \mathscr{G}_2 \subseteq \mathscr{P}$ *or* $[y_0, y_1, y_2] \cup \mathscr{G}_1 \cup \mathscr{G}_2 \subseteq \mathscr{A}$*, in which cases the problem is in P.*

This dichotomy is an important reduction step in the proof of our dichotomy theorem for Holant^c. It is also interesting in its own right as a connection between Holant and #CSP. The assumption on signature normalization in the statement of the theorem is without loss of generality. For a non-normalized signature, we can first normalize it and then apply the dichotomy criterion. (Note that when $y_0 = y_2 = 0$ the normalization on $[a, b]$ keeps $[y_0, y_1, y_2]$ normalized.) The additional assumption of the existence of a unary signature $[a, b]$ with $ab \neq 0$ circumvents a technical difficulty, and finds a circuitous route to the proof of our main dichotomy theorem for Holant^c. For Holant^c, the needed unary signature will be produced from $[1, 0]$ and $[0, 1]$. We also note that we do not require the signatures in $\mathscr{G}_1$ and $\mathscr{G}_2$ to be symmetric.

One direction in Theorem 5, from Holant to #CSP, is straightforward. Thus our main claim is a reduction from #CSP to these bipartite Holant problems. Start with #CSP($[y_0, y_1, y_2] \cup \mathscr{G}_1 \cup \mathscr{G}_2$) $\equiv_{\mathrm{T}}$ Holant($[y_0, y_1, y_2] \cup \mathscr{G}_1 \cup \mathscr{G}_2 | \{=_k : k \geq 1\}$). The approach is to construct the binary equality $[1, 0, 1] = (=_2)$ in LHS in the Holant problem. As soon as we have $[1, 0, 1]$ in LHS, together with $[1, 0, 0, 1] = (=_3)$ in RHS, we get equality gates of all arities $(=_k)$ in RHS. Then with the help of $[1, 0, 1]$ in LHS we can transfer $\mathscr{G}_2$ to LHS.

If the problem Holant($[y_0, y_1, y_2] | [1, 0, 0, 1]$) is already #P-hard, then for any $\mathscr{G}_1$ and $\mathscr{G}_2$, it is #P-hard. So we only need to consider the cases, where Holant($[y_0, y_1, y_2] | [1, 0, 0, 1]$) is not #P-hard. For this, we again use Theorem 2 from [22]. The first tractable case $y_1^2 = y_0 y_2$ is degenerate, which does not apply here. The remaining three tractable cases are proved separately and the proofs can be found in the full paper [6].

5 Dichotomy Theorem for Complex Holant^c Problems

In this section, we prove our main result, a dichotomy theorem for Holant^c problems with complex valued symmetric signatures over Boolean variables, which is stated as Theorem 6. The proof crucially uses the dichotomies proved in the previous two sections. We first prove in Lemma 2 that we can always realize a non-degenerate ternary signature except in some trivial cases. With this non-degenerate ternary signature, we can immediately prove #P-hardness if it is not of one of the tractable cases in Theorem 3. For tractable ternary signatures, we use Theorem 5 to extend the dichotomy theorem to the whole signature set. In Theorem 5, we only considered the generic case of the ternary function. The double-root case is handled here in Lemma 3.

Lemma 2. *Given any set of symmetric signatures $\mathscr{F}$ which contains $[1, 0]$ and $[0, 1]$, we can construct a non-degenerate symmetric ternary signature $X = [x_0, x_1, x_2, x_3]$, except in the following two trivial cases:*

1. *Any non-degenerate signature in $\mathscr{F}$ is of arity at most 2;*
2. *In $\mathscr{F}$, all unary signatures are of form $[x, 0]$ or $[0, x]$; all binary signatures are of form $[x, 0, y]$ or $[0, x, 0]$; and all signatures of arity greater than 2 are of form $[x, 0, \ldots, 0, y]$.*

Proof (Sketch). Suppose Case 1 does not hold, and let $X \triangleq [x_0, x_1, \ldots, x_m] \in \mathscr{F}$ be a non-degenerate signature of arity at least 3. It must be that all ternary sub-signatures are degenerate, otherwise we are done. Then we can show that X must be of the form $[x_0, 0, \ldots, 0, x_m]$, where $x_0 x_m \neq 0$. If we have a unary signature, or a unary sub-signature of a binary signature, of the form $[a, b]$ $(ab \neq 0)$, we can connect this signature to $m-3$ dangling edges of X to get a non-degenerate ternary signature $[x, 0, 0, y]$, and we are done. Otherwise, we are in Case 2.

