A full dichotomy for Holant^c, inspired by quantum computation

Backens, Miriam

A full dichotomy for Holant$^{\mathrm{c}}$, inspired by quantum computation[*]

Miriam Backens

August 11, 2021

Abstract

Holant problems are a family of counting problems parameterised by sets of algebraic-complex valued constraint functions, and defined on graphs. They arise from the theory of holographic algorithms, which was originally inspired by concepts from quantum computation.

Here, we employ quantum information theory to explain existing results about holant problems in a concise way and to derive two new dichotomies: one for a new family of problems, which we call Holant$^+$, and, building on this, a full dichotomy for Holant^c. These two families of holant problems assume the availability of certain unary constraint functions – the two pinning functions in the case of Holant^c, and four functions in the case of Holant$^+$ – and allow arbitrary sets of algebraic-complex valued constraint functions otherwise. The dichotomy for Holant$^+$ also applies when inputs are restricted to instances defined on planar graphs. In proving these complexity classifications, we derive an original result about entangled quantum states.

1 Introduction

Quantum computation provided the inspiration for holographic algorithms [37], which in turn inspired the holant framework for computational counting problems (first introduced in the conference version of [18]). Computational counting problems include a variety of computational problems, from combinatorial problems defined on graphs to the problems of computing partition functions in statistical physics and computing amplitudes in quantum computation. They are being analysed in different frameworks, including that of counting constraint satisfaction problems (counting CSPs) and that of holant problems. Computational counting problems are an area of active research, yet so far there appear to have been no attempts to apply knowledge from quantum information theory or quantum computation to their analysis. Nevertheless, as we show in the following, quantum information theory, and particularly the theory of quantum entanglement, offer promising new avenues of research into holant problems.

A holant problem is parameterised by a set of functions $\mathcal{F}$; in this paper we consider finite sets of algebraic complex-valued functions of Boolean inputs. The restriction to finite sets follows the standard in the counting CSP community. We use it to avoid issues around efficient computability that arise when allowing problems to be parameterised by infinite sets of functions. In the following, the set of all algebraic complex-valued functions of Boolean inputs is denoted Υ. We also write $\Upsilon_n := \{f \in \Upsilon \mid \mathrm{arity}(f) = n\}$ for the restriction of Υ to functions of arity n. An instance of the problem $\mathsf{Holant}(\mathcal{F})$ consist of a multigraph $G = (V, E)$ with vertices V and edges E, and a map π. This map assigns to each vertex $v \in V$ a function $\pi(v) = f_v \in \mathcal{F}$. The map also sets up a bijection between the edges incident on v and the arguments of f_v, so the degree of v must equal the arity of f_v. Given the map π, any assignment $\sigma : E \to \{0, 1\}$ of Boolean values to edges induces a weight

$$\mathrm{wt}(\sigma) := \prod_{v \in V} f_v(\sigma|_{E(v)}),$$

---

[*]This paper combines extended versions of the conference publications [1] and [3] with some new results.

where $\sigma|_{E(v)}$ is the restriction of σ to the edges incident on v. The desired output of the holant problem is the total weight $\sum_{\sigma:E\to\{0,1\}} \text{wt}(\sigma)$, where the sum is over all assignments σ. Formally, we define the problem $\mathsf{Holant}(\mathcal{F})$ as follows.

Name $\mathsf{Holant}(\mathcal{F})$
Instance A tuple $(G, \mathcal{F}, \pi)$.
Output $\text{Holant}_\Omega = \sum_{\sigma:E\to\{0,1\}} \prod_{v\in V} f_v(\sigma|_{E(v)})$.

For example, let $N \in \mathbb{N}_{>0}$, and define $\mathcal{M}_N := \{\text{ONE}_n \mid 1 \le n \le N\}$, where

$$\text{ONE}_n(x_1,\dots,x_n) = \begin{cases} 1 & \text{if } \sum_{k=1}^n x_k = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathsf{Holant}(\mathcal{M}_N)$ corresponds to counting the number of perfect matchings on graphs of maximum degree N. To see this, consider some graph $G = (V, E)$ with maximum degree N. The map π must assign to each vertex a function of appropriate arity and since all functions are invariant under permutations of the arguments, this means π is fully determined by this requirement. Consider an assignment $\sigma : E \to \{0, 1\}$, this corresponds to a subset of edges $E_\sigma := \{e \in E \mid \sigma(e) = 1\}$. Note that $\text{wt}(\sigma) \in \{0, 1\}$ for all σ, since each weight is a product of functions taking values in the set $\{0, 1\}$. Now suppose $\text{wt}(\sigma) = 1$, then for each $v \in V$, σ assigns the value 1 to exactly one edge in $E(v)$. Thus, E_σ is a perfect matching. Conversely, if $\text{wt}(\sigma) = 0$ then there exists a vertex for which $f_v(\sigma|_{E(v)}) = 0$. This implies $\sum_{e \in E(v)} \sigma(e) \ne 1$, i.e. either 0 or at least 2 of the edges incident on v are assigned 1 by σ. Thus, E_σ contains either no edges incident on v, or at least 2 edges incident on v, so E_σ is not a perfect matching. We have shown that $\text{wt}(\sigma) = 1$ if σ corresponds to a perfect matching and $\text{wt}(\sigma) = 0$ otherwise. The correspondence between assignments $\sigma : E \to \{0, 1\}$ and subsets $E_\sigma \subseteq E$ is a bijection. Therefore, the total weight – and hence the output of $\mathsf{Holant}(\mathcal{F})$ – is exactly the number of perfect matchings of G.

Sometimes it is interesting to restrict holant problems to instances defined on planar graphs. In that case, the correspondence between the edges incident on a vertex v and the arguments of its assigned function f_v must respect the ordering of the edges in the given planar embedding of the graph and the ordering of the arguments of the function. This means that choosing an edge to correspond to the first argument of f_v fixes the bijection between edges and arguments for all the remaining edges and arguments. Given a finite set of functions $\mathcal{F}$, the problem of computing holant values on planar graphs is denoted $\mathsf{Pl\text{-}Holant}(\mathcal{F})$. For example, $\mathsf{Pl\text{-}Holant}(\mathcal{M}_N)$ is the problem of counting perfect matchings on planar graphs of maximum degree N.

The complexity of holant problems has not been fully classified yet, but there exist a number of classifications for subfamilies of holant problems. These classifications restrict the family of problems in one or more of the following ways:

- by considering only symmetric functions, i.e. functions that are invariant under any permutation of the arguments (these functions depend only on the Hamming weight of their input), or

- by restricting the codomain, e.g. to algebraic real numbers or even non-negative real numbers, or

- by considering only sets of functions that contain some specified subset of functions, such as arbitrary unary functions or the two 'pinning functions' $\delta_0(x) = 1 - x$ and $\delta_1(x) = x$.

Known classification include a full dichotomy for holant problems where all unary functions are assumed to be available [12], and a classification for $\mathsf{Holant}^c(\mathcal{F}) := \mathsf{Holant}(\mathcal{F} \cup \{\delta_0, \delta_1\})$ with symmetric functions [17]. There is also a dichotomy for real-valued Holant^c, where functions need

not be symmetrical but must take values in $\mathbb{R}$ instead of $\mathbb{C}$ [19]. Both existing results about Holant^c are proved via dichotomies for counting CSPs with complex-valued, not necessarily symmetric functions (see Section 3.2 for a formal definition): in the first case, a dichotomy for general #CSP [18], and in the second case, a dichotomy for #CSP$_2^c$, a subfamily of #CSP in which each variable must appear an even number of times and variables can be pinned to 0 or 1 [19]. The only broad classifications not assuming availability of certain functions are the dichotomy for symmetric holant [11] and the dichotomy for non-negative real-valued holant [33].

The problem of computing amplitudes in quantum computation can also be expressed as a holant problem. This makes certain holant problems examples of *(strong) classical simulation of quantum computations*, an area of active research in the theory of quantum computation. If a quantum computation can be simulated efficiently on a classical computer, the quantum computer yields no advantage. If, on the other hand, the complexity of classically simulating a certain family of quantum computations can be shown to satisfy some hardness condition, this provides evidence that quantum computations are indeed more powerful than classical computers.

Indeed, while the origin of holant problems is not related to the notion of classical simulation of quantum computations, quantum computing did provide the inspiration for its origins [37, 18]. Nevertheless, so far there have been no attempts to apply knowledge from quantum information theory or quantum computation to the analysis of holant problems. Yet, as we show in the following, quantum information theory, and particularly the theory of quantum entanglement, offer promising new avenues of research into holant problems.

The complexity of Holant $(\mathcal{F})$ may depend on *decomposability* properties of the functions in $\mathcal{F}$. A function $f \in \Upsilon_n$ is considered to be decomposable if there exists a permutation $\rho : [n] \to [n]$, an integer k satisfying $1 \leq k < n$, and functions f_1, f_2 such that

$$f(x_1, \ldots, x_n) = f_1(x_{\rho(1)}, \ldots, x_{\rho(k)}) f_2(x_{\rho(k+1)}, \ldots, x_{\rho(n)}).$$

For example, the constant function $f(x_1, x_2) = 1$ is equal to $f_1(x_1) f_2(x_2)$, where $f_1(x) = f_2(x) = 1$. Not all functions are decomposable: there are no functions $g_1, g_2 \in \Upsilon_1$ such that $\mathrm{EQ}_2(x_1, x_2) = g_1(x_1) g_2(x_2)$.

The functions in Υ_n are in bijection with the algebraic complex-valued vectors of length 2^n by mapping each function to a list of its values, or conversely. Under this map, the notion of a function being non-decomposable corresponds exactly to the quantum-theoretical notion of a vector being *genuinely entangled*. We therefore draw on the large body of research about the properties of entangled vectors from quantum theory, and apply it to the complexity classification of holant problems. In the process, we also derive a new result about entanglement previously unknown in the quantum theory literature, this is Theorem 46.

For relating the complexity of two computational problems, the main technique used in this paper is that of *gadget reductions*. An n-ary gadget over some set of functions $\mathcal{F}$ is a fragment of a signature grid using functions from $\mathcal{F}$, which is connected to the rest of the graph by n external edges. This signature grid is assigned an effective function in Υ_n by treating the external edges as variables and summing over all possible assignments of Boolean values to the internal edges. If the function g is the effective function of some gadget over $\mathcal{F}$, then we say g is *realisable* over $\mathcal{F}$. Allowing functions realisable over $\mathcal{F}$ in addition to the functions in $\mathcal{F}$ does not affect the complexity of computing the holant [12].

Now, the problem Holant$(\mathcal{F})$ is known to be polynomial-time computable if $\mathcal{F}$ contains only functions that decompose as products of unary and binary functions [12, Theorem 2.1]. For Holant$(\mathcal{F})$ to be #P-hard, $\mathcal{F}$ must thus contain a function that does not decompose in this way. Borrowing terminology from quantum theory, we say functions that do not decompose as products of unary and binary functions have *multipartite entanglement*. By applying results from quantum theory, we show how to build a gadget for a non-decomposable ternary function

using an arbitrary function with multipartite entanglement and the four unary functions

$$\delta_0(x) = 1 - x, \qquad \delta_1(x) = x, \qquad \delta_+(x) = 1, \quad \text{and} \quad \delta_-(x) = (-1)^x.$$

We further apply entanglement theory to analyse a family of symmetric ternary gadgets, and show that given some non-decomposable ternary function, this family lets us realise a symmetric non-decomposable ternary function. Furthermore, we show how to realise symmetric non-decomposable binary functions with specific properties. Together, these gadgets allow us to classify the complexity of the problem $\mathsf{Holant}^+(\mathcal{F}) := \mathsf{Holant}(\mathcal{F} \cup \{\delta_0, \delta_1, \delta_+, \delta_-\})$.

Theorem 1 (Informal statement of Theorem 56). *Let $\mathcal{F} \subseteq \Upsilon$ be finite. Then $\mathsf{Holant}^+(\mathcal{F})$ can be computed in polynomial time if $\mathcal{F}$ satisfies one of five explicit conditions. In all other cases, $\mathsf{Holant}^+(\mathcal{F})$ is #P-hard. The same dichotomy holds when the problem is restricted to instances defined on planar graphs.*

By combining these entanglement-based techniques with techniques from the real-valued Holant^c classification in [13], we then develop a complexity classification for $\mathsf{Holant}^c(\mathcal{F}) := \mathsf{Holant}(\mathcal{F} \cup \{\delta_0, \delta_1\})$. The problem Holant^c had first been described about a decade ago, with a full complexity classification remaining open until now.

Theorem 2 (Informal statement of Theorem 59). *Let $\mathcal{F} \subseteq \Upsilon$ be finite. Then $\mathsf{Holant}^c(\mathcal{F})$ can be computed in polynomial time if $\mathcal{F}$ satisfies one of six explicit conditions. In all other cases, $\mathsf{Holant}^c(\mathcal{F})$ is #P-hard.*

The remainder of this paper is structured as follows. We introduce preliminary definitions in Section 2 and give an overview over known holant complexity classifications in Section 3. In Section 4, we introduce relevant notions and results from quantum theory, particularly about entanglement. The complexity classification for Holant^+ is derived in Section 5 and the complexity classification for Holant^c is derived in Section 6.

2 Preliminaries

Holant problems are a framework for computational counting problems on graphs, introduced by Cai *et al.* [18], and based on the theory of holographic algorithms developed by Valiant [37].

Throughout, we consider algebraic complex-valued functions with Boolean inputs. Let $\mathbb{A}$ be the set of algebraic complex numbers. For any non-negative integer k, let Υ_k be the set of functions $\{0,1\}^k \to \mathbb{A}$, and define $\Upsilon := \bigcup_{k \in \mathbb{N}} \Upsilon_k$. Let $\mathcal{F} \subseteq \Upsilon$ be a finite set of functions, and let $G = (V, E)$ be an undirected graph with vertices V and edges E. Throughout, graphs are allowed to have parallel edges and self-loops. A *signature grid* is a tuple $\Omega = (G, \mathcal{F}, \pi)$ where π is a function that assigns to each n-ary vertex $v \in V$ a function $f_v : \{0,1\}^n \to \mathbb{A}$ in $\mathcal{F}$, specifying which edge incident on v corresponds to which argument of f_v. The Holant for a signature grid Ω is:

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

where σ is an assignment of Boolean values to each edge and $\sigma|_{E(v)}$ is the application of σ to the edges incident on v.

Name $\mathsf{Holant}(\mathcal{F})$
Instance A signature grid $\Omega = (G, \mathcal{F}, \pi)$.
Output Holant_Ω.

Remark. We restrict the definition of holant problems to finite sets of functions, as is common in the counting CSP literature and some of the holant literature. This is to avoid issues with the representation of the function values, and it is also relevant for some of the results in Section 5.4.

4

For any positive integer n, let $[n] := \{1, 2, \ldots, n\}$. Suppose $f \in \Upsilon_n$ and suppose $\rho : [n] \to [n]$ is a permutation, then $f_\rho(x_1, \ldots, x_n) := f(x_{\rho(1)}, \ldots, x_{\rho(n)})$.

A function is called *symmetric* if it depends only on the Hamming weight of the input, i.e. the number of non-zero bits in the input bit string. An n-ary symmetric function is often written as $f = [f_0, f_1, \ldots, f_n]$, with f_k for $k \in [n]$ being the value f takes on inputs of Hamming weight k. For any positive integer k, define the *equality function* of arity k as $\mathrm{EQ}_k := [1, 0, \ldots, 0, 1]$, where there are $(k-1)$ zeroes in the expression. Furthermore, define $\mathrm{ONE}_k := [0, 1, 0, \ldots, 0]$, where there are $(k-1)$ zeroes at the end. Let $\delta_0 := [1, 0]$ and $\delta_1 := [0, 1]$, these are called the *pinning functions* because they allow 'pinning' inputs to 0 or 1, respectively. Finally, define $\delta_+ := [1, 1]$ (this is equal to EQ_1), $\delta_- := [1, -1]$, and $\mathrm{NEQ} = [0, 1, 0]$. We will also occasionally use the unary functions $\delta_i := [1, i]$ and $\delta_{-i} = [1, -i]$, where i is the imaginary unit, i.e. $i^2 = -1$.

An n-ary function is called *degenerate* if it is a product of unary functions in the following sense: there exist functions $u_1, \ldots, u_n \in \Upsilon_1$ such that $f(x_1, \ldots, x_n) = u_1(x_1) \ldots u_n(x_n)$. Any function that cannot be expressed as a product of unary functions in this way is called *non-degenerate*. Suppose $f \in \Upsilon_n$, k is an integer satisfying $1 \leq k < n$, and $\rho : [n] \to [n]$ is a permutation such that $f_\rho(x_1, \ldots, x_n) = g(x_1, \ldots, x_k)h(x_{k+1}, \ldots, x_n)$ for some functions $g \in \Upsilon_k, h \in \Upsilon_{n-k}$, then f is said to be *decomposable*. Any way of writing f as a product of two or more functions with disjoint sets of arguments is called a *decomposition* of f. Borrowing terminology from linear algebra (which will be justified later), we say g and h in the above equation are *tensor factors* of f. A function that cannot be decomposed in this way is called *non-decomposable*. Any degenerate function of arity $n \geq 2$ is decomposable, but not all decomposable functions are degenerate. For example, $f(x_1, x_2, x_3, x_4) = \mathrm{EQ}_2(x_1, x_2)\mathrm{NEQ}(x_3, x_4)$ is decomposable but not degenerate since f cannot be written as a product of unary functions.

For any $f \in \Upsilon_n$, the *support* of f is $\mathrm{supp}(f) := \{\mathbf{x} \in \{0, 1\}^n \mid f(\mathbf{x}) \neq 0\}$.

Suppose $f \in \Upsilon_4$, $g \in \Upsilon_3$, and $h \in \Upsilon_2$. As a shorthand, let $f_{xyzw} := f(x, y, z, w)$, $g_{xyz} := g(x, y, z)$, and $h_{xy} := h(x, y)$ for any $x, y, z, w \in \{0, 1\}$. We sometimes identify these functions with the following matrices of their values:

$$f = \begin{pmatrix} f_{0000} & f_{0001} & f_{0010} & f_{0011} \\ f_{0100} & f_{0101} & f_{0110} & f_{0111} \\ f_{1000} & f_{1001} & f_{1010} & f_{1011} \\ f_{1100} & f_{1101} & f_{1110} & f_{1111} \end{pmatrix}, \quad g = \begin{pmatrix} g_{000} & g_{001} & g_{010} & g_{011} \\ g_{100} & g_{101} & g_{110} & g_{111} \end{pmatrix} \quad \text{and} \quad h = \begin{pmatrix} h_{00} & h_{01} \\ h_{10} & h_{11} \end{pmatrix}.$$

Here, for binary functions, matrix rows are labelled by the first input and columns are labelled by the second input. For ternary functions, rows are labelled by the first input and columns by the second and third inputs in lexicographic order. For functions of arity four, rows are labelled by the first two inputs and columns by the last two inputs, again in lexicographic order.

For any pair of counting problems A and B, we say A reduces to B and write $A \leq_T B$ if there exists a polynomial-time Turing reduction from problem A to problem B. If $A \leq_T B$ and $B \leq_T A$, we say A and B are *interreducible* and write $A \equiv_T B$.

The following result is well-known in the literature, see e.g. [13, p. 12].

Lemma 3. *Suppose $\mathcal{F} \subseteq \Upsilon$ is finite, $g \in \Upsilon$, and $c \in \mathbb{A} \setminus \{0\}$. Then* $\mathsf{Holant}\,(\mathcal{F} \cup \{c \cdot g\}) \equiv_T$ $\mathsf{Holant}\,(\mathcal{F} \cup \{g\})$.

Given a bipartite graph $G = (V, W, E)$, with vertex partitions V and W, we can define a *bipartite signature grid*. Let $\mathcal{F}$ and $\mathcal{G}$ be two finite subsets of Υ and let $\pi : V \cup W \to \mathcal{F} \cup \mathcal{G}$ be a map assigning functions to vertices, with the property that $\pi(v) \in \mathcal{F}$ for all $v \in V$ and $\pi(w) \in \mathcal{G}$ for all $w \in W$. The bipartite signature grid specified in this way is denoted by the tuple $(G, \mathcal{F}|\mathcal{G}, \pi)$. The corresponding *bipartite holant problem* is $\mathsf{Holant}\,(\mathcal{F} \mid \mathcal{G})$.

The following reductions relate bipartite and non-bipartite holant problems.

Proposition 4 ([17, Proposition 2]). *For any finite $\mathcal{F} \subseteq \Upsilon$, we have*

$$\mathsf{Holant}\,(\mathcal{F}) \equiv_T \mathsf{Holant}\,(\mathcal{F} \mid \{\mathrm{EQ}_2\}).$$

Proposition 5 ([17, Proposition 3]). *Suppose $\mathcal{G}_1, \mathcal{G}_2 \subseteq \Upsilon$ are finite, then*

$$\text{Holant}\left(\mathcal{G}_1 \cup \{\text{EQ}_2\} \mid \mathcal{G}_2 \cup \{\text{EQ}_2\}\right) \equiv_T \text{Holant}\left(\mathcal{G}_1 \cup \mathcal{G}_2\right).$$

A signature grid $\Omega = (\mathcal{F}, G, \pi)$ is called *planar* if G is a plane graph and, for each v, the arguments of f_v are ordered counterclockwise starting from an edge specified by π [15, Section 2.1]. We denote by Pl-Holant$(\mathcal{F})$ the problem Holant$(\mathcal{F})$ restricted to planar signature grids.

2.1 Signature grids in terms of vectors

As noted in [14], any function $f \in \Upsilon_n$ can be considered as a vector in $\mathbb{A}^{2^n}$, which is the list of values of f, indexed by $\{0,1\}^n$. Let $\{|\mathbf{x}\rangle\}_{\mathbf{x} \in \{0,1\}^n}$ be an orthonormal basis[1] for $\mathbb{A}^{2^n}$. The vector corresponding to the function f is then denoted by $|f\rangle := \sum_{\mathbf{x} \in \{0,1\}^n} f(\mathbf{x}) |\mathbf{x}\rangle$.

Denote by $\otimes$ the Kronecker product of matrices, which we usually call the *tensor product*. Based on this notion of tensor product, we define *tensor powers* of a matrix M as follows: $M^{\otimes 1} := M$ and $M^{\otimes k+1} = M^{\otimes k} \otimes M$ for any positive integer k. The operations of tensor product and tensor power can be extended to vectors by considering them as single-column matrices. Denote by M^T the transpose of the matrix M.

Suppose $\Omega = (G, \mathcal{F}|\mathcal{G}, \pi)$ is a bipartite signature grid, where $G = (V, W, E)$ has vertex partitions V and W. Then the holant for Ω can be written as:

$$\text{Holant}_\Omega = \left(\bigotimes_{w \in W} (|g_w\rangle)^T \right) \left(\bigotimes_{v \in V} |f_v\rangle \right) = \left(\bigotimes_{v \in V} (|f_v\rangle)^T \right) \left(\bigotimes_{w \in W} |g_w\rangle \right), \qquad (2)$$

where the tensor products are assumed to be ordered such that, in each inner product, two components associated with the same edge meet.

2.2 Holographic reductions

Holographic transformations are the origin of the name 'holant problems'. Let $\text{GL}_2(\mathbb{A})$ be the set of all invertible 2 by 2 matrices over $\mathbb{A}$, and let $\mathcal{O} := \{M \in \text{GL}_2(\mathbb{A}) \mid M^T M = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)\}$ be the set of orthogonal matrices.

Suppose $M \in \text{GL}_2(\mathbb{A})$, then for any $f \in \Upsilon_0$ let $M \circ f = f$, and for any $f \in \Upsilon_n$ with $n > 0$ let $M \circ f$ denote the function corresponding to the vector $M^{\otimes n} |f\rangle$. Furthermore, for any set of functions $\mathcal{F}$, define $M \circ \mathcal{F} := \{M \circ f \mid f \in \mathcal{F}\}$.

Theorem 6 (Valiant's Holant Theorem [37] as stated in [17, Proposition 4]). *For any $M \in \text{GL}_2(\mathbb{A})$ and any finite sets $\mathcal{F}, \mathcal{G} \subseteq \Upsilon$,*

$$\text{Holant}\left(\mathcal{F} \mid \mathcal{G}\right) \equiv_T \text{Holant}\left(M \circ \mathcal{F} \mid (M^{-1})^T \circ \mathcal{G}\right).$$

Corollary 7 ([17, Proposition 5]). *Let $O \in \mathcal{O}$ and let $\mathcal{F} \subseteq \Upsilon$ be finite, then*

$$\text{Holant}\left(\mathcal{F}\right) \equiv_T \text{Holant}\left(O \circ \mathcal{F}\right).$$

Going from a set of functions $\mathcal{F} \mid \mathcal{G}$ to $M \circ \mathcal{F} \mid (M^{-1})^T \circ \mathcal{G}$ or from $\mathcal{F}$ to $O \circ \mathcal{F}$ is a *holographic reduction*.

---

[1]In using this notation for vectors, called *Dirac notation* and common in quantum computation and quantum information theory, we anticipate the interpretation of the vectors associated with functions as quantum states, cf. Section 4.

2.3 Gadgets and realisability

A *gadget* over a finite set of functions $\mathcal{F}$ (also called $\mathcal{F}$-gate) is a fragment of a signature grid with some 'dangling' edges. Any such gadget can be assigned an effective function g.

Formally, let $G = (V, E, E')$ be a graph with vertices V, (normal) edges E, and dangling edges E', where $E \cap E' = \emptyset$. Unlike a normal edge, each dangling edge has only one end incident on a vertex in V, the other end is dangling. The gadget is determined by a tuple $\Gamma = (\mathcal{F}, G, \pi)$, where $\pi : V \to \mathcal{F}$ assigns a function to each vertex v in such a way that each argument of the function corresponds to one of the edges (normal or dangling) incident on v. Suppose $E' = \{e_1, \ldots, e_n\}$, then the effective function associated with this gadget is

$$g_\Gamma(y_1, \ldots, y_n) = \sum_{\sigma: E \to \{0,1\}} \prod_{v \in V} f_v(\hat{\sigma}|_{E(v)}),$$

where $\hat{\sigma}$ is the extension of σ to domain $E \cup E'$ which satisfies $\hat{\sigma}(e_k) = y_k$ for all $k \in [n]$, and $\hat{\sigma}|_{E(v)}$ is the restriction of $\hat{\sigma}$ to edges (both normal and dangling) which are incident on v.

If g is the effective function of some gadget over $\mathcal{F}$, then g is said to be *realisable over $\mathcal{F}$*. This notion is sometimes extended to say g is realisable over $\mathcal{F}$ if there exists a gadget over $\mathcal{F}$ with effective function $c \cdot g$ for some $c \in \mathbb{A} \setminus \{0\}$. By Lemma 3, the extended definition does not affect the validity of the following lemma.

Lemma 8 ([12, p. 1717])**.** *Suppose $\mathcal{F} \subseteq \Upsilon$ is finite and g is realisable over $\mathcal{F}$. Then*

$$\mathsf{Holant}\,(\mathcal{F} \cup \{g\}) \equiv_T \mathsf{Holant}\,(\mathcal{F})\,.$$

Following [33], we define $S(\mathcal{F}) = \{g \mid g \text{ is realisable over } \mathcal{F}\}$ for any set of functions $\mathcal{F}$. Then Lemma 8 implies that, for any finite subset $\mathcal{F}' \subseteq S(\mathcal{F})$, we have $\mathsf{Holant}\,(\mathcal{F}') \leq_T \mathsf{Holant}\,(\mathcal{F})$.

The following lemma will be useful later. This result is stated e.g. in [19, Lemma 2.1], but as it is not proved there, we give a quick proof here. Note this lemma uses the scaled definition of realisability.

Lemma 9. *Suppose $f(x_1, \ldots, x_{n+m}) = g(x_1, \ldots, x_n)h(x_{n+1}, \ldots, x_{n+m})$, where none of these functions are identically zero. Then $g, h \in S(\{f, \delta_0, \delta_1\})$.*

Proof. As g is not identically zero, there exists $\mathbf{a} \in \{0,1\}^n$ such that $g(\mathbf{a}) \neq 0$. But then

$$h(y_1, \ldots, y_m) = \sum_{x_1, \ldots, x_n \in \{0,1\}} f(x_1, \ldots, x_n, y_1, \ldots, y_m) \prod_{k \in [n]} \delta_{a_k}(x_k).$$

The right-hand side is the effective function of some gadget over $\{f, \delta_0, \delta_1\}$, which consists of one copy of f connected to n unary functions, so $h \in S(\{f, \delta_0, \delta_1\})$.

An analogous argument with the roles of g and h swapped shows that $g \in S(\{f, \delta_0, \delta_1\})$. $\square$

When considering a bipartite Holant problem $\mathsf{Holant}\,(\mathcal{F} \mid \mathcal{G})$ for some finite $\mathcal{F}, \mathcal{G} \subseteq \Upsilon$, we need to use gadgets that respect the bipartition. Suppose $\Gamma = (\mathcal{F}|\mathcal{G}, G, \pi)$ where $G = (V, W, E, E')$ is a bipartite graph with vertex partitions V and W, (normal) edges E and dangling edges E', and suppose $\pi : V \cup W \to \mathcal{F} \cup \mathcal{G}$ satisfies $\pi(v) \in \mathcal{F}$ for all $v \in V$ and $\pi(w) \in \mathcal{G}$ for all $w \in W$. If furthermore all dangling edges are incident on vertices from V, then Γ is called a *left-hand side (LHS) gadget* over $\mathcal{F}|\mathcal{G}$. Otherwise, if all dangling edges are incident on vertices from W, then Γ is called a *right-hand side (RHS) gadget* over $\mathcal{F}|\mathcal{G}$. The following result is a straightforward extension of Lemma 8.

Lemma 10. *Let $\mathcal{F}, \mathcal{G} \subseteq \Upsilon$ be two finite sets of functions. Suppose f is an LHS gadget over $\mathcal{F}|\mathcal{G}$ and g is a RHS gadget over $\mathcal{F}|\mathcal{G}$. Then*

$$\mathsf{Holant}\,(\mathcal{F} \cup \{f\} \mid \mathcal{G}) \equiv_T \mathsf{Holant}\,(\mathcal{F} \mid \mathcal{G}) \quad and \quad \mathsf{Holant}\,(\mathcal{F} \mid \mathcal{G} \cup \{g\}) \equiv_T \mathsf{Holant}\,(\mathcal{F} \mid \mathcal{G})\,.$$

A gadget is called *planar* if it is defined by a plane graph and if the dangling edges, ordered counterclockwise corresponding to the order of the arguments of the effective function, are in the outer face in a planar embedding [15, Section 2.4]. In reductions between planar holant problems, only planar gadgets may be used.

2.4 Polynomial interpolation

Finally, there is the technique of *polynomial interpolation*. Let $\mathcal{F}$ be a set of functions and suppose g is a function that cannot be realised over $\mathcal{F}$. If, given any signature grid over $\mathcal{F} \cup \{g\}$, it is possible to set up a family of signature grids over $\mathcal{F}$ such that the holant for the original problem instance can be determined efficiently from the holant values of the family by solving a system of linear equations, then g is said to be *interpolatable* over $\mathcal{F}$. We do not directly use polynomial interpolation here, though the technique is employed by many of the results we build upon. A rigorous definition of polynomial interpolation can be found in [18].

2.5 Linear algebra lemmas for holographic transformations

Holographic transformations are, at their core, linear maps. In this section, we give a few lemmas about decompositions of matrices that will significantly simplify later arguments about these transformations. In particular, we extend the orthogonal QR decomposition from real to complex matrices and prove two further lemmas building on it. These result are straightforward and may not be novel, but we have not been able to find a reference for them in the literature, so we provide proofs for completeness.

Let $K := \left(\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right)$, $X := \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $T := \left(\begin{smallmatrix} 1 & 0 \\ 0 & \exp(i\pi/4) \end{smallmatrix}\right)$ where i is the imaginary unit; these are all elements of $\mathrm{GL}_2(\mathbb{A})$.

Recall that any real square matrix M can be written as $M = QR$ where Q is an orthogonal matrix and R is upper (or lower) triangular. The equivalent result for complex matrices requires Q to be unitary instead of orthogonal. Nevertheless, many complex 2 by 2 matrices do admit a decomposition with a complex orthogonal matrix and an upper or lower triangular matrix. Where this is not possible, we give an alternative decomposition using the matrix K defined above instead of an orthogonal matrix.

Lemma 11 (Orthogonal QR decomposition for complex matrices). *Let M be an invertible 2 by 2 complex matrix, write $\mathcal{O}$ for the set of all 2 by 2 complex orthogonal matrices, and let K be as defined above. Then the following hold:*

- *There exists $Q \in \mathcal{O} \cup \{K, KX\}$ such that $Q^{-1}M$ is upper triangular.*

- *There exists $Q \in \mathcal{O} \cup \{K, KX\}$ such that $Q^{-1}M$ is lower triangular.*

- *If $Q^{-1}M$ is neither lower nor upper triangular for any orthogonal Q, then $M = KD$ or $M = KXD$, where D is diagonal.*

Proof. Write M as:
$$M = \begin{pmatrix} x & y \\ z & w \end{pmatrix}.$$

We assumed M was invertible, so $\det M = xw - yz \neq 0$. Note that $K^{-1} = \frac{1}{2}\left(\begin{smallmatrix} 1 & -i \\ 1 & i \end{smallmatrix}\right)$ and $(KX)^{-1} = \frac{1}{2}\left(\begin{smallmatrix} 1 & i \\ 1 & -i \end{smallmatrix}\right)$.

For a lower triangular decomposition, we want the top right element of $Q^{-1}M$ to vanish, this is $(Q^{-1})_{00}y + (Q^{-1})_{01}w$. The elements y and w cannot both be zero since M is invertible. Suppose first $y^2 + w^2 = 0$, i.e. $w = \pm iy$. Then

$$\begin{pmatrix} 1 & \pm i \\ 1 & \mp i \end{pmatrix} \begin{pmatrix} x & y \\ z & \pm iy \end{pmatrix} = \begin{pmatrix} x \pm iz & 0 \\ x \mp iz & 2y \end{pmatrix}$$

so $Q^{-1}M$ is lower triangular for some $Q \in \{K, KX\}$. If instead $y^2 + w^2 \neq 0$, then

$$\frac{1}{\sqrt{y^2 + w^2}} \begin{pmatrix} w & -y \\ y & w \end{pmatrix} \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \frac{1}{\sqrt{y^2 + w^2}} \begin{pmatrix} wx - yz & 0 \\ yx + wz & y^2 + w^2 \end{pmatrix},$$

and $\frac{1}{\sqrt{y^2 + w^2}} \begin{pmatrix} w & -y \\ y & w \end{pmatrix} \in \mathcal{O}$.

Analogously, for an upper triangular decomposition, some $Q \in \{K, KX\}$ works if $x^2 + z^2 = 0$, and some $Q \in \mathcal{O}$ works otherwise.