We next consider the double root case for a non-degenerate $X = [x_0, x_1, x_2, x_3]$. By Lemma 1, Holant(X) is already #P-hard unless the double eigenvalue is i or $-i$. Then, $x_{k+2} + \alpha x_{k+1} - x_k = 0$ for $k = 0, 1$, where $\alpha = \pm 2i$.

Lemma 3. *Let $X = [x_0, x_1, x_2, x_3]$ be a non-degenerate complex signature satisfying $x_{k+2} + \alpha x_{k+1} - x_k = 0$ for $k = 0, 1$, where $\alpha = \pm 2i$. Let $Y = [y_0, y_1, y_2]$ be a non-degenerate binary signature. Then Holant$(Y|X)$ is #P-hard unless $y_2 + \alpha y_1 - y_0 = 0$ (in which case Holant$(\{X, Y\})$ is in P by Fibonacci gates).*

Proof (Sketch). We prove this result for $\alpha = -2i$. The other case is similar.

We have $X = T^{\otimes 3}[1, 1, 0, 0]^{\mathrm{T}}$, where $T = \left[\begin{smallmatrix} 1 & \frac{B-1}{3} \\ i & A + \frac{B-1}{3} i \end{smallmatrix}\right]$, $A \neq 0$. By expressing $\left[\begin{smallmatrix} y_0 & y_1 \\ y_1 & y_2 \end{smallmatrix}\right] = T_0^{\mathrm{T}} T_0$, which is always possible for some non-singular $T_0 = \left[\begin{smallmatrix} a & c \\ b & d \end{smallmatrix}\right]$, we have $Y = [1, 0, 1] T_0^{\otimes 2}$. Thus we apply a holographic reduction and have Holant$(Y|X) \equiv_{\mathrm{T}}$ Holant$([1, 0, 1] | (T_0 T)^{\otimes 3} [1, 1, 0, 0]^{\mathrm{T}})$. Next, we try to use an orthogonal matrix to transform $T_0 T$ to be upper-triangular. We show that we can do this, except for the tractable cases. This leads to a reduction from Holant$([v, 1, 0, 0])$ to Holant$([1, 0, 1] | (T_0 T)^{\otimes 3} [1, 1, 0, 0]^{\mathrm{T}})$ and therefore to Holant$(Y|X)$, for some v. By Lemma 1, Holant$([v, 1, 0, 0])$ is #P-hard.

Theorem 6. *Let $\mathscr{F}$ be a set of complex symmetric signatures.* Holant$^c(\mathscr{F})$ *is #P-hard unless $\mathscr{F}$ satisfies one of the following conditions, in which case it is tractable:*

1. Holant$^*(\mathscr{F})$ *is tractable (for which we have an effective dichotomy in [10]);*
2. *There exists a $T \in \mathscr{T}$ such that $\mathscr{F} \subseteq T\mathscr{A}$, where*

$$\mathscr{T} \triangleq \{T \mid [1, 0, 1] T^{\otimes 2}, [1, 0] T, [0, 1] T \in \mathscr{A}\}$$

Proof. First of all, if $\mathscr{F}$ is an exceptional case of Lemma 2, we know that Holant$^*(\mathscr{F})$ is tractable and we are done. Now we can assume that we can construct a non-degenerate symmetric ternary signature $X = [x_0, x_1, x_2, x_3]$ and the problem is equivalent to Holant$^c(\mathscr{F} \cup \{X\})$. As discussed in Section 3, there are three categories for X and we only need to consider the first two: (1) $x_k = \alpha_1^{3-k} \alpha_2^k + \beta_1^{3-k} \beta_2^k$; (2) $x_k = A k \alpha^{k-1} + B \alpha^k$, where $A \neq 0$.

Case 1: $x_k = \alpha_1^{3-k} \alpha_2^k + \beta_1^{3-k} \beta_2^k$. In this case, $X = T^{\otimes 3}[1, 0, 0, 1]^T$, where $T = \left[\begin{smallmatrix} \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \end{smallmatrix}\right]$. (Note that we can replace T by $T\left[\begin{smallmatrix} 1 & 0 \\ 0 & \omega^j \end{smallmatrix}\right]$, $0 \leq j \leq 2$, and $X = T^{\otimes 3}[1, 0, 0, 1]^T$ still holds.) So we have the following reduction chain,

$$\begin{aligned} \text{Holant}^c(\mathscr{F}) &\equiv_{\mathrm{T}} \text{Holant}^c(\mathscr{F} \cup \{X\}) \equiv_{\mathrm{T}} \text{Holant}(\mathscr{F} \cup \{X, [1,0], [0,1]\}) \\ &\equiv_{\mathrm{T}} \text{Holant}(\{[1,0,1], [1,0], [0,1]\} | \mathscr{F} \cup \{X\}) \\ &\equiv_{\mathrm{T}} \text{Holant}(\{[1,0,1] T^{\otimes 2}, [1,0] T, [0,1] T\} | [1,0,0,1] \cup T^{-1} \mathscr{F}). \end{aligned}$$