We have $Q \in \{K, KX\}$ for both decompositions if and only if $x^2 + z^2 = 0$ and $y^2 + w^2 = 0$ simultaneously. Write $z = \pm ix$. Then, by invertibility of M, $w = \mp iy$. Thus, letting $D = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$:

$$M = \begin{pmatrix} x & y \\ \pm ix & \mp iy \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ \pm i & \mp i \end{pmatrix} \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} = \begin{cases} KD & \text{if } \pm \text{ goes to } +, \text{ or} \\ KXD & \text{if } \pm \text{ goes to } -. \end{cases}$$

This completes the proof. □

Lemma 12. *Suppose M is a 2 by 2 invertible complex matrix such that $M^T M = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ for some $\lambda, \mu \in \mathbb{C} \setminus \{0\}$. Then there exists a 2 by 2 orthogonal matrix Q and a diagonal matrix D such that $M = QD$.*

Proof. Given invertibility of M, the property $(M^T M)_{01} = (M^T M)_{10} = 0$ means that the columns of M are orthogonal to each other (under the 'real' inner product). Hence there must exists $Q \in \mathcal{O}$ such that the columns of M are scalings of the columns of Q, which implies $M = QD$. □

Using the complex QR decomposition, we can also consider the solutions of $A^T A \doteq X$, where '$\doteq$' denotes equality up to scalar factor.

Lemma 13. *The solutions of:*
$$A^T A \doteq X \tag{3}$$
are exactly those matrices A satisfying $A = KD$ or $A = KXD$ for some invertible diagonal matrix D.

Proof. First, we check that matrices of the form KD or KXD for some invertible diagonal matrix D satisfy (3). Indeed:

$$(KD)^T KD = D^T K^T KD = D \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} D = 2DXD = 2xyX \doteq X, \tag{4}$$

where $D = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$ with $x, y \in \mathbb{C} \setminus \{0\}$. Similarly:

$$(KXD)^T KXD = D^T X^T K^T KXD = 2DX^3 D = 2DXD = 2xyX \doteq X. \tag{5}$$

This completes the first part of the proof.

It remains to be shown that these are the only solutions of $A^T A \doteq X$. Assume, for the purposes of deriving a contradiction, that there is a solution that has an orthogonal QR decomposition. In particular, suppose $A = QR$ for some upper triangular matrix R. Then, by orthogonality of Q:

$$A^T A = R^T Q^T QR = R^T R = \begin{pmatrix} R_{00} & 0 \\ R_{01} & R_{11} \end{pmatrix} \begin{pmatrix} R_{00} & R_{01} \\ 0 & R_{11} \end{pmatrix} = \begin{pmatrix} R_{00}^2 & R_{00}R_{01} \\ R_{00}R_{01} & R_{01}^2 + R_{11}^2 \end{pmatrix}. \tag{6}$$

The only way for the top left component of this matrix to be zero, as required, is if R_{00} is zero. Yet, in that case, the top right and bottom left components of $A^T A$ are zero too, hence

$A^T A$ cannot be invertible. That is a contradiction because any non-zero scalar multiple of X is invertible.

A similar argument applies if R is lower triangular.

Thus, all solutions of (3) have to fall into the third case of Lemma 11: i.e. all solutions must be of the form $A = KD$ or $A = KXD$. $\qquad\square$

For any $M \in \mathrm{GL}_2(\mathbb{A})$, denote by $f_M(x, y) := M_{xy}$ the binary function corresponding to this matrix. The following is straightforward, but we give a proof for completeness.

Lemma 14. *Suppose $M \in \mathrm{GL}_2(\mathbb{A})$ and $g \in \Upsilon$, then $M \circ g, M^T \circ g \in S(\{g, f_M\})$.*

Proof. Let $n = \mathrm{arity}(g)$. We have

$$
(M \circ g)(x_1, \ldots, x_n) = \sum_{y_1, \ldots, y_n \in \{0,1\}} \left(\prod_{j=1}^{n} M_{x_j y_j} \right) g(y_1, \ldots, y_n)
$$

$$
= \sum_{y_1, \ldots, y_n \in \{0,1\}} \left(\prod_{j=1}^{n} f_M(x_j, y_j) \right) g(y_1, \ldots, y_n),
$$

so $M \circ g \in S(\{g, f_M\})$. Furthermore, for any $M \in \mathrm{GL}_2(\mathbb{A})$, $f_{M^T}(x, y) = f_M(y, x)$; therefore

$$
(M^T \circ g)(x_1, \ldots, x_n) = \sum_{y_1, \ldots, y_n \in \{0,1\}} \left(\prod_{j=1}^{n} f_M(y_j, x_j) \right) g(y_1, \ldots, y_n)
$$

and $M^T \circ g \in S(\{g, f_M\})$. $\qquad\square$

3 Known results about holant problems

We now introduce the existing families of holant problems and their associated dichotomy results. Gadget constructions (which are at the heart of many reductions) are easier the more functions are known to be available. As a result, several families of holant problems have been defined, in which certain sets of functions are freely available, i.e. are always added to the set of functions parameterising the holant problem. More formally, suppose $\mathcal{G} \subseteq \Upsilon$ is a finite set of functions and denote by $\mathsf{Holant}^{\mathcal{G}}$ the holant problem where functions in $\mathcal{G}$ are freely available, then $\mathsf{Holant}^{\mathcal{G}}(\mathcal{F}) := \mathsf{Holant}(\mathcal{F} \cup \mathcal{G})$ for any set of functions $\mathcal{F}$. Effectively, the problem $\mathsf{Holant}^{\mathcal{G}}$ restricts analysis to cases where the set of constraint functions contains $\mathcal{G}$.

3.1 Conservative holant

Write $\mathcal{U} := \Upsilon_1$ for compatibility with earlier notation. We will not use the notation Holant^* [18, 12] here as we do not define holant problems for infinite sets of constraint functions. Instead, as is common in the counting CSP literature, we refer to holant problems where arbitrary finite subsets of $\mathcal{U}$ are freely available as 'conservative'.

We begin with some definitions. Given a bit string $\mathbf{x}$, let $\bar{\mathbf{x}}$ be its bit-wise complement and let $|\mathbf{x}|$ denote its Hamming weight. Denote by $\langle \mathcal{F} \rangle$ the closure of a set of functions $\mathcal{F}$ under tensor products. Furthermore, define:

- the set of all unary and binary functions

$$
\mathcal{T} := \Upsilon_1 \cup \Upsilon_2,
$$

- the set of functions which are non-zero on at most two complementary inputs

$$
\mathcal{E} := \{ f \in \Upsilon \mid \exists \mathbf{a} \in \{0,1\}^{\mathrm{arity}(f)} \text{ such that } f(\mathbf{x}) = 0 \text{ if } \mathbf{x} \notin \{\mathbf{a}, \bar{\mathbf{a}}\} \},
$$

- the set of functions which are non-zero only on inputs of Hamming weight at most 1

$$\mathcal{M} := \{f \in \Upsilon \mid f(\mathbf{x}) = 0 \text{ if } |\mathbf{x}| > 1\}.$$

Note that $\mathcal{U} \subseteq \mathcal{T}$, $\mathcal{U} \subseteq \mathcal{E}$ and $\mathcal{U} \subseteq \mathcal{M}$. The following result has been adapted to our notation.

Theorem 15 ([12, Theorem 2.2]). *Suppose $\mathcal{F}$ is a finite subset of Υ. If*

- $\mathcal{F} \subseteq \langle \mathcal{T} \rangle$, *or*

- *there exists $O \in \mathcal{O}$ such that $\mathcal{F} \subseteq \langle O \circ \mathcal{E} \rangle$, or*

- $\mathcal{F} \subseteq \langle K \circ \mathcal{E} \rangle = \langle KX \circ \mathcal{E} \rangle$, *or*

- $\mathcal{F} \subseteq \langle K \circ \mathcal{M} \rangle$ *or* $\mathcal{F} \subseteq \langle KX \circ \mathcal{M} \rangle$,

then, for any finite subset $\mathcal{U}' \subseteq \mathcal{U}$, the problem $\mathsf{Holant}(\mathcal{F},\mathcal{U}')$ is polynomial-time computable. Otherwise, there exists a finite subset $\mathcal{U}' \subseteq \mathcal{U}$ such that $\mathsf{Holant}(\mathcal{F},\mathcal{U}')$ is #P-hard. The dichotomy is still true even if the inputs are restricted to planar graphs.

To get an intuition for the polynomial-time computable cases, first note that every tensor product can be thought of as a gadget over its factors, and that tensor closure commutes with holographic transformations. Furthermore, if a signature grid is not connected, its holant is just the product of the holant values of the individual connected components.

Now, in the first polynomial-time computable case of Theorem 15, $\mathcal{F} \subseteq \langle \mathcal{T} \rangle$, the signature grid can be transformed to one in which every vertex has degree at most 2 by replacing every function with a disconnected gadget over unary and binary functions. Then each connected component is a path or cycle, and its holant value can be computed by matrix multiplication.

In the second polynomial-time computable case, $\mathcal{F} \subseteq \langle O \circ \mathcal{E} \rangle$, again replace decomposable functions by disconnected gadgets. Suppose O is the identity matrix, then an assignment of a Boolean value to one edge contributes to at most one non–zero-weight assignment of values to all edges in the same connected component. Thus the holant value for any connected graph is a sum of at most two terms, which can be computed efficiently, and hence the overall value can be found. If O is not the identity, apply Corollary 7 to reduce to the identity case.

For the third polynomial-time computable case, $\mathcal{F} \subseteq \langle K \circ \mathcal{E} \rangle$, note that given any finite $\mathcal{G} \subseteq \Upsilon$,

$$\mathsf{Holant}\,(K \circ \mathcal{G}) \equiv_T \mathsf{Holant}(K \circ \mathcal{G} \mid \{\mathrm{EQ}_2\}) \equiv_T \mathsf{Holant}(\mathcal{G} \mid \{\mathrm{NEQ}\}) \leq_T \mathsf{Holant}(\mathcal{G} \cup \{\mathrm{NEQ}\}),$$

where the first equivalence is Proposition 4, the second is Theorem 6, and the final reduction is because dropping the restriction to bipartite signature grids cannot make the problem easier. Hence, since $\mathrm{NEQ} \in \mathcal{E}$, this case reduces to the second one.

For the final polynomial-time computable case, $\mathcal{F} \subseteq \langle K \circ \mathcal{M} \rangle$ or $\mathcal{F} \subseteq \langle KX \circ \mathcal{M} \rangle$, use the first two steps of the above reduction. It can be shown that all LHS gadgets over $\mathcal{M}|\{\mathrm{NEQ}\}$ are in $\langle \mathcal{M} \rangle$ (cf. e.g. [4, Lemma 43]). Thus, vertices from the right-hand side partition can be removed one-by-one, updating the signature grid in the process, until it decomposes into a discrete graph whose holant value can be computed straightforwardly.

The polynomial-time computable sets of Theorem 15 will also be related to quantum entanglement in Section 4.2. If $\mathcal{F}$ is not one of the exceptional sets defined above, then the closure of $\mathcal{F} \cup \mathcal{U}$ under taking gadgets contains all functions [4, Theorem 67].

3.2 Counting constraint satisfaction problems

The family of complex-weighted Boolean #CSP (counting constraint satisfaction problems) is defined as follows. Let $\mathcal{F} \subseteq \Upsilon$ be a finite set of functions and let V be a finite set of variables. A *constraint* c over $\mathcal{F}$ is a tuple consisting of a function $f_c \in \mathcal{F}$ and a *scope*, which is a tuple of arity(f) (not necessarily distinct) elements of V. If C is a set of constraints, then any assignment $\mathbf{x} : V \to \{0,1\}$ of values to the variables induces a *weight* $\mathrm{wt}_{(V,C)}(\mathbf{x}) := \prod_{c \in C} f_c(\mathbf{x}|_c)$, where $\mathbf{x}|_c$ denotes the restriction of $\mathbf{x}$ to the scope of c.

Name #CSP$(\mathcal{F})$
Instance A tuple (V, C), where V is a finite set of variables and C is a finite set of constraints over $\mathcal{F}$.
Output The value $Z_{(V,C)} = \sum_{\mathbf{x}:V \to \{0,1\}} \prod_{c \in C} f_c(\mathbf{x}|_c)$.

Counting constraint satisfaction problems are closely related to holant problems. In particular, a counting constraint satisfaction problem can be thought of as a bipartite holant problem with the constraints as one part and the variables as the other part. Each vertex corresponding to a variable is assigned the function EQ_k, with k the total number of times that variable appears across the scopes of all constraints. The straightforward formalisation of this idea would require parameterising the holant problem with an infinite set of functions, the set containing all equality functions of any arity. Yet the bipartite structure is not necessary: if two variable vertices are adjacent to each other, they can be merged into one, and if two constraint vertices are adjacent to each other, a new variable vertex assigned EQ_2 can be introduced between them. Now, $S(\{\mathrm{EQ}_3\})$ is exactly the set of all equality functions, hence:

Proposition 16 ([17, Proposition 1]). #CSP$(\mathcal{F}) \equiv_T$ Holant $(\mathcal{F} \cup \{\mathrm{EQ}_3\})$.

This proposition shows that any counting constraint satisfaction problem can be expressed as a holant problem. On the other hand, some holant problems can only be expressed as counting constraint satisfaction problems with the additional restriction that every variable must appear exactly twice across the scopes of all the constraints. This means the holant framework is more general than counting constraint satisfaction problems. For example, as is well known, the problem of counting matchings on a graph can be expressed in the holant framework but not in the standard #CSP framework as defined above (cf. [4, pp. 23:3–4]).

The dichotomies for #CSP and its variants feature families of tractable functions which do not appear in Theorem 15.

Definition 17. A function $f : \{0,1\}^n \to \mathbb{A}$ for some non-negative integer n is called *affine* if it has the form:

$$f(\mathbf{x}) = c i^{l(\mathbf{x})} (-1)^{q(\mathbf{x})} \chi_{A\mathbf{x}=\mathbf{b}}(\mathbf{x}), \qquad (7)$$

where $c \in \mathbb{A}$, $i^2 = -1$, $l : \{0,1\}^n \to \{0,1\}$ is a linear function, $q : \{0,1\}^n \to \{0,1\}$ is a quadratic function, A is an m by n matrix with Boolean entries for some $0 \le m \le n$, $\mathbf{b} \in \{0,1\}^m$, and χ is an indicator function which takes value 1 on inputs satisfying $A\mathbf{x} = \mathbf{b}$, and 0 otherwise.

The set of all affine functions is denoted by $\mathcal{A}$.

The support of the function $\chi_{A\mathbf{x}=\mathbf{b}}$ is an affine subspace of $\{0,1\}^n$, hence the name 'affine functions'. There are different definitions of this family in different parts of the literature, but they are equivalent. For the reader familiar with quantum information theory, the affine functions correspond – up to scaling – to stabiliser states (cf. Section 4.2 and also independently [16]).

Lemma 18 ([16, Lemma 3.1]). *If* $f(x_1, \ldots, x_n), g(y_1, \ldots, y_m) \in \mathcal{A}$, *then so are*

1. $(f \otimes g)(x_1, \ldots, x_n, y_1, \ldots, y_m) = f(x_1, \ldots, x_n)g(y_1, \ldots, y_m)$,

2. $f(x_{\rho(1)}, \ldots, x_{\rho(n)})$ *for any permutation* $\rho : [n] \to [n]$,

3. $f^{x_j = x_\ell}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) = f(x_1, \ldots, x_{j-1}, x_\ell, x_{j+1}, \ldots, x_n)$, *setting the variable* x_j *to be equal to* x_ℓ, *and*

4. $f^{x_j = *}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) = \sum_{x_j \in \{0,1\}} f(x_1, \ldots, x_n)$.

Definition 19 (adapted from [19, Definition 10]). A function $f : \{0,1\}^n \to \mathbb{A}$ for some non-negative integer n is called *local affine* if it satisfies $\left(\bigotimes_{j=1}^n T^{a_j} \right) |f\rangle \in \mathcal{A}$ for any $\mathbf{a} \in \mathrm{supp}(f)$, where

$$T^1 = T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix} \quad \text{and} \quad T^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The set of all local affine functions is denoted $\mathcal{L}$.

Both $\mathcal{A}$ and $\mathcal{L}$ are closed under tensor products, i.e. $\langle \mathcal{A} \rangle = \mathcal{A}$ and $\langle \mathcal{L} \rangle = \mathcal{L}$.

Theorem 20 ([18, Theorem 3.1]). *Suppose* $\mathcal{F} \subseteq \Upsilon$ *is finite. If* $\mathcal{F} \subseteq \mathcal{A}$ *or* $\mathcal{F} \subseteq \langle \mathcal{E} \rangle$, *then* $\#\mathsf{CSP}(\mathcal{F})$ *is computable in polynomial time. Otherwise,* $\#\mathsf{CSP}(\mathcal{F})$ *is* $\#P$-*hard.*

Unlike a holant problem, the complexity of a counting CSP does not change if the pinning functions δ_0 and δ_1 are added to the set of constraint functions.

Lemma 21 (Pinning lemma [23, Lemma 8]). *For any finite* $\mathcal{F} \subseteq \Upsilon$,

$$\#\mathsf{CSP}^c(\mathcal{F}) := \#\mathsf{CSP}(\mathcal{F} \cup \{\delta_0, \delta_1\}) \equiv_T \#\mathsf{CSP}(\mathcal{F}).$$

The dichotomy of Theorem 20 also holds for a variant counting constraint satisfaction problem called $\#\mathsf{R_3\text{-}CSP}$ with a restriction on the number of times each variable may appear [18, Theorem 3.2].

Name $\#\mathsf{R_3\text{-}CSP}(\mathcal{F})$
Instance A tuple (V, C), where V is a finite set of variables and C is a finite set of constraints over $\mathcal{F}$ such that each variable appears at most three times across all scopes.
Output The value $Z_{(V,C)} = \sum_{\mathbf{x}:V \to \{0,1\}} \prod_{c \in C} f_c(\mathbf{x}|_c)$.

For any finite $\mathcal{F} \subseteq \Upsilon$, the problem $\#\mathsf{R_3\text{-}CSP}(\mathcal{F})$ is equivalent to the bipartite holant problem $\mathsf{Holant}\,(\mathcal{F} \mid \{\mathrm{EQ}_1, \mathrm{EQ}_2, \mathrm{EQ}_3\})$ [18, Section 2]. Its dichotomy follows immediately from that for $\#\mathsf{CSP}$ if $\mathcal{F}$ contains the binary equality function (or indeed any equality function of arity at least 2), but this interreduction is non-trivial if $\mathcal{F}$ does not contain any equality function of arity at least 2.

We will also consider a variant of counting CSPs in which each variable has to appear an even number of times.

Name $\#\mathsf{CSP}_2(\mathcal{F})$
Instance A tuple (V, C), where V is a finite set of variables and C is a finite set of constraints over $\mathcal{F}$ such that each variable appears an even number of times across all scopes.
Output The value $Z_{(V,C)} = \sum_{\mathbf{x}:V \to \{0,1\}} \prod_{c \in C} f_c(\mathbf{x}|_c)$.

Based on the above definition, we also define $\#\mathsf{CSP}_2^c(\mathcal{F}) := \#\mathsf{CSP}_2(\mathcal{F} \cup \{\delta_0, \delta_1\})$. The dichotomy for this problem differs from that for plain $\#\mathsf{CSP}$.

Theorem 22 ([19, Theorem 4.1]). *Suppose* $\mathcal{F} \subseteq \Upsilon$ *is finite. A* $\#\mathsf{CSP}_2^c(\mathcal{F})$ *problem has a polynomial time algorithm if one of the following holds:*

- $\mathcal{F} \subseteq \langle \mathcal{E} \rangle$,

- $\mathcal{F} \subseteq \mathcal{A}$,

- $\mathcal{F} \subseteq T \circ \mathcal{A}$, or

- $\mathcal{F} \subseteq \mathcal{L}$.

Otherwise, it is #P-hard.

For any finite $\mathcal{F} \subseteq \Upsilon$, we have

$$\#\mathsf{CSP}(\mathcal{F} \cup \{\delta_+\}) \leq_T \#\mathsf{CSP}_2(\mathcal{F} \cup \{\delta_+\}),$$

where, for each variable y that appears an odd number of times in the original instance, we add a new constraint $(\delta_+, (y))$ to make it appear an even number of times instead. Yet if no non-zero scaling of δ_+ is present or realisable (via gadgets or other methods), then there is no general reduction from $\#\mathsf{CSP}(\mathcal{F})$ to $\#\mathsf{CSP}_2(\mathcal{F})$. This lack of a general reduction can be seen for example by noting the differences between Theorem 22 (which classifies the complexity of $\#\mathsf{CSP}_2^c$) and Theorem 20 (which by Lemma 21 also applies to $\#\mathsf{CSP}^c$).

By analogy with planar holant problems, we also define planar counting CSPs.

Definition 23. $\mathsf{Pl\text{-}\#CSP}(\mathcal{F}) := \mathsf{Pl\text{-}Holant}\,(\mathcal{F} \cup \{\mathrm{EQ}_3\})$, i.e. $\mathsf{Pl\text{-}\#CSP}(\mathcal{F})$ is the restriction of $\#\mathsf{CSP}(\mathcal{F})$ to planar instances of the corresponding holant problem according to Proposition 16.

The complexity dichotomy for $\mathsf{Pl\text{-}\#CSP}$ involves an additional tractable family as compared to the general case. This new family is called *matchgate functions* and consists of those functions which correspond to computing a weighted sum of perfect matchings. We denote the set of all matchgate functions by[2] $\mathcal{H} \subseteq \Upsilon$. As the rigorous definition of this set is somewhat intricate and not required for our work we do not reproduce it here; the interested reader can find it in [15, pp. STOC17-65f]. Indeed, the only property of $\mathcal{H}$ which we will require is the following lemma, which is adapted to avoid having to define concepts not used in this paper.

Lemma 24 ([15, first part of Lemma 2.29]). *If $f \in \Upsilon$ has arity ≤ 3, then $f \in \mathcal{H}$ if and only if one of the following parity conditions is satisfied:*

- $f(\mathbf{x}) = 0$ *whenever $|\mathbf{x}|$ is even, or*

- $f(\mathbf{x}) = 0$ *whenever $|\mathbf{x}|$ is odd.*

For $f \in \Upsilon_1$, this means f is a matchgate function if and only if $f = c \cdot \delta_0$ or $f = c \cdot \delta_1$ for some $c \in \mathbb{A}$.

We can now state the dichotomy for $\mathsf{Pl\text{-}\#CSP}$.

Theorem 25 ([15, Theorem 6.1′]). *Let $\mathcal{F}$ be any finite set of complex-valued functions in Boolean variables. Then $\mathsf{Pl\text{-}\#CSP}(\mathcal{F})$ is #P-hard unless $\mathcal{F} \subseteq \mathcal{A}$, $\mathcal{F} \subseteq \langle \mathcal{E} \rangle$, or $\mathcal{F} \subseteq \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix} \right) \circ \mathcal{H}$, in which case the problem is computable in polynomial time.*

3.3 Partial results for Holant^c and Holant

Holant^c is the holant problem in which the unary functions pinning edges to 0 or 1 are freely available [18, 17], i.e. $\mathsf{Holant}^c(\mathcal{F}) := \mathsf{Holant}\,(\mathcal{F} \cup \{\delta_0, \delta_1\})$ for any finite $\mathcal{F} \subseteq \Upsilon$. We will give a full dichotomy for this problem in Section 6, building on the following dichotomies for symmetric functions and for real-valued functions.

---

[2]Note that we use a different symbol than [15] for the set of matchgate functions, to avoid clashing with other established notation.

Theorem 26 ([17, Theorem 6]). *Let $\mathcal{F} \subseteq \Upsilon$ be a finite set of symmetric functions.* $\mathsf{Holant}^c(\mathcal{F})$ *is #P-hard unless $\mathcal{F}$ satisfies one of the following conditions, in which case it is polynomial-time computable:*

- $\mathcal{F} \subseteq \langle \mathcal{T} \rangle$, *or*

- *there exists $O \in \mathcal{O}$ such that $\mathcal{F} \subseteq \langle O \circ \mathcal{E} \rangle$, or*

- $\mathcal{F} \subseteq \langle K \circ \mathcal{E} \rangle = \langle KX \circ \mathcal{E} \rangle$, *or*

- $\mathcal{F} \subseteq \langle K \circ \mathcal{M} \rangle$ *or* $\mathcal{F} \subseteq \langle KX \circ \mathcal{M} \rangle$, *or*

- *there exists $B \in \mathcal{B}$ such that $\mathcal{F} \subseteq B \circ \mathcal{A}$, where:*

$$\mathcal{B} = \left\{ M \mid M^T \circ \{\mathrm{EQ}_2, \delta_0, \delta_1\} \subseteq \mathcal{A} \right\}. \tag{8}$$

Note that the first four polynomial-time computable cases are exactly the ones appearing in Theorem 15.

The preceding results all apply to algebraic complex-valued functions, but the following theorem is restricted to algebraic real-valued functions.

Theorem 27 ([19, Theorem 5.1]). *Let $\mathcal{F}$ be a set of algebraic real-valued functions. Then* $\mathsf{Holant}^c(\mathcal{F})$ *is #P-hard unless $\mathcal{F}$ is a tractable family for conservative holant or for $\#\mathsf{CSP}_2^c$.*

In the case of Holant with no freely available functions, there exists a dichotomy for complex-valued symmetric functions [11, Theorem 31] and a dichotomy for (not necessarily symmetric) functions taking non-negative real values [33, Theorem 19]. We will not explore those results in any detail here.

3.4 Results about ternary symmetric functions

The computational complexity of problems of the form $\mathsf{Holant}\left(\{[y_0, y_1, y_2]\} \mid \{[x_0, x_1, x_2, x_3]\}\right)$, where $[y_0, y_1, y_2] \in \Upsilon_2$ and $[x_0, x_1, x_2, x_3] \in \Upsilon_3$, has been fully determined. These are holant problems on bipartite graphs where one partition only contains vertices of degree 2, the other partition only contains vertices of degree 3, and all vertices of the same arity are assigned the same symmetric function.

If $[x_0, x_1, x_2, x_3]$ is degenerate, the problem is tractable by the first case of Theorem 15. If $[x_0, x_1, x_2, x_3]$ is non-degenerate, it can always be mapped to either $[1, 0, 0, 1]$ or $[1, 1, 0, 0]$ by a holographic transformation [17, Section 3], cf. also Section 4 below. By Theorem 6, it thus suffices to consider the cases $\{[y_0, y_1, y_2]\} \mid \{[1, 0, 0, 1]\}$ and $\{[y_0, y_1, y_2]\} \mid \{[1, 1, 0, 0]\}$.

There are also some holographic transformations which leave the function $[1, 0, 0, 1]$ invariant. In particular, if $M = \left(\begin{smallmatrix} 1 & 0 \\ 0 & \omega \end{smallmatrix}\right)$, where $\omega^3 = 1$, i.e. ω is a third root of unity, then $M \circ \mathrm{EQ}_3 = \mathrm{EQ}_3$ [17, Section 4]. By applying Theorem 6 with this M,

$$\mathsf{Holant}\left(\{[y_0, y_1, y_2]\} \mid \{[1, 0, 0, 1]\}\right) \equiv_T \mathsf{Holant}\left(\{[y_0, \omega y_1, \omega^2 y_2]\} \mid \{[1, 0, 0, 1]\}\right). \tag{9}$$

This relationship can be used to reduce the number of symmetric binary functions needing to be considered in this section. Following [17, Section 4], a symmetric binary function $[y_0, y_1, y_2]$ is called *ω-normalised*[3] if

- $y_0 = 0$, or

- there does not exist a primitive $(3t)$-th root of unity λ, where the greatest common divisor $\gcd(t, 3) = 1$, such that $y_2 = \lambda y_0$.

---

[3]We use the term ω-normalisation to distinguish it from other notions of normalisation, e.g. ones relating to the norm of the vector associated with a function.

Similarly, a unary function $[a, b]$ is called ω-normalised if

- $a = 0$, or

- there does not exist a primitive $(3t)$-th root of unity λ, where $\gcd(t, 3) = 1$, such that $b = \lambda a$.

If a binary function is not ω-normalised, it can be made so through application of a holographic transformation of the form given in (9). Unary functions will only be required when the binary function has the form $[0, y_1, 0]$; in that case the binary function is automatically ω-normalised, and it remains so under a holographic transformation that ω-normalises the unary function.

These definitions allow a complexity classification of all holant problems on bipartite signature grids where there is a ternary equality function on one partition and a non-degenerate symmetric binary function on the other partition.

Theorem 28 ([17, Theorem 5]). *Let $\mathcal{G}_1, \mathcal{G}_2 \subseteq \Upsilon$ be finite and let $[y_0, y_1, y_2] \in \Upsilon_2$ be an ω-normalised and non-degenerate function. In the case of $y_0 = y_2 = 0$, further assume that $\mathcal{G}_1$ contains a unary function $[a, b]$ which is ω-normalised and satisfies $ab \neq 0$. Then:*

$$\mathsf{Holant}\left(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \mid \{[1, 0, 0, 1]\} \cup \mathcal{G}_2\right) \equiv_T \#\mathsf{CSP}(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \cup \mathcal{G}_2).$$

Theorem 29 ([17, Theorem 4]). $\mathsf{Holant}\left(\{[y_0, y_1, y_2]\} \mid \{[x_0, x_1, x_2, x_3]\}\right)$ *is #P-hard unless $[y_0, y_1, y_2]$ and $[x_0, x_1, x_2, x_3]$ satisfy one of the following conditions, in which case the problem is polynomial-time computable:*

- $[x_0, x_1, x_2, x_3]$ *is degenerate, or*

- *there is $M \in \mathrm{GL}_2(\mathbb{A})$ such that:*

  - $[x_0, x_1, x_2, x_3] = M \circ [1, 0, 0, 1]$ *and $M^T \circ [y_0, y_1, y_2]$ is in $\mathcal{A} \cup \langle \mathcal{E} \rangle$,*
  - $[x_0, x_1, x_2, x_3] = M \circ [1, 1, 0, 0]$ *and $[y_0, y_1, y_2] = (M^{-1})^T \circ [0, a, b]$ for some $a, b \in \mathbb{A}$.*

Here, we have combined the last two cases of [17, Theorem 4] into one case, since one can be mapped to the other by a bit flip: a holographic transformation using the matrix X.

We will also use a result about planar holant problems involving the ternary equality function. As the notation used in the original statement of this theorem differs significantly from the notation used in this paper, we first state the original theorem and then prove a corollary which translates the theorem into our notation.

Theorem 30 ([30, Theorem 7]). *Let $a, b \in \mathbb{A}$ and define $X := ab$, $Y := a^3 + b^3$. The problem $\mathsf{Pl\text{-}Holant}\left(\{[a, 1, b]\} \mid \{\mathrm{EQ}_3\}\right)$ is #P-hard for all $a, b \in \mathbb{A}$ except in the following cases, for which the problem is polynomial-time computable:*

1. $X = 1$

2. $X = Y = 0$

3. $X = -1$ *and $Y \in \{0, \pm 2i\}$, or*

4. $4X^3 = Y^2$.

Corollary 31. *Suppose $g \in \Upsilon_2$ is symmetric. The problem $\mathsf{Pl\text{-}Holant}\left(\{g\} \mid \{\mathrm{EQ}_3\}\right)$ is #P-hard except in the following cases, for which the problem is polynomial-time computable:*

1. $g \in \langle \mathcal{E} \rangle$,

2. $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ g \in \mathcal{A}$ *for some $\lambda \in \mathbb{A}$ such that $\lambda^3 = 1$, or*

3. $g = c \cdot [a, 1, b]$, *where $a, b, c \in \mathbb{A} \setminus \{0\}$ and $a^3 = b^3$.*

Remark. Note that the three exceptional cases of Corollary 31 overlap, e.g. $g = \mathrm{EQ}_2$ satisfies both Case 1 and Case 2, and $g = [1, 1, 1]$ satisfies all three cases.

We also include binary functions of the form $[a, 0, b]$ in the corollary for completeness.

Proof of Corollary 31. First, we show that the problem is tractable in the exceptional cases. Let $M = \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix}$ with $\lambda^3 = 1$ be such that $g' := M \circ g$ is ω-normalised. Then

$$
\begin{aligned}
\mathsf{Pl\text{-}Holant}\left(\{g\} \mid \{\mathrm{EQ}_3\}\right) &\equiv_T \mathsf{Pl\text{-}Holant}\left(\{g'\} \mid \{\mathrm{EQ}_3\}\right) \\
&\leq_T \mathsf{Holant}\left(\{g', \mathrm{EQ}_3\}\right) \\
&\leq_T \#\mathsf{CSP}(\{g'\}),
\end{aligned}
$$

where the first step is by Theorem 6, the second step is by forgetting about the bipartition and the planarity constraint, and the third step is by Proposition 16. Hence if $\#\mathsf{CSP}(\{g'\})$ is polynomial-time computable, then so is $\mathsf{Pl\text{-}Holant}\left(\{g\} \mid \{\mathrm{EQ}_3\}\right)$. By Theorem 20, $\#\mathsf{CSP}(\{g'\})$ is polynomial-time computable if $g' \in \mathcal{A}$ or $g' \in \langle \mathcal{E} \rangle$. Since M is diagonal, $g' \in \langle \mathcal{E} \rangle$ is equivalent to $g \in \langle \mathcal{E} \rangle$, so we have tractability for Case 1. Furthermore, $g' \in \mathcal{A}$ is equivalent to $\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} \circ g \in \mathcal{A}$ for some $\lambda^3 = 1$, which establishes tractability for Case 2.

This leaves Case 3. Assume $g = c \cdot [a, 1, b]$ for some $a, b, c \in \mathbb{A} \setminus \{0\}$ with $a^3 = b^3$, and define $g'' := [a, 1, b]$. The functions g and g'' differ only by a non-zero factor, so by a straightforward extension of Lemma 3 to bipartite signature grids we have

$$
\mathsf{Pl\text{-}Holant}\left(\{g\} \mid \{\mathrm{EQ}_3\}\right) \equiv_T \mathsf{Pl\text{-}Holant}\left(\{g''\} \mid \{\mathrm{EQ}_3\}\right). \tag{10}
$$

Now, $a^3 = b^3$ implies

$$
0 = (a^3 - b^3)^2 = (a^3 + b^3)^2 - 4(ab)^3,
$$

hence g'' satisfies condition 4 of Theorem 30. This establishes tractability for Case 3.

It remains to prove the hardness part of the theorem. If $g = [a, 0, b]$ for some $a, b \in \mathbb{A}$, then $g \in \langle \mathcal{E} \rangle$ and $\mathsf{Pl\text{-}Holant}\left(\{g\} \mid \{\mathrm{EQ}_3\}\right)$ is polynomial-time computable by the above arguments. Thus, from now on, we may assume $g = c \cdot [a, 1, b]$, where $a, b, c \in \mathbb{A}$ and $c \neq 0$.

Let $g'' := [a, 1, b]$ as before, then again (10) holds. We will show that if g'' satisfies one of the tractability conditions of Theorem 30, then g satisfies one of Cases 1–3. Consider each tractability condition of Theorem 30 in turn.

1. $X = 1$ is equivalent to $ab = 1$. This is exactly the condition for the function $[a, 1, b]$ to be degenerate. But then $g = c \cdot [a, 1, b]$ is degenerate as well, so $g \in \langle \mathcal{E} \rangle$ and g satisfies Case 1.

2. $X = Y = 0$ is equivalent to $ab = 0$ and $a^3 + b^3 = 0$. Together, the two equalities imply that $a = b = 0$. Thus $g'' = \mathrm{NEQ}$, which implies $g = c \cdot \mathrm{NEQ} \in \langle \mathcal{E} \rangle$, so g satisfies Case 1.

3. $X = -1$ implies that $a, b \neq 0$. We can therefore rewrite $X = ab = -1$ to $b = -a^{-1}$, which in turn implies $Y = a^3 - a^{-3}$. This case therefore reduces to $a^3 - a^{-3} \in \{0, \pm 2i\}$. We distinguish subcases.