Since $[1,0,1]T^{\otimes 2}$ is a non-degenerate binary signature, we can apply Theorem 5. (We replace T by $T\begin{bmatrix} 1 & 0 \\ 0 & \omega^j \end{bmatrix}$, $0 \le j \le 2$, to normalize $[1,0,1]T^{\otimes 2}$, if needed.) We need to verify is that when $[1,0,1]T^{\otimes 2} = [\alpha_1^2 + \alpha_2^2, \alpha_1\beta_1 + \alpha_2\beta_2, \beta_1^2 + \beta_2^2]$ is of the form $[0,*,0]$, at least one of $[1,0]T = [\alpha_1, \beta_1]$ or $[0,1]T = [\alpha_2, \beta_2]$ has both entries non-zero. If not, we would have $\alpha_1\beta_1 = 0$ and $\alpha_2\beta_2 = 0$, which implies that $[1,0,1]T^{\otimes 2} = [0,0,0]$, a contradiction. (We may again replace T by $T\begin{bmatrix} 1 & 0 \\ 0 & \omega^j \end{bmatrix}$, $0 \le j \le 2$, to normalize this unary, if needed, which does not conflict with the normalization of $[1,0,1]T^{\otimes 2}$.) Therefore, by Theorem 5, we know that the problem is #P-hard unless $[1,0,1]T^{\otimes 2} \cup T^{-1}\mathscr{F} \subseteq \mathscr{P}$ or $\{[1,0,1]T^{\otimes 2}, [1,0]T, [0,1]T\} \cup T^{-1}\mathscr{F} \subseteq \mathscr{A}$. In the first case, Holant$^*(\mathscr{F})$ is tractable; and the second case is equivalent to having $T \in \mathscr{T}$ satisfying $\mathscr{F} \subseteq T\mathscr{A}$.

Case 2: $x_k = Ak\alpha^{k-1} + B\alpha^k$, where $A \neq 0$. In this case, if $\alpha \neq \pm i$, the problem is #P-hard by Lemma 1 and we are done. Now we consider the case $\alpha = i$ (the case $\alpha = -i$ is similar). Consider the following Equation

$$z_{k+2} - 2iz_{k+1} - z_k = 0. \tag{1}$$

We note that $X = [x_0, x_1, x_2, x_3]$ satisfies this equation for $k = 0, 1$. If all non-degenerate signatures $Z = [z_0, z_1, \ldots, z_m]$ in $\mathscr{F}$ with arity $m \ge 2$ fulfill

Condition: Z satisfies Equation (1) for $k = 0, 1, \ldots, m-2$

then this is the second tractable case in the Holant* dichotomy theorem in [10] and we are done. So suppose this is not the case, and $Z = [z_0, z_1, \ldots, z_m] \in \mathscr{F}$, for some $m \ge 2$, is a non-degenerate signature that does not satisfy this Condition. By Lemma 3, if any non-degenerate sub-signature $[z_k, z_{k+1}, z_{k+2}]$ does not satisfy Equation (1), then, together with X which does satisfy (1), we know that the problem is #P-hard and we are done. So we assume every non-degenerate sub-signature $[z_k, z_{k+1}, z_{k+2}]$ of Z satisfies (1). In particular $m \ge 3$, and there exists some binary sub-signature of Z that is degenerate and does not satisfy (1).

If all binary sub-signatures of Z are degenerate (but Z itself is not), we claim that Z has the form $[z_0, 0, \ldots, 0, z_m]$, where $z_0 z_m \neq 0$. Then we can produce some $[a, 0, 0, b]$, $ab \neq 0$, and reduce to Case 1. Otherwise, we can find a ternary sub-signature $[z_k, z_{k+1}, z_{k+2}, z_{k+3}]$ (or its reversal) where $[z_k, z_{k+1}, z_{k+2}]$ is degenerate and $[z_{k+1}, z_{k+2}, z_{k+3}]$ is non-degenerate and thus satisfies $-z_{k+1} - 2iz_{k+2} + z_{k+3} = 0$. Then either we have got an instance of Case 1 or we could prove #P-hardness directly by Lemma 1. (Details are in the full paper [6].)