   - Suppose $a^3 - a^{-3} = 0$. Then $a^6 = 1$, i.e. $a = e^{ik\pi/3}$ for some $k \in \{0, 1, 2, 3, 4, 5\}$.
     - If $k = 0$, let $\lambda = 1$, then $\lambda^3 = 1$ and $\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} \circ g = c \cdot [1, 1, -1] \in \mathcal{A}$.
     - If $k = 1$, let $\lambda = e^{4i\pi/3}$, then $\lambda^3 = 1$ and

       $$
       \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} \circ g = c \cdot [e^{i\pi/3}, \lambda, -\lambda^2 e^{-i\pi/3}] = c \cdot [e^{i\pi/3}, e^{4i\pi/3}, -e^{7i\pi/3}] = e^{4i\pi/3} c \cdot [-1, 1, 1]
       $$

       so $\begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} \circ g \in \mathcal{A}$.

– If $k = 2$, let $\lambda = e^{2i\pi/3}$, then $\lambda^3 = 1$ and

$$\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ g = c \cdot [e^{2i\pi/3}, \lambda, -\lambda^2 e^{-2i\pi/3}] = c \cdot [e^{2i\pi/3}, e^{2i\pi/3}, -e^{2i\pi/3}] = e^{2i\pi/3} c \cdot [1, 1, -1]$$

so $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ g \in \mathcal{A}.$

The remaining subcases are similar.

- Suppose $a^3 - a^{-3} = 2i$. Then $a^6 - 1 = 2ia^3$, or equivalently $a^6 - 2ia^3 - 1 = (a^3 - i)^2 = 0$. Thus $a = e^{(4k+1)i\pi/6}$ for some $k \in \{0, 1, 2\}$.

  – If $k = 0$, let $\lambda = e^{2i\pi/3}$, then $\lambda^3 = 1$ and

  $$\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ g = c \cdot [e^{i\pi/6}, \lambda, -\lambda^2 e^{-i\pi/6}] = c \cdot [e^{i\pi/6}, e^{2i\pi/3}, -e^{7i\pi/6}] = e^{i\pi/6} c \cdot [1, -1, 1]$$

  so $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ g \in \mathcal{A}.$

  – If $k = 1$, let $\lambda = e^{4i\pi/3}$, then $\lambda^3 = 1$ and

  $$\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ g = c \cdot [e^{5i\pi/6}, \lambda, -\lambda^2 e^{-5i\pi/6}] = c \cdot [e^{5i\pi/6}, e^{4i\pi/3}, -e^{11i\pi/6}] = e^{5i\pi/6} c \cdot [1, -1, 1]$$

  so $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ g \in \mathcal{A}.$

  – If $k = 2$, let $\lambda = 1$, then $\lambda^3 = 1$ and

  $$\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ g = c \cdot [e^{9i\pi/6}, \lambda, -\lambda^2 e^{-9i\pi/6}] = c \cdot [e^{3i\pi/2}, 1, -e^{-3i\pi/2}] = c \cdot [-i, 1, -i]$$

  so $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ g \in \mathcal{A}.$

- Suppose $a^3 - a^{-3} = -2i$. This subcase is analogous to the subcase $a^3 - a^{-3} = 2i$.

In each subcase, we were able to find $\lambda \in \mathbb{A}$ such that $\lambda^3 = 1$ and $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ g \in \mathcal{A}$, i.e. the function g satisfies Case 2.

4. $4X^3 = Y^2$ implies $4(ab)^3 = (a^3 + b^3)^2$, which is equivalent to $(a^3 - b^3)^2 = 0$. Hence this condition immediately implies that g satisfies Case 3.

This completes the case distinction.

We have shown that if $g'' = [a, 1, b]$ satisfies any of the tractability conditions of Theorem 30, then $g = c \cdot g''$ satisfies one of Cases 1–3. Hence, conversely, if $g = c \cdot [a, 1, b]$ and g does not satisfy any of Cases 1–3, then $g'' = [a, 1, b]$ does not satisfy any of the tractability conditions of Theorem 30. In that case, Theorem 30 implies that $\mathsf{Pl\text{-}Holant}(\{g''\} \mid \{\mathrm{EQ}_3\})$ is #P-hard. Thus, by (10), $\mathsf{Pl\text{-}Holant}(\{g\} \mid \{\mathrm{EQ}_3\})$ is #P-hard. $\qquad\square$

3.5 Results about functions of arity 4

Besides the above results about ternary functions, we will also make use of the following result about realising or interpolating the arity-4 equality function from a more general function of arity 4.

Lemma 32 ([15, Part of Lemma 2.40]). *Suppose $\mathcal{F} \subseteq \Upsilon$ is finite and contains a function f of arity 4 with matrix*

$$\begin{pmatrix} a & 0 & 0 & b \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ c & 0 & 0 & d \end{pmatrix}$$

where $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ has full rank. Then $\mathsf{Pl\text{-}Holant}(\{\mathrm{EQ}_4\} \cup \mathcal{F}) \leq_T \mathsf{Pl\text{-}Holant}(\mathcal{F})$.

The lemma can of course also be used in the non-planar setting.

Functions in $\mathcal{E}$ are often called 'generalised equality functions'. Recall from Section 3.2 that #CSP$_2$ is a counting CSP in which each variable has to appear an even number of times.

Lemma 33 ([19, Lemma 5.2])**.** *Suppose $\mathcal{F} \subseteq \Upsilon$ is finite and contains a generalised equality function f of arity 4, then*

$$\mathsf{Holant}\,(\mathcal{F}) \equiv_T \#\mathsf{CSP}_2(\mathcal{F}).$$

Remark. Note that the statement of Lemma 5.2 in [19] has 'Holant^c' instead of plain 'Holant', but the proof does not use pinning functions, so this is presumably a typo.

4　The quantum state perspective

In Section 2.1, we introduced the idea of considering functions as complex vectors. This perspective is not only useful for proving Valiant's Holant Theorem (which is at the heart of the theory of holant problems), it also gives a connection to the theory of quantum computation.

In quantum computation and quantum information theory, the basic system of interest is a *qubit* (quantum bit), which takes the place of the usual bit in standard computer science. The state of a qubit is described by a vector[4] in the two-dimensional complex Hilbert space $\mathbb{C}^2$. State spaces compose by tensor product, i.e. the state of n qubits is described by a vector in $\left(\mathbb{C}^2\right)^{\otimes n}$, which is isomorphic to $\mathbb{C}^{2^n}$. Here, $\otimes$ denotes the tensor product of Hilbert spaces and tensor powers are defined analogously to tensor powers of matrices in Section 2.1. Thus, the vector associated with an n-ary function can be considered to be a quantum state of n qubits. The vectors describing quantum states are usually required to have norm 1, but for the methods used here, multiplication by a non-zero complex number does not make a difference, so we can work with states having arbitrary norms.

Let $\{|0\rangle, |1\rangle\}$ be an orthonormal basis for $\mathbb{C}^2$. This is usually called the *computational basis*. The induced basis on $\left(\mathbb{C}^2\right)^{\otimes n}$ is labelled by $\{|\mathbf{x}\rangle\}_{\mathbf{x} \in \{0,1\}^n}$ as a short-hand, e.g. we write $|00\ldots0\rangle$ instead of $|0\rangle \otimes |0\rangle \otimes \ldots \otimes |0\rangle$. This is exactly the same as the basis introduced in Section 2.1.

Holographic transformations also have a natural interpretation in quantum information theory: going from an n-qubit state $|f\rangle$ to $M^{\otimes n} |f\rangle$, where M is some invertible 2 by 2 matrix, is a 'stochastic local operation with classical communication' (SLOCC) [6, 22]. These are physical operations that can be applied locally (without needing access to more than one qubit at a time) using classical (i.e. non-quantum) communication between the sites where the different qubits are held, and which succeed with non-zero probability. If two quantum states are equivalent under SLOCC, they can be used for the same quantum information tasks, albeit potentially with different probabilities of success. Two n-qubit states $|\psi\rangle$ and $|\phi\rangle$ are equivalent under SLOCC if and only if there exist invertible complex 2 by 2 matrices $M_1, M_2, \ldots, M_n$ such that $|\psi\rangle = (M_1 \otimes M_2 \otimes \ldots \otimes M_n) |\phi\rangle$ [22, Section II.A]. In particular, SLOCC operations do not need to be symmetric under permutations.

From now on, we will sometimes mix standard holant terminology (or notation) and quantum terminology (or notation).

4.1　Entanglement and its classification

One major difference between quantum theory and preceding theories of physics (known as 'classical physics') is the possibility of *entanglement* in states of multiple systems.

Definition 34. A state of multiple systems is *entangled* if it cannot be written as a tensor product of states of individual systems.

Example. In the case of two qubits,

$$|00\rangle + |01\rangle + |10\rangle + |11\rangle \tag{11}$$

---

[4]Strictly speaking, vectors only describe *pure* quantum states: there are also *mixed* states, which need to be described differently; but we do not consider those here.

is a product state – it can be written as $(|0\rangle + |1\rangle) \otimes (|0\rangle + |1\rangle)$. On the other hand, consider the state

$$|00\rangle + |11\rangle. \tag{12}$$

It is impossible to find single-qubit states $|f\rangle, |g\rangle \in \mathbb{C}^2$ such that $|f\rangle \otimes |g\rangle = |00\rangle + |11\rangle$. In function notation, this can be seen by noting that if $h(x,y) = f(x)g(y)$, then

$$h(0,0)h(1,1) - h(0,1)h(1,0) = f(0)g(0)f(1)g(1) - f(0)g(1)f(1)g(0) = 0,$$

whereas for $h(x,y) = \mathrm{EQ}_2(x,y)$, we have $h(0,0)h(1,1) - h(0,1)h(1,0) = 1$. Thus, $|00\rangle + |11\rangle$ is entangled.

Where a state involves more than two systems, it is possible for some of the systems to be entangled with each other and for other systems to be in a product state with respect to the former. We sometimes use the term *genuinely entangled state* to refer to a state in which no subsystem is in a product state with the others. The term *multipartite entanglement* refers to entangled states in which more than two qubits are mutually entangled.

Under the bijection between functions in vectors described in Section 2.1, a state vector is entangled if and only if the corresponding function is non-degenerate. A state is genuinely entangled if and only if the corresponding function is non-decomposable. Finally, a state has multipartite entanglement if and only if the corresponding function has a non-decomposable factor of arity at least three. In other words, entangled states correspond to functions in $\Upsilon \setminus \langle \mathcal{U} \rangle$ and multipartite entangled states correspond to functions in $\Upsilon \setminus \langle \mathcal{T} \rangle$.

Entanglement is an important resource in quantum computation, where it has been shown that quantum speedups are impossible without the presence of unboundedly growing amounts of entanglement [29]. Similarly, it is a resource in quantum information theory [35], featuring in protocols such as quantum teleportation [5] and quantum key distribution [24]. Many quantum information protocols have the property that two quantum states can be used to perform the same task if one can be transformed into the other by SLOCC, motivating the following equivalence relation.

Definition 35. Two n-qubit states are *equivalent under SLOCC* if one can be transformed into the other using SLOCC. More formally: suppose $|f\rangle, |g\rangle \in (\mathbb{C}^2)^{\otimes n}$ are two n-qubit states. Then $|f\rangle \sim_{SLOCC} |g\rangle$ if and only if there exist invertible complex 2 by 2 matrices $M_1, M_2, \ldots, M_n$ such that $(M_1 \otimes M_2 \otimes \ldots \otimes M_n)|f\rangle = |g\rangle$.

The equivalence classes of this relation are called *entanglement classes* or *SLOCC classes*. This definition is justified because SLOCC does not affect the decomposition of a state into tensor factors. To see this, suppose $|f\rangle, |g\rangle$ are two n-qubit states satisfying Definition 35. Furthermore, suppose $|f\rangle$ decomposes as $\left|f_1^k\right\rangle \otimes \left|f_{k+1}^n\right\rangle$ for some k with $1 \le k < n$, where $\left|f_1^k\right\rangle$ is a state on the first k qubits and $\left|f_{k+1}^n\right\rangle$ is a state on the last $(n-k)$ qubits. Then

$$\begin{aligned}
|g\rangle &= (M_1 \otimes M_2 \otimes \ldots \otimes M_n)|f\rangle \\
&= (M_1 \otimes M_2 \otimes \ldots \otimes M_n)\left(\left|f_1^k\right\rangle \otimes \left|f_{k+1}^n\right\rangle\right) \\
&= \left((M_1 \otimes \ldots \otimes M_k)\left|f_1^k\right\rangle\right) \otimes \left((M_{k+1} \otimes \ldots \otimes M_n)\left|f_{k+1}^n\right\rangle\right),
\end{aligned}$$

so $|g\rangle$ decomposes in the same way as $|f\rangle$. Since the matrices $M_1, M_2, \ldots, M_n$ are invertible, the converse also holds. Hence $|g\rangle$ can be decomposed as a tensor product according to some partition if and only if $|f\rangle$ can be decomposed according to the same partition. Therefore SLOCC does not affect entanglement. Due to the correspondences between vectors and functions outlined in the first part of the section, the same holds for holographic transformations.

For two qubits, there is only one class of entangled states: all entangled two-qubit states are equivalent to $|00\rangle + |11\rangle$ (the vector corresponding to EQ_2) under SLOCC. For three qubits,

there are two classes of genuinely entangled states, the GHZ class and the W class [22]. The former contains states that are equivalent under SLOCC to the GHZ state:

$$|\text{GHZ}\rangle := \frac{1}{\sqrt{2}}(|000\rangle + |111\rangle), \tag{13}$$

the latter those equivalent to the W state:

$$|W\rangle := \frac{1}{\sqrt{3}}(|001\rangle + |010\rangle + |100\rangle). \tag{14}$$

Note that, up to scalar factor, $|\text{GHZ}\rangle$ is the vector corresponding to EQ_3 and $|W\rangle$ is the vector corresponding to ONE_3.

We say that a function has GHZ type if it is equivalent to the GHZ state under local holographic transformations and that a function has W type if it is equivalent to the GHZ state under local holographic transformations. Local holographic transformations include invertible scaling, so non-zero scalar factors do not affect the entanglement classification. In the holant literature, GHZ-type functions have been called the *generic case* and W-type functions have been called the *double-root case*, cf. [17, Section 3].

The two types of genuinely entangled ternary functions can be distinguished as follows.

Lemma 36 ([32]). *Let f be a ternary function and write $f_{klm} := f(k, l, m)$ for all $k, \ell, m \in \{0, 1\}$. Then f has GHZ type if the following polynomial in the function values is non-zero:*

$$(f_{000}f_{111} - f_{010}f_{101} + f_{001}f_{110} - f_{011}f_{100})^2 - 4(f_{010}f_{100} - f_{000}f_{110})(f_{011}f_{101} - f_{001}f_{111}). \tag{15}$$

The function f has W type if the polynomial (15) is zero, and furthermore each of the following three expressions is satisfied:

$$(f_{000}f_{011} \neq f_{001}f_{010}) \vee (f_{101}f_{110} \neq f_{100}f_{111}) \tag{16}$$

$$(f_{001}f_{100} \neq f_{000}f_{101}) \vee (f_{011}f_{110} \neq f_{010}f_{111}) \tag{17}$$

$$(f_{011}f_{101} \neq f_{001}f_{111}) \vee (f_{010}f_{100} \neq f_{000}f_{110}). \tag{18}$$

If the polynomial (15) is zero and at least one of the three expressions evaluates to false, f is decomposable. In fact, for any decomposable f, at least two of the expressions are false.

The above lemma can be specialised to symmetric functions as follows.

Lemma 37. *Let f be a ternary symmetric function and write $f = [f_0, f_1, f_2, f_3]$. Then f has GHZ type if the following polynomial in the function values is non-zero:*

$$(f_0 f_3 - f_1 f_2)^2 - 4(f_1^2 - f_0 f_2)(f_2^2 - f_1 f_3) \neq 0. \tag{19}$$

The function f has W type if the above polynomial is zero and furthermore

$$(f_1^2 \neq f_0 f_2) \vee (f_2^2 \neq f_1 f_3).$$

If the polynomial in (19) is zero and the above expression evaluates to false, f is decomposable.

For joint states of more than three qubits, there are infinitely many SLOCC classes. It is possible to partition these into families which share similar properties. Yet, so far, there is no consensus on how to partition the classes: there are different schemes for partitioning even the four-qubit entanglement classes, yielding different families [38, 31, 2].

A *generalised GHZ state* is a vector corresponding (up to scalar factor) to EQ_k for some integer $k \geq 3$. A *generalised W state* is a vector corresponding (up to scalar factor) to ONE_k for some integer $k \geq 3$.

4.2 The existing results in the quantum picture

Using the correspondence between vectors and functions, several of the existing dichotomies have straightforward descriptions in the quantum picture. The tractable cases of Theorem 15 can be described as follows:

- The case $\mathcal{F} \subseteq \langle \mathcal{T} \rangle$ corresponds to vectors with no multipartite entanglement. Unbounded multipartite entanglement is needed for quantum computation to offer any advantage over non-quantum computation [29], so it makes sense that its absence would lead to a holant problem that is polynomial-time computable.

- In the cases $\mathcal{F} \subseteq \langle O \circ \mathcal{E} \rangle$ or $\mathcal{F} \subseteq \langle K \circ \mathcal{E} \rangle$, assuming $\mathcal{F} \nsubseteq \langle \mathcal{T} \rangle$, there is GHZ-type multipartite entanglement. To see this, note first that if $f \in \langle \mathcal{E} \rangle \setminus \langle \mathcal{T} \rangle$ is non-decomposable, then f must have arity at least 3 and be non-zero on exactly two complementary inputs. Suppose f has arity n, and let $\mathbf{a} \in \{0,1\}^n$ be such that $f(\mathbf{a}) \neq 0$. Without loss of generality, assume $a_1 = 0$ (if $a_1 = 1$, replace $\mathbf{a}$ by $\bar{\mathbf{a}}$). Then

$$|f\rangle = \left(\begin{pmatrix} f(\mathbf{a}) & 0 \\ 0 & f(\bar{\mathbf{a}}) \end{pmatrix} \otimes \bigotimes_{k=2}^{n} X^{a_n} \right) |\mathrm{EQ}_n\rangle ,$$

where X^0 is the identity matrix; hence $|f\rangle \sim_{SLOCC} |\mathrm{EQ}_n\rangle$. Further holographic transformations do not affect the equivalence under SLOCC. As formally shown in [4, Lemma 46], the sets $\langle O \circ \mathcal{E} \rangle$ and $\langle K \circ \mathcal{E} \rangle$ are already closed under taking gadgets, so every non-decomposable function of arity n is SLOCC-equivalent to EQ_n. This means it is impossible to realise W-type multipartite entanglement from $\mathcal{F}$ via gadgets, which again indicates these cases are insufficient to describe full quantum computation.

- Finally, in the case $\mathcal{F} \subseteq \langle K \circ \mathcal{M} \rangle$ or $\mathcal{F} \subseteq \langle KX \circ \mathcal{M} \rangle$, again assuming $\mathcal{F} \nsubseteq \langle \mathcal{T} \rangle$, there is W-type multipartite entanglement. To see this, note first that if $f \in \langle \mathcal{M} \rangle \setminus \langle \mathcal{T} \rangle$ is non-decomposable, then f must have arity at least 3. Suppose $n := \mathrm{arity}(f) \geq 3$ and there exists an index $k \in [n]$ such that f is 0 on the bit string that has a 1 only on position k and zeroes elsewhere. Then by the definition of $\mathcal{M}$, the function f is 0 whenever input k is non-zero; hence f can be decomposed as $f(x_1, \ldots, x_n) = \delta_0(x_k) f'(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$. Thus, any non-decomposable $f \in \langle \mathcal{M} \rangle$ has support on all bit strings of Hamming weight exactly 1. Now suppose $f \in \mathcal{M}$ is a non-decomposable function of arity n, then if $f(0 \ldots 0) = 0$, we have

$$|f\rangle = \left(\bigotimes_{k=1}^{n} \begin{pmatrix} 1 & 0 \\ 0 & f(\mathbf{e}_k) \end{pmatrix} \right) |\mathrm{ONE}_n\rangle ,$$

where $\mathbf{e}_k$ for $1 \leq k \leq n$ is the n-bit string that contains a single 1 in position k and zeroes elsewhere. If $f(0 \ldots 0) \neq 0$, then

$$|f\rangle = f(0 \ldots 0) \left(\bigotimes_{k=1}^{n} \begin{pmatrix} 1 & \frac{1}{n} \\ 0 & \frac{f(\mathbf{e}_k)}{f(0 \ldots 0)} \end{pmatrix} \right) |\mathrm{ONE}_n\rangle .$$

In both cases $|f\rangle \sim_{SLOCC} |\mathrm{ONE}_n\rangle$, i.e. any non-decomposable function in $\mathcal{M}$ has W-type entanglement. Further holographic transformations do not affect SLOCC-equivalence. Again, it was shown in [4, Lemma 46] that $\langle K \circ \mathcal{M} \rangle$ and $\langle KX \circ \mathcal{M} \rangle$ are closed under taking gadgets, so every non-decomposable function of arity n they contain is SLOCC-equivalent to ONE_n. In particular, it is impossible to realise GHZ-type multipartite entanglement.

The family of affine functions (cf. Definition 17) also has a natural description in quantum information theory: the quantum states corresponding to affine functions are known as *stabiliser states* [20]. These states and the associated operations play an important role in the context of quantum error-correcting codes [28] and are thus at the core of most attempts to build large-scale quantum computers [21]. The fragment of quantum theory consisting of stabiliser states and operations that preserve the set of stabiliser states can be efficiently simulated on a classical computer [28]; this result is known as the Gottesman-Knill theorem. The connection between affine functions and stabiliser quantum mechanics has recently been independently noted and explored in [16].

The examples in this section show that holant problems and quantum information theory are linked not only by quantum algorithms being an inspiration for holographic ones: instead, many of the known tractable set of functions of various holant problems correspond to state sets that are of independent interest in quantum computation and quantum information theory. The one exception are local affine functions, which seem not to have been described in the quantum literature, possibly because this set does not contain any interesting unitary operations.

The restriction to algebraic numbers is not a problem from the quantum perspective, not even when considering the question of universal quantum computation: there exist (approximately) universal sets of quantum operations where each operation can be described using algebraic complex coefficients. One such example is the Clifford+T gate set [7, 27], which is generated by the operators

$$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}, \qquad \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

4.3 Affine functions, holographic transformations, and entanglement

The lemmas in this section, like those in Section 2.5 are straightforward. They will be useful in the complexity classification proofs later.

The first results is well known in quantum theory. It is also closely related to the theory of functional clones [9, 8] and holant clones [4], though we will not introduce the full formalisms of those frameworks here. Instead, we give a proof in the language of gadgets.

Lemma 38. *The set of affine functions is closed under taking gadgets, i.e. $S(\mathcal{A}) = \mathcal{A}$.*

Proof. Suppose $G = (V, E, E')$ is a graph with vertices V, (normal) edges E, and dangling edges $E' = \{e_1, \ldots, e_n\}$, where $E \cap E' = \emptyset$. Let $\Gamma = (\mathcal{A}, G, \pi)$ be a gadget with effective function

$$g_\Gamma(y_1, \ldots, y_n) = \sum_{\sigma: E \to \{0,1\}} \prod_{v \in V} f_v(\hat{\sigma}|_{E(v)}),$$

where $\hat{\sigma}$ is the extension of σ to domain $E \cup E'$ which satisfies $\hat{\sigma}(e_k) = y_k$ for all $k \in [n]$, and $\hat{\sigma}|_{E(v)}$ is the restriction of $\hat{\sigma}$ to edges (both normal and dangling) which are incident on v.

We prove $g_\Gamma \in \mathcal{A}$ by induction on the number of normal edges $m := |E|$.

The base case $m = 0$ implies that all edges are dangling and $g_\Gamma = \bigotimes_{v \in V} f_v$. Then by associativity of $\otimes$ and by repeated application of Lemma 18 (1), we have $g_\Gamma \in \mathcal{A}$.

For the inductive step, assume the desired property holds if there are m normal edges. Consider a gadget $\Gamma = (\mathcal{A}, G, \pi)$ with $m + 1$ normal edges and n dangling edges. Pick some $e = \{u, v\} \in E$ and 'cut it', i.e. replace it by two dangling edges e_{n+1}, e_{n+2}, where e_{n+1} is incident on u and e_{n+2} is incident on v. Let $\bar{E} = E \setminus \{e\}$ and let $E'' = E' \cup \{e_{n+1}, e_{n+2}\}$. The resulting graph is $G' = (V, \bar{E}, E'')$. Since G' has the same vertices as G and each vertex has

the same degree in both graphs, $\Gamma' = (\mathcal{A}, G', \pi)$ is a valid gadget, where π is the same map as before. Then

$$g_{\Gamma'}(y_1, \ldots, y_{n+2}) = \sum_{\sigma : \bar{E} \to \{0,1\}} \prod_{v \in V} f_v(\hat{\sigma}'|_{E(v)}),$$

where $\hat{\sigma}'$ is the extension of σ to domain $\bar{E} \cup E''$ which satisfies $\hat{\sigma}'(e_k) = y_k$ for all $k \in [n+2]$. Now Γ' is a gadget with m normal edges, so by the inductive hypothesis, $g_{\Gamma'}(y_1, \ldots, y_{n+2}) \in \mathcal{A}$.

But $g_{\Gamma}(y_1, \ldots, y_n) = \sum_{y_{n+1} \in \{0,1\}} g_{\Gamma'}(y_1, \ldots, y_n, y_{n+1}, y_{n+1})$, i.e. $g_{\Gamma} = (g_{\Gamma'}^{y_{n+2} = y_{n+1}})^{y_{n+1} = *}$. Thus, by Lemma 18 (3) and (4), we have $g_{\Gamma} \in \mathcal{A}$. $\qquad\square$

Lemma 39. *The set $\mathcal{B}_{\mathcal{A}} := \{L \in \mathrm{GL}_2(\mathbb{A}) \mid f_L \in \mathcal{A}\}$ is a group under matrix multiplication.*

Proof. Closure under matrix multiplication follows directly from Lemma 38. The identity matrix corresponds to EQ_2, which is affine, so $\mathcal{B}_{\mathcal{A}}$ contains the identity. For closure under inverse, note that by Definition 17 any matrix $A \in \mathcal{B}_{\mathcal{A}}$ corresponds to a function

$$f_A(x, y) = c i^{\ell(x,y)} (-1)^{q(x,y)} \chi(x, y),$$

where $c \in \mathbb{A} \setminus \{0\}$ by invertibility of A, $\ell, q : \{0,1\}^2 \to \{0,1\}$ with ℓ being linear and q being quadratic, and $\chi \in \{[1,1,1], [1,0,1], [0,1,0]\}$ is the indicator function for an affine support. (With support on some other affine subspace of $\{0,1\}^2$, A could not be invertible.) Constant terms in ℓ or q can be absorbed into c, and for Boolean variables, $x^2 = x$, so non–cross-terms from q can be absorbed into ℓ. Thus, without loss of generality, $\ell(x, y) = \lambda x + \mu y$ and $q(x, y) = \kappa xy$ for some $\lambda, \mu \in \{0, 1, 2, 3\}$ and $\kappa \in \{0, 1\}$. Let $P := \left(\begin{smallmatrix} 1 & 0 \\ 0 & i \end{smallmatrix} \right)$.

Now if $\chi = [1, 1, 1]$, then

$$A = c \begin{pmatrix} 1 & i^\mu \\ i^\lambda & (-1)^\kappa i^{\lambda+\mu} \end{pmatrix} = cP^\lambda \begin{pmatrix} 1 & 1 \\ 1 & (-1)^\kappa \end{pmatrix} P^\mu.$$

For A to be invertible, κ must be 1. But then $A^{-1} = c^{-1} P^{4-\mu} \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix} \right) P^{4-\lambda} \in \mathcal{B}_{\mathcal{A}}$. On the other hand, if $\chi = [1, 0, 1]$ or $\chi = [0, 1, 0]$, then

$$A = c \begin{pmatrix} 1 & 0 \\ 0 & (-1)^\kappa i^{\lambda+\mu} \end{pmatrix} \qquad \text{or} \qquad A = c \begin{pmatrix} 0 & i^\mu \\ i^\lambda & 0 \end{pmatrix}.$$

Again, in both cases A^{-1} has the same form as A, so $A^{-1} \in \mathcal{B}_{\mathcal{A}}$. Hence $\mathcal{B}_{\mathcal{A}}$ is a group. $\qquad\square$

Up to scaling, the unary elements of $\mathcal{A}$ are

$$\delta_0 := [1, 0], \quad \delta_1 := [0, 1], \quad \delta_+ := [1, 1], \quad \delta_- := [1, -1], \quad \delta_i := [1, i], \quad \text{and} \quad \delta_{-i} := [1, -i].$$

We say the pairs $\{\delta_0, \delta_1\}$, $\{\delta_+, \delta_-\}$, and $\{\delta_i, \delta_{-i}\}$ are orthogonal pairs[5], any other pair of distinct unary functions $u, v \in \mathcal{A}$ is called non-orthogonal. It is straightforward to see that if $u, u^\perp$ are an orthogonal pair and $v \in \mathcal{A}$ is a unary function that is not a scaling of u or $u^\perp$, then $v(x) = \alpha \cdot u(x) + \beta \cdot u^\perp(x)$ where $\alpha^4 = \beta^4$.

Lemma 40. *Suppose $M \in \mathcal{B} = \left\{ L \in \mathrm{GL}_2(\mathbb{A}) \,\big|\, L^T \circ \{\mathrm{EQ}_2, \delta_0, \delta_1\} \subseteq \mathcal{A} \right\}$ and $M^T \circ \delta_+, M^T \circ \delta_- \in \mathcal{A}$. Then $M \in \mathcal{B}_{\mathcal{A}} = \{L \in \mathrm{GL}_2(\mathbb{A}) \mid f_L \in \mathcal{A}\}$.*

Proof. The property $M \in \mathcal{B}$ implies there are $u, v \in \{\delta_0, \delta_1, \delta_+, \delta_-, \delta_i, \delta_{-i}\}$ and $\lambda, \mu \in \mathbb{A} \setminus \{0\}$ such that $M^T \circ \delta_0 = \lambda \cdot u$ and $M^T \circ \delta_1 = \mu \cdot v$. But $\delta_\pm(x) = \delta_0(x) \pm \delta_1(x)$, so by linearity $(M^T \circ \delta_\pm)(x) = \lambda \cdot u(x) \pm \mu \cdot v(x)$.

First, suppose u and v are orthogonal. Then $M^T = M' \left(\begin{smallmatrix} \lambda & 0 \\ 0 & \mu \end{smallmatrix} \right)$ where

$$M' \in \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \right\},$$

---

[5]This is because the two corresponding vectors are orthogonal under the complex inner product.

since this set contains all matrices that map δ_0 and δ_1 to a pair of orthogonal functions in $\mathcal{A}$, including permutations. It is straightforward to check that if u and v are orthogonal, then $M^T \circ \delta_\pm \in \mathcal{A}$ if and only if $\lambda^4 = \mu^4$. But then M^T is a product of two matrices corresponding to functions in $\mathcal{A}$, so $M \in \mathcal{B}_\mathcal{A}$ by Lemma 38.

Now suppose u and v are not orthogonal; denote by $u^\perp$ the function which forms an orthogonal pair with u. Then there exist $\alpha, \beta \in \mathbb{A} \setminus \{0\}$ with $\alpha^4 = \beta^4$ such that $v(x) = \alpha \cdot u(x) + \beta \cdot u^\perp(x)$. Thus

$$(M^T \circ \delta_\pm)(x) = (\lambda \pm \mu\alpha) \cdot u(x) \pm \mu\beta \cdot u^\perp(x).$$

These two functions are in $\mathcal{A}$ if and only if both $(\lambda + \mu\alpha)^4 = \mu^4\beta^4$ and $(\lambda - \mu\alpha)^4 = \mu^4\beta^4$. That means

$$(\lambda + \mu\alpha)^4 = (\lambda - \mu\alpha)^4 \quad \Longleftrightarrow \quad \lambda\mu\alpha(\lambda^2 + \mu^2\alpha^2) = 0 \quad \Longleftrightarrow \quad \lambda = \pm i\mu\alpha$$

Then $\mu^4\beta^4 = (1+i)^4\mu^4\alpha^4$, so since all of the numbers are non-zero, we have $\beta^4 = -4\alpha^4$. This contradicts the assumption $\alpha^4 = \beta^4$. Therefore this case cannot happen and we always have $M \in \mathcal{B}_\mathcal{A}$ by the previous case. $\qquad\square$

The following makes more precise some of the arguments about entanglement types in Section 4.2 in the context of ternary functions, and extends the argument to functions in $\mathcal{A}$.

Lemma 41. *Suppose $f \in \langle \mathcal{E} \rangle$ is a non-decomposable ternary function, then f has GHZ type. Similarly, suppose $g \in \langle \mathcal{M} \rangle$ is a non-decomposable ternary function, then g has W type. Finally, suppose $h \in \mathcal{A}$ is a non-decomposable ternary function, then h has GHZ type.*

Proof. Suppose $f \in \langle \mathcal{E} \rangle$ is a non-decomposable ternary function. Non-decomposability implies $f \in \mathcal{E}$, so there exists $\mathbf{a} \in \{0,1\}^3$ such that $f(\mathbf{x}) = 0$ unless $\mathbf{x} \in \{\mathbf{a}, \bar{\mathbf{a}}\}$. Since f is non-decomposable, $f(\mathbf{a})$ and $f(\bar{\mathbf{a}})$ must both be non-zero. We can thus find matrices $A, B, C \in \{I, X\}$ such that $(A \otimes B \otimes C)|f\rangle = f_\mathbf{a}|000\rangle + f_{\bar{\mathbf{a}}}|111\rangle$, which is clearly a GHZ-type state. But this is an SLOCC operation, which does not affect the entanglement class, so f has GHZ type.

Now suppose $g \in \langle \mathcal{M} \rangle$ is a non-decomposable ternary function. Non-decomposability implies $g \in \mathcal{M}$, hence $g(\mathbf{x}) = 0$ whenever $|\mathbf{x}| > 1$. The polynomial in (15) becomes

$$(g_{000}g_{111} - g_{010}g_{101} + g_{001}g_{110} - g_{011}g_{100})^2 - 4(g_{010}g_{100} - g_{000}g_{110})(g_{011}g_{101} - g_{001}g_{111})$$
$$= (0 - 0 + 0 - 0)^2 - 4(g_{010}g_{100} - 0)(0 - 0) = 0.$$

Yet g is non-decomposable by assumption, therefore Lemma 36 implies that g must have W type.

Finally, suppose $h \in \mathcal{A}$ is a non-decomposable ternary function. Then by Definition 17, $h(\mathbf{x}) = c\,i^{l(\mathbf{x})}(-1)^{q(\mathbf{x})}\chi_{A\mathbf{x}=\mathbf{b}}(\mathbf{x})$ where $c \in \mathbb{A} \setminus \{0\}$ is a constant, l is a linear Boolean function, q is a quadratic Boolean function, and χ is a 0-1 valued indicator function for an affine subspace of $\{0,1\}^3$.