References

1. Bulatov, A.: A dichotomy theorem for constraint satisfaction problems on a 3-element set. *J. ACM*, 53(1):66–120, 2006.
2. Bulatov, A.: The complexity of the counting constraint satisfaction problem. In *ICALP (1)*, volume 5125 of *Lecture Notes in Computer Science*, 646–661, 2008.
3. Bulatov, A., Dalmau, V.: Towards a dichotomy theorem for the counting constraint satisfaction problem. *Inf. Comput.*, 205(5):651–678, 2007.
4. Bulatov, A., Grohe, M.: The complexity of partition functions. *Theor. Comput. Sci.*, 348(2-3):148–186, 2005.

5. Cai, J-Y., Chen, X., Lu, P.: Graph homomorphisms with complex values: A dichotomy theorem. In Proceedings of *ICALP(1)*, pages 275–286, 2010.
6. Cai, J-Y., Huang, S., Lu, P.: From Holant To #CSP And Back: Dichotomy For Holant^c Problems. CoRR, abs/1004.0803, 2010, http://arxiv.org/abs/1004.0803.
7. Cai, J-Y., Lu, P.: Holographic algorithms: from art to science. In *STOC '07: Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 401–410, 2007.
8. Cai, J-Y., Lu, P., Xia, M.: 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*, pages 644–653, 2008.
9. Cai, J-Y., Lu, P., Xia, M.: A Computational Proof of Complexity of Some Restricted Counting Problems In Proceedings of *TAMC*, pages 138-149, 2009.
10. Cai, J-Y., Lu, P., Xia, M.: Holant Problems and Counting CSP. In *STOC '09: Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 715-724, 2009.
11. Creignou, N., Khanna, N., Sudan, M.: *Complexity classifications of boolean constraint satisfaction problems.* SIAM Monographs on Discrete Mathematics and Applications, 2001.
12. Dagum, P., Luby, M.: Approximating the permanent of graphs with large factors. *Theor. Comput. Sci.*, 102:283–305, 1992.
13. Dyer, M.E., Goldberg, L.A., Jerrum, M.: The complexity of weighted boolean #CSP. *CoRR*, abs/0704.3683, 2007.
14. Dyer, M.E., Goldberg, L.A., Paterson, M.: On counting homomorphisms to directed acyclic graphs. *J. ACM*, 54(6), 2007.
15. Dyer, M.E., Greenhill, C.S.: The complexity of counting graph homomorphisms (extended abstract). In Proceedings of *SODA*, pages 246–255, 2000.
16. Dyer, M.E., Greenhill, C.S.: Corrigendum: The complexity of counting graph homomorphisms. *Random Struct. Algorithms*, 25(3):346–352, 2004.
17. Feder, T., Vardi, M.Y.: The computational structure of monotone monadic snp and constraint satisfaction: A study through datalog and group theory. *SIAM J. Comput.*, 28(1):57–104, 1998.
18. Freedman, M., Lovász, L., Schrijver, A.: Reflection positivity, rank connectivity, and homomorphism of graphs. *J. AMS*, 20:37–51, 2007.
19. Goldberg, L.A., Grohe, M., Jerrum, M., Thurley, M.: A complexity dichotomy for partition functions with mixed signs. In Proceedings of *STACS*, pages 493–504, 2009. *CoRR*, abs/0804.1932, 2008.
20. Hell, P., Nešetřil, J.: On the complexity of h-coloring. *Journal of Combinatorial Theory, Series B*, 48(1):92–110, 1990.
21. Kowalczyk, M.: Classification of a Class of Counting Problems Using Holographic Reductions. In *Proceedings of COCOON '09*, pages 472-485, 2009.
22. Kowalczyk, M., Cai, J-Y.: Holant Problems for Regular Graphs with Complex Edge Functions. In Proceedings of *STACS*, pages, 525-536, 2010.
23. Vadhan, S.P.: The complexity of counting in sparse, regular, and planar graphs. *SIAM J. Comput.*, 31(2):398–427, 2001.
24. Valiant, L.G.: The complexity of enumeration and reliability problems. *SIAM J. Comput.*, 8(3):410–421, 1979.
25. Valiant, L.G.: Holographic algorithms. In *FOCS '04: Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 306–315,2004. SIAM J. Comput. 37(5): 1565-1594 (2008).
26. Valiant, L.G.: Accidental algorthims. In *FOCS '06: Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 509–517, 2006.