By [34], for any function $h' \in \Upsilon_n$ there exist matrices $M_1, \ldots, M_n \in \{I, H\}$, where $H = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ is the Hadamard matrix, such that $(M_1 \otimes \ldots \otimes M_n)|h'\rangle$ is everywhere non-zero. Both I and H correspond to affine functions, so by Lemmas 14 and 38, if $h' \in \mathcal{A}$ then the function corresponding to $(M_1 \otimes \ldots \otimes M_n)|h'\rangle$ is also in $\mathcal{A}$. Hence, since SLOCC operations do not affect the entanglement class, we may assume without loss of generality that h has full support by replacing it with the function transformed according to [34] if necessary. Then χ is the constant-1 function and can be ignored. Now, a SLOCC transformation by $P := \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}$ on argument x_k contributes a factor i^{x_k} to the overall function. Thus, by such transformations, we can make l trivial and remove all terms of the form x_k^2 from q without changing the entanglement. It thus suffices to consider the function $h'(x_1, x_2, x_3) = (-1)^{\gamma_{12}x_1x_2 + \gamma_{13}x_1x_3 + \gamma_{23}x_2x_3}$, where $\gamma_{12}, \gamma_{13}, \gamma_{23} \in \{0, 1\}$. Then the first term of the polynomial in (15) becomes

$$(h_{000}h_{111} - h_{010}h_{101} + h_{001}h_{110} - h_{011}h_{100})^2$$
$$= c^4\left((-1)^{\gamma_{12}+\gamma_{13}+\gamma_{23}} - (-1)^{\gamma_{13}} + (-1)^{\gamma_{12}} - (-1)^{\gamma_{23}}\right)^2$$

25

The second term becomes

$$-4(h_{010}h_{100} - h_{000}h_{110})(h_{011}h_{101} - h_{001}h_{111})$$

$$= -4c^4\Big(1 - (-1)^{\gamma_{12}}\Big)\Big((-1)^{\gamma_{13}+\gamma_{23}} - (-1)^{\gamma_{12}+\gamma_{13}+\gamma_{23}}\Big)$$

$$= -4c^4(-1)^{\gamma_{13}+\gamma_{23}}\Big(1 - (-1)^{\gamma_{12}}\Big)^2$$

Thus, if $\gamma_{12} = 0$, the polynomial in (15) is equal to

$$c^4\Big((-1)^{\gamma_{13}+\gamma_{23}} - (-1)^{\gamma_{13}} + 1 - (-1)^{\gamma_{23}}\Big)^2,$$

which is $16c^4$ if $\gamma_{13} = \gamma_{23} = 1$, and 0 otherwise. If $\gamma_{12} = 1$, the polynomial becomes

$$c^4\Big(-(-1)^{\gamma_{13}+\gamma_{23}} - (-1)^{\gamma_{13}} - 1 - (-1)^{\gamma_{23}}\Big)^2 - 16c^4(-1)^{\gamma_{13}+\gamma_{23}},$$

which is 0 if $\gamma_{13} = \gamma_{23} = 0$, and non-zero otherwise. Hence the function has GHZ-type if and only if $\gamma_{12} + \gamma_{13} + \gamma_{23} \geq 2$.

It remains to see what happens if $\gamma_{12} + \gamma_{13} + \gamma_{23} < 2$. Now, the condition (16), $(h_{000}h_{011} \neq h_{001}h_{010}) \vee (h_{101}h_{110} \neq h_{100}h_{111})$, becomes

$$\Big((-1)^{\gamma_{23}} \neq 1\Big) \vee \Big((-1)^{\gamma_{12}+\gamma_{13}} \neq (-1)^{\gamma_{12}+\gamma_{13}+\gamma_{23}}\Big),$$

which reduces to the single inequality $\gamma_{23} \neq 0$. Similarly, (17) becomes $\gamma_{13} \neq 0$ and (18) becomes $\gamma_{12} \neq 0$. Hence h either has GHZ type or it decomposes. Thus, any non-decomposable ternary function in $\mathcal{A}$ has GHZ type. □

5 Holant$^+$

Before deriving the dichotomy for Holant^c, we consider a new family of holant problems, called Holant$^+$, which fits between conservative holant problems and Holant^c: It has four freely available functions, which are all unary and include the pinning functions. Using results from quantum information theory, these four functions can be shown to be sufficient for constructing the gadgets required to apply the dichotomies in Section 3.4. Formally, for any finite $\mathcal{F} \subseteq \Upsilon$:

$$\mathsf{Holant}^+(\mathcal{F}) := \mathsf{Holant}(\mathcal{F} \cup \{\delta_0, \delta_1, \delta_+, \delta_-\}). \tag{20}$$

Note that the vectors $|0\rangle$ and $|1\rangle$ corresponding to δ_0 and δ_1 are orthogonal to each other. Similarly, the vectors $|+\rangle$ and $|-\rangle$ corresponding, up to scalar factor, to δ_+ and δ_- are orthogonal to each other. In quantum theory, the set $\{|+\rangle, |-\rangle\}$ is known as the *Hadamard basis* of $\mathbb{C}^2$, since these vectors are related to the computational basis vectors by a Hadamard transformation: $\{|+\rangle, |-\rangle\} \doteq H \circ \{|0\rangle, |1\rangle\}$, where $H = \frac{1}{\sqrt{2}}\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix}\right)$. Hence $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$, i.e. the Hadamard basis vectors differ from the vectors corresponding to δ_+ and δ_- by a factor of $\frac{1}{\sqrt{2}}$, which does not affect any of the following arguments.

In the next subsection, we first state and extend a result from quantum theory that is used in the later proofs. Specifically, we prove that if $\mathcal{F} \nsubseteq \langle \mathcal{T} \rangle$, then $S(\mathcal{F} \cup \{\delta_0, \delta_1, \delta_+, \delta_-\})$ contains a non-decomposable ternary function. It is vital for the proof to have all four unary functions available, e.g. if $\mathcal{F} = \{\mathrm{EQ}_4\}$, it would be impossible to produce a non-decomposable ternary function using only pinning.

In Section 5.2, we furthermore show that, under some mild assumptions on $\mathcal{F}$, the set $S(\mathcal{F} \cup \{\delta_0, \delta_1, \delta_+, \delta_-\})$ actually contains a *symmetric* non-decomposable ternary function. In Section 5.3, we exhibit gadget constructions for certain binary functions and show that the

assumptions of the previous section are satisfied if $\mathcal{F}$ is not one of the exceptional cases of Theorem 15. All the gadgets in these subsections are planar. To ensure that the full complexity classification works for planar holant problems, we next give a reduction between planar holant problems and planar counting CSPs in Section 5.4. This is based on results sketched in the literature, but to our knowledge the full proof has not been written out before. Finally, in Section 5.5, we combine all of the parts to prove the complexity classification for Holant^+, which holds even when restricted to the planar case.

5.1 Why these free functions?

The definition of Holant^+ is motivated by the following results from quantum theory. We first state the results in quantum terminology and translate them into holant terminology at the end of the section.

Theorem 42 ([36, Lemma on p. 296],[26])**.** *Let $|\Psi\rangle$ be an n-system genuinely entangled quantum state. For any two of the n systems, there exists a projection, onto a tensor product of states of the other $(n-2)$ systems, that leaves the two systems in an entangled state.*

Here, 'projection' means a (partial) inner product between $|\Psi\rangle$ and the tensor product of single-system states. The n systems do not have to be qubits, but for the purposes of the following arguments, it suffices to think of them as n qubits. The original proof of this statement in [36] was flawed but it was recently corrected [26]. The following corollary is not stated explicitly in either paper, but can be seen to hold by inspecting the proof in [26].

Corollary 43. *Let $|\Psi\rangle$ be an n-qubit genuinely entangled quantum state. For any two of the n qubits, there exists a projection, onto a tensor product of computational and Hadamard basis states of the other $(n-2)$ qubits, that leaves the remaining two qubits in an entangled state.*

In other words, Theorem 42 holds when the systems are restricted to qubits and the projectors are restricted to products of computational and Hadamard basis states. Here, it is crucial to have projectors taken from two bases that are linked by the Hadamard transformation: the proof applies only in that case. Intuitively, the corollary states that if n parties share a genuinely entangled n-qubit state, then this can be converted into an entangled 2-qubit state shared by two of the parties using local projections[6].

In holant terminology, the corollary corresponds to the following proposition about producing binary non-decomposable functions from a higher-arity non-decomposable function via gadgets with unary functions.

Proposition 44 (Restatement of Corollary 43)**.** *Let f be a non-decomposable function of arity $n \geq 2$. Suppose $j, k \in [n]$ with $j < k$. Then there exist $u_m \in \{\delta_0, \delta_1, \delta_+, \delta_-\}$ for all $m \in [n]\backslash\{j,k\}$ such that the following binary function is non-decomposable:*

$$g(x_j, x_k) = \sum_{x_s \in \{0,1\} \ for \ s \in [n]\backslash\{j,k\}} f(x_1, \ldots, x_n) \prod_{m \in [n]\backslash\{j,k\}} u_m(x_m).$$

This g is the effective function of the gadget in Figure 1a.

Remark. Proposition 44 can be considered an alternative definition of what it means to be non-decomposable. To see this, suppose the function $f \in \Upsilon_n$ is decomposable, i.e. there exists $1 \leq k < n$ and functions f_1, f_2 such that:

$$f(x_1, \ldots, x_n) = f_1(x_{\rho(1)}, \ldots, x_{\rho(k)}) f_2(x_{\rho(k+1)}, \ldots, x_{\rho(n)}).$$

---

[6]In the real world, these projections correspond to post-selected measurements, so without post-selection the protocol may fail in some runs.
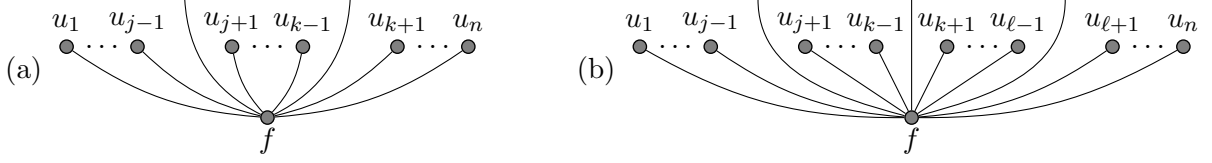
Figure 1: (a) The gadget from Proposition 44 and (b) the gadget from Theorem 45.

Choose one argument from each partition, say $x_{\rho(1)}$ and $x_{\rho(n)}$; the argument is analogous for any other choice. Then for all $u_m \in \{\delta_0, \delta_1, \delta_+, \delta_-\}$ we have

$$g(x_{\rho(1)}, x_{\rho(n)}) = \sum_{x_s \in \{0,1\} \text{ for } s \in [n] \setminus \{\rho(1), \rho(n)\}} f(x_1, \ldots, x_n) \prod_{m \in [n] \setminus \{\rho(1), \rho(n)\}} u_m(x_m)$$

$$= \left(\sum_{x_{\rho(2)}, \ldots, x_{\rho(k)} \in \{0,1\}} f_1(x_{\rho(1)}, \ldots, x_{\rho(k)}) \prod_{m=2}^{k} u_m(x_{\rho(m)}) \right)$$

$$\times \left(\sum_{x_{\rho(k+1)}, \ldots, x_{\rho(n-1)} \in \{0,1\}} f_2(x_{\rho(k+1)}, \ldots, x_{\rho(n)}) \prod_{m=k+1}^{n-1} u_m(x_{\rho(m)}) \right)$$

which is clearly decomposable. Thus if the conclusion of Proposition 44 holds for some f, then f must be non-decomposable.

We extend this proposition as follows. Note that all gadgets are planar.

Theorem 45. *Let f be a non-decomposable function of arity $n \geq 3$. Then there exist $j, k, \ell \in [n]$ with $j < k < \ell$, and $u_m \in \{\delta_0, \delta_1, \delta_+, \delta_-\}$ for all $m \in [n] \setminus \{j, k, \ell\}$, such that the following ternary function is non-decomposable:*

$$g(x_j, x_k, x_\ell) = \sum_{x_s \in \{0,1\} \text{ for } s \in [n] \setminus \{j,k,\ell\}} f(x_1, \ldots, x_n) \prod_{m \in [n] \setminus \{j,k,\ell\}} u_m(x_m).$$

This g is the effective function of the gadget in Figure 1b.

Proof. The result is proved by induction on n. If $n = 3$, f itself is the desired non-decomposable ternary function; this is the base case. Now suppose the result holds for all n satisfying $3 \leq n \leq N$.

We prove the result for $n = N + 1$ by contradiction, i.e. we begin by assuming that for some non-decomposable function f of arity $n = N + 1$ there does not exist any choice $j, k, \ell \in [n]$ with $j < k < \ell$, and $u_m \in \{\delta_0, \delta_1, \delta_+, \delta_-\}$ for all $m \in [n] \setminus \{j, k, \ell\}$ such that the function g defined in the theorem statement is non-decomposable. Note the function f cannot be identically zero since such functions are trivially decomposable.

First, consider the family of gadgets that arise by composing one input of f with one of the allowed unary functions:

$$h_{j,u}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n) := \sum_{x_j \in \{0,1\}} f(x_1, \ldots, x_n) u(x_j)$$

where $j \in [n]$ and $u \in \{\delta_0, \delta_1, \delta_+, \delta_-\}$. Note that

$$f(x_1, \ldots, x_n) = h_{j,\delta_0}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)\delta_0(x_j) + h_{j,\delta_1}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)\delta_1(x_j)$$

$$= h_{j,\delta_+}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)\delta_+(x_j) + h_{j,\delta_-}(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)\delta_-(x_j)$$

for any $j \in [n]$. Hence, if $h_{j,u}$ was identically zero for some j and u, then f would be decomposable. But we assumed f was non-decomposable, therefore the functions $h_{j,u}$ cannot be identically zero.

Furthermore, if one of the functions $h_{j,u}$ has a non-decomposable tensor factor of arity at least 3, then we can remove the other tensor factors by Lemma 9, replace f with the resulting function of arity between 3 and N (inclusive), and be done by the inductive hypothesis. Thus, for all j and u, we must have $h_{j,u} \in \langle \mathcal{T} \rangle$, i.e. $h_{j,u}$ decomposes as a tensor product of unary and binary functions.

Now by Proposition 44, we can find $u_m \in \{\delta_0, \delta_1, \delta_+, \delta_-\}$ for all $m \in [n] \setminus \{1,2\}$ such that

$$b(x_1, x_2) := \sum_{x_s \in \{0,1\} \text{ for } s \in [n] \setminus \{1,2\}} f(x_1, \ldots, x_n) \prod_{m \in [n] \setminus \{1,2\}} u_m(x_m)$$

is non-decomposable. Since b arises from h_{n,u_n} by contraction with unary functions, the arguments x_1 and x_2 must appear in the same tensor factor of h_{n,u_n}. Yet $h_{n,u_n} \in \langle \mathcal{T} \rangle$ by the argument of the previous paragraph, so we must have $h_{n,u_n}(x_1, \ldots, x_{n-1}) = b(x_1, x_2)h(x_3, \ldots, x_{n-1})$ for some $h \in \langle \mathcal{T} \rangle$, which is not identically zero. Similarly, by Proposition 44, we can find $v_m \in \{\delta_0, \delta_1, \delta_+, \delta_-\}$ for all $m \in [n] \setminus \{2,3\}$ such that

$$b'(x_2, x_3) := \sum_{x_s \in \{0,1\} \text{ for } s \in [n] \setminus \{2,3\}} f(x_1, \ldots, x_n) \prod_{m \in [n] \setminus \{2,3\}} v_m(x_m)$$

is non-decomposable. Then, analogous to the above, $h_{1,v_1}(x_2, \ldots, x_n) = b'(x_2, x_3)h'(x_4, \ldots, x_n)$ for some $h' \in \langle \mathcal{T} \rangle$, which is not identically zero.

Now consider the gadget

$$f'(x_2, \ldots, x_{n-1}) := \sum_{x_1, x_n \in \{0,1\}} f(x_1, \ldots, x_n)v_1(x_1)u_n(x_n).$$

If we perform the sum over x_n first, we find

$$f'(x_2, \ldots, x_{n-1}) = \sum_{x_1 \in \{0,1\}} h_{n,u_n}(x_1, \ldots, x_{n-1})v_1(x_1) = \sum_{x_1 \in \{0,1\}} b(x_1, x_2)h(x_3, \ldots, x_{n-1})v_1(x_1),$$
(21)

which is not identically zero since h is not, and $v'(x_2) := \sum_{x_1 \in \{0,1\}} b(x_1, x_2)v_1(x_1)$ being identically zero would imply b is decomposable. By inspection, the arguments x_2 and x_3 appear in different tensor factors of f'.

If, on the other hand, we perform the sum over x_1 first, we find

$$f'(x_2, \ldots, x_{n-1}) = \sum_{x_n \in \{0,1\}} h_{1,v_1}(x_2, \ldots, x_n)u_n(x_n) = \sum_{x_n \in \{0,1\}} b'(x_2, x_3)h'(x_4, \ldots, x_n)u_n(x_n).$$
(22)

This could be identically zero if $h'(x_4, \ldots, x_n) = h''(x_4, \ldots, x_{n-1})u_n^\perp(x_n)$ for some h'', where the function $u_n^\perp$ satisfies $\sum_{x_n \in \{0,1\}} u_n^\perp(x_n)u_n(x_n) = 0$. Yet from (21) we deduced that f' is not identically zero, so this cannot happen. Then, inspection of (22) shows that the arguments x_2 and x_3 appear in the same non-decomposable tensor factor of f'. This contradicts the finding from (21) that they appear in different tensor factors.

Hence the assumption must have been wrong and we have $h_{1,v_1} \notin \langle \mathcal{T} \rangle$ or $h_{n,u_n} \notin \langle \mathcal{T} \rangle$. Thus, by Lemma 9 and the induction hypothesis, we can realise the desired non-decomposable ternary function. $\qquad \square$

In quantum terminology, this corresponds to the following theorem.

Theorem 46 (Restatement of Theorem 45). *Let $|\Psi\rangle$ be an n-qubit genuinely entangled state with $n \geq 3$. There exists some choice of three of the n qubits and a projection of the other $(n-3)$ qubits onto a tensor product of computational and Hadamard basis states that leaves the three qubits in a genuinely entangled state.*
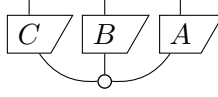
Figure 2: A 'virtual gadget' for a non-decomposable ternary function. The white vertex represents either EQ_3 or ONE_3 and the boxes represent (not necessarily symmetric) binary functions corresponding to the matrices A, B, and C, respectively.

This result, which was not previously known in the quantum information theory literature, is stronger than Corollary 43 in that we construct entangled three-qubit states rather than two-qubit ones. On the other hand, our result may not hold for arbitrary choices of three qubits: all we show is that there exists some choice of three qubits for which it does hold.

The original proof of this theorem in an earlier version of this paper was long and involved; this new shorter proof was suggested by Gachechiladze and Gühne [25].

5.2 Symmetrising ternary functions

The dichotomies given in Section 3.4 apply to symmetric ternary non-decomposable functions. The functions constructed according to Theorem 45 are ternary and non-decomposable, but they are not generally symmetric. Yet, these general ternary non-decomposable functions can be used to realise symmetric ones, possibly with the help of an additional binary non-decomposable function. We prove this by distinguishing cases according to whether the ternary non-decomposable function constructed using Theorem 45 is in the GHZ or the W entanglement class (cf. Section 4.1).

Consider a function $f \in \Upsilon_3$ which is in the GHZ class. By definition, there exist matrices $A, B, C \in GL_2(\mathbb{A})$ such that $|f\rangle = (A \otimes B \otimes C)|GHZ\rangle$, i.e.

$$f(x_1, x_2, x_3) = \sum_{y_1, y_2, y_3 \in \{0,1\}} A_{x_1,y_1} B_{x_2,y_2} C_{x_3,y_3} EQ_3(y_1, y_2, y_3).$$

We can thus draw f as the 'virtual gadget' shown in Figure 2. The 'boxes' denoting the matrices are non-symmetric to indicate that A, B, C are not in general symmetric. The white dot is assigned EQ_3. This notation is not meant to imply that the binary functions associated with A, B, C or the ternary equality function are available on their own. Instead, thinking of the function as such a composite will simply make future arguments more straightforward.

Similarly, if f is in the W class then there exist matrices $A, B, C \in GL_2(\mathbb{A})$ such that $|f\rangle = (A \otimes B \otimes C)|W\rangle$, or equivalently,

$$f(x_1, x_2, x_3) = \sum_{y_1, y_2, y_3 \in \{0,1\}} A_{x_1,y_1} B_{x_2,y_2} C_{x_3,y_3} ONE_3(y_1, y_2, y_3).$$

In this case, f can again be represented as a virtual gadget, but now the white dot in Figure 2 is assigned the function ONE_3.

In both the GHZ and the W case, three vertices assigned f can be connected to form the rotationally symmetric gadget shown in Figure 3a. In fact, since f is a function over the Boolean domain, the effective function g of that gadget is fully symmetric: its value depends only on the Hamming weight of the inputs. On the other hand, g may be decomposable (in fact, any symmetric decomposable function must be degenerate) and it may be the all-zero function. For a general non-symmetric f there are three such symmetric gadgets that can be constructed by 'rotating' f, i.e. by replacing $f(x_1, x_2, x_3)$ with $f(x_2, x_3, x_1)$ or $f(x_3, x_1, x_2)$. Rotating f in this way does not affect the planarity of the gadget. The idea leads to the following lemmas.

Lemma 47. *Suppose $f \in \Upsilon_3$ has GHZ type, i.e. $|f\rangle = (A \otimes B \otimes C)|GHZ\rangle$ for some matrices $A, B, C \in GL_2(\mathbb{A})$. Then there exists a non-decomposable symmetric ternary function $g \in S(\{f\})$, which is furthermore realisable by a planar gadget.*
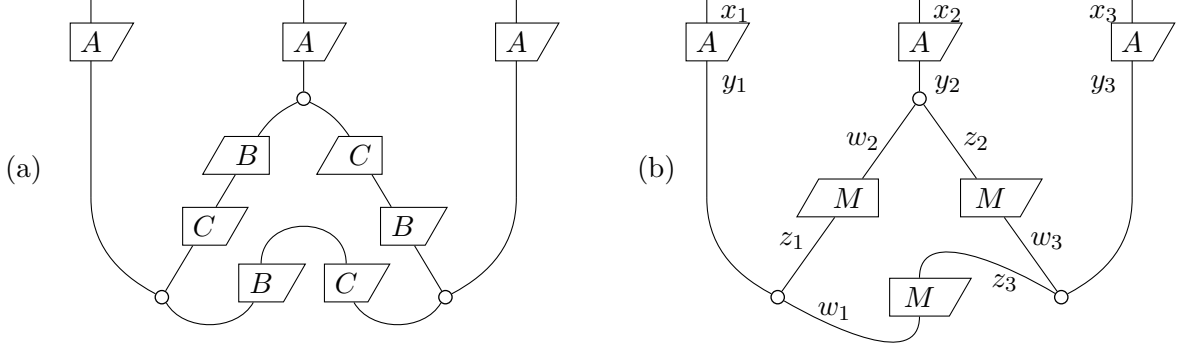
Figure 3: (a) A symmetric gadget constructed from three copies of the ternary function from Figure 2. (b) A simplified version of the same gadget, where $M := C^T B$. The variable names next to the edges are those used in (24).

Proof. The function f can be represented by a virtual gadget as in Figure 2, where the white vertex is assigned EQ_3 and the matrices A, B, and C are those appearing in the statement of the lemma. We will realise a symmetric ternary function by using the triangle gadget in Figure 3a. Note that there are three different planar versions of this gadget by cyclically permuting the inputs of f (and thus the roles of A, B, and C) in the gadget. The desired result will follow from arguing that either at least one of these three gadgets is non-decomposable and therefore yields the desired function g, or f is already symmetric. In the latter case, we simply take $g := f$.

Consider the gadget in Figure 3a. To simplify the argument, we will not parameterise each of the three matrices A, B, and C individually; instead we let

$$M := C^T B = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{23}$$

yielding the gadget in Figure 3b. The effective function of that gadget, which we will denote $h(x_1, x_2, x_3)$ is equal to

$$\sum A_{x_1,y_1} A_{x_2,y_2} A_{x_3,y_3} \mathrm{EQ}_3(y_1, z_1, w_1) \mathrm{EQ}_3(y_2, z_2, w_2) \mathrm{EQ}_3(y_3, z_3, w_3) M_{z_1,w_2} M_{z_2,w_3} M_{z_3,w_1}$$

$$= \sum_{y_1,y_2,y_3 \in \{0,1\}} A_{x_1,y_1} A_{x_2,y_2} A_{x_3,y_3} M_{y_1,y_2} M_{y_2,y_3} M_{y_3,y_1}, \tag{24}$$

where the first sum is over all $y_1, y_2, y_3, z_1, z_2, z_3, w_1, w_2, w_3 \in \{0,1\}$. Let the function h' be such that $h = A \circ h'$, then, by invertibility of A,

$$h'(y_1, y_2, y_3) = M_{y_1,y_2} M_{y_2,y_3} M_{y_3,y_1}.$$

Plugging in the values from (23), we find that $h' = [a^3, abc, bcd, d^3]$. Since h and h' are connected by a holographic transformation, they must be in the same entanglement class: in particular, h is non-decomposable if and only if h' is non-decomposable. Hence we may work with h' instead of h for the remainder of the proof.

Recall from Section 4.1 that a symmetric ternary function is non-decomposable if and only if it has either GHZ type or W type. By Lemma 37, h' has GHZ type if and only if:

$$P_{h'} := a^2 d^2 (ad + 3bc)(ad - bc)^3 \neq 0.$$

It has W type if $P_{h'} = 0$ and furthermore

$$\left((abc)^2 \neq a^3 bcd \right) \vee \left((bcd)^2 \neq abcd^3 \right). \tag{25}$$

If neither of these conditions is satisfied, h' is decomposable (and in fact degenerate).

Now, as M is invertible, we have $ad - bc \neq 0$. Thus, h' fails to have GHZ type only if at least one of a, d, or $(ad + 3bc)$ is zero. We consider the cases individually.

31

- Suppose $a = 0$. Then $P_{h'} = 0$ and (25) becomes $(0 \neq 0) \vee (bcd \neq 0)$. Hence h' has W type if $bcd \neq 0$ and is degenerate otherwise. Since M is invertible, $a = 0$ implies $bc \neq 0$. Thus, h' is degenerate (in fact, it is identically zero) if $a = d = 0$ and it has W type otherwise.

- Suppose $a \neq 0$ and $d = 0$. Then $P_{h'} = 0$ and (25) becomes $(abc \neq 0) \vee (0 \neq 0)$. Hence h' has W type if $abc \neq 0$. But invertibility of M together with $d = 0$ implies that $bc \neq 0$, so in this case a, b, c are all guaranteed to be non-zero. Therefore h' always has W type.

- Suppose $a, d \neq 0$ and $ad + 3bc = 0$, i.e. $bc = -\frac{1}{3}ad$. Then $P_{h'} = 0$ and by substituting for bc, (25) becomes
$$\left(\frac{1}{9}a^4 d^2 \neq -\frac{1}{3}a^4 d^2 \right) \vee \left(\frac{1}{9}a^2 d^4 \neq -\frac{1}{3}a^2 d^4 \right),$$
which is true for all $a, d \neq 0$. Therefore h' always has W type.

By combining the three cases, we find that h' is degenerate if and only if $a = d = 0$, or equivalently if and only if $(C^T B)_{00} = (C^T B)_{11} = 0$.

As noted before, there are three different planar gadgets that can be constructed from the same non-symmetric ternary function, by 'rotating' it. Assume that all three gadget constructions yield a decomposable function. For the original gadget, this means that $(C^T B)_{00} = (C^T B)_{11} = 0$. For the rotated versions of the gadget, the assumption furthermore implies that

$$(B^T A)_{00} = 0 = (B^T A)_{11} \quad \text{and} \quad (A^T C)_{00} = 0 = (A^T C)_{11}.$$

These equalities indicate that $C^T B$, $B^T A$, and $A^T C$ are purely off-diagonal matrices. Additionally, since A, B, C are invertible, $C^T B$, $B^T A$, and $A^T C$ must also be invertible. Hence there exist invertible diagonal matrices $D_1, D_2, D_3 \in \mathrm{GL}_2(\mathbb{A})$ such that:

$$B^T A = X D_1, \qquad C^T B = X D_2, \quad \text{and} \quad A^T C = X D_3. \tag{26}$$

By rearranging the second equation of (26), we find $B = (C^T)^{-1} X D_2$. Now, by transposing, rearranging, and then inverting, the third equation of (26) is equivalent to

$$C^T A = D_3 X \iff C^T = D_3 X A^{-1} \iff (C^T)^{-1} = A X D_3^{-1},$$

so $B = A X D_3^{-1} X D_2$. Similarly, transposing and then rearranging the second equation of (26) yields $C = (B^T)^{-1} D_2 X$. By rearranging and inverting, the first equation of (26) is equivalent to

$$B^T = X D_1 A^{-1} \iff (B^T)^{-1} = A D_1^{-1} X,$$

so $C = A D_1^{-1} X D_2 X$.

Both $D_1^{-1} X D_2 X$ and $X D_3^{-1} X D_2$ are diagonal matrices; write them as

$$D_C = D_1^{-1} X D_2 X = \begin{pmatrix} \gamma_0 & 0 \\ 0 & \gamma_1 \end{pmatrix} \quad \text{and} \quad D_B = X D_3^{-1} X D_2 = \begin{pmatrix} \beta_0 & 0 \\ 0 & \beta_1 \end{pmatrix} \tag{27}$$

for some $\beta_0, \beta_1, \gamma_0, \gamma_1 \in \mathbb{A} \setminus \{0\}$. Then $B = A D_B$ and $C = A D_C$, so:

$$|f\rangle = (A \otimes B \otimes C) |\mathrm{GHZ}\rangle = A^{\otimes 3}(I \otimes D_B \otimes D_C) |\mathrm{GHZ}\rangle = A^{\otimes 3} \left(\beta_0 \gamma_0 |000\rangle + \beta_1 \gamma_1 |111\rangle \right), \tag{28}$$

where I is the 2 by 2 identity matrix. Hence the assumption that all three gadgets are decomposable implies that $f = A \circ [\beta_0\gamma_0, 0, 0, \beta_1\gamma_1]$, i.e. the assumption implies that f is already symmetric.

We have shown that either one of the three possible planar triangle gadgets yields a symmetric non-degenerate function, or f itself is already symmetric. Thus there always exists a non-degenerate symmetric ternary function in $S(\{f\})$ which is realisable by a planar gadget. $\qquad \square$
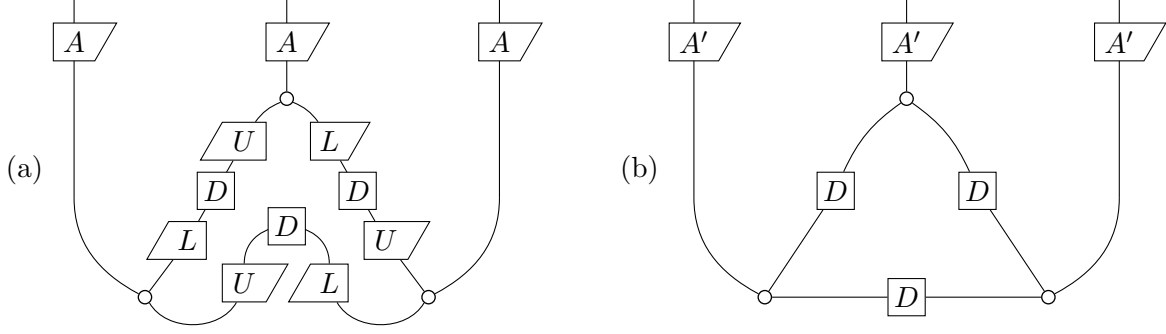
32

Figure 4: (a) The symmetrisation gadget after the matrix $C^T B$ has been converted into LDU decomposition, where L is lower triangular, D is diagonal (and hence symmetric), and U is upper triangular. (b) If the white vertices are assigned the function ONE_3, the triangular matrices can be replaced by a different matrix on the outer leg, which is absorbed into A to form A'.

Lemma 48. *Suppose $f \in \Upsilon_3$ has W type, i.e. $|f\rangle = (A \otimes B \otimes C)\,|W\rangle$ for some matrices $A, B, C \in \mathrm{GL}_2(\mathbb{A})$. If $f \notin (K \circ \mathcal{M}) \cup (KX \circ \mathcal{M})$, there exists a non-degenerate symmetric ternary function $g \in S(\{f\})$. The function g has GHZ type and is realisable by a planar gadget.*

Proof. Since f has W type, we can write $|f\rangle = (A \otimes B \otimes C)\,|W\rangle$: i.e. the function f can be thought of as the 'virtual gadget' given in Figure 2, where the white dot is now assigned ONE_3. We can therefore combine three copies of f into the triangle gadget given in Figure 3a, analogous to the previous lemma. By basic linear algebra, we can apply the LDU decomposition to $C^T B$:

$$C^T B = LDU = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix} \begin{pmatrix} b & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix},$$

where $a, b, c, d \in \mathbb{A}$ with $b, c \neq 0$ and $L = \left(\begin{smallmatrix} 1 & 0 \\ a & 1 \end{smallmatrix}\right)$ is lower triangular, $D = \left(\begin{smallmatrix} b & 0 \\ 0 & c \end{smallmatrix}\right)$ is diagonal, and $U = \left(\begin{smallmatrix} 1 & d \\ 0 & 1 \end{smallmatrix}\right)$ is upper triangular. This is illustrated in Figure 4a.

Now, note that each of the white dots assigned ONE_3 is connected to a copy of U and a transposed copy of L. It is straightforward to show that

$$(I \otimes U \otimes L^T)\,|W\rangle = (a + d)\,|000\rangle + |W\rangle = \left(\begin{pmatrix} 1 & a + d \\ 0 & 1 \end{pmatrix} \otimes I \otimes I\right)|W\rangle,$$

so the diagram can be transformed to the one of Figure 4b, where $A' := A \left(\begin{smallmatrix} 1 & a+d \\ 0 & 1 \end{smallmatrix}\right)$.

The effective function of the gadget in Figure 4b, which we will denote $g(x_1, x_2, x_3)$, equals

$$\sum A'_{x_1, y_1} A'_{x_2, y_2} A'_{x_3, y_3} \text{ONE}_3(y_1, z_1, w_1) \text{ONE}_3(y_2, z_2, w_2) \text{ONE}_3(y_3, z_3, w_3) D_{z_1, w_2} D_{z_2, w_3} D_{z_3, w_1},$$

where the sum is over all $y_1, y_2, y_3, z_1, z_2, z_3, w_1, w_2, w_3 \in \{0, 1\}$. Let $g' := (A')^{-1} \circ g$, then

$$g'(y_1, y_2, y_3) = \sum_{z_1, z_2, z_3 \in \{0,1\}} \text{ONE}_3(y_1, z_1, z_3) \text{ONE}_3(y_2, z_2, z_1) \text{ONE}_3(y_3, z_3, z_2) D_{z_1, z_1} D_{z_2, z_2} D_{z_3, z_3},$$

where we have used the property that D is diagonal.

Recall that ONE_3 is the 'perfect matchings' constraint. This means the gadget for g' is a symmetric matchgate on three vertices, where every internal edge that is in the matching contributes a factor c to the weight, and every internal edge that is not in the matching contributes a factor b to the weight. Clearly, $g'(1, 1, 1) = b^3$ since if the three external edges are in the matching, there can be no internal edges in the matching. As a matchgate, g' must satisfy a parity condition: in particular, given the odd number of vertices, it is 0 on inputs

Figure 5: Gadget for constructing a ternary function that is not in $K \circ \mathcal{M}$ (or $KX \circ \mathcal{M}$). The degree-3 vertex is assigned the function f and the degree-2 vertex is assigned the function h, with the second input of h connected to the first input of f.

of even Hamming weight. The remaining value on inputs of Hamming weight 1, having one external edge and thus one internal edge to make a matching, is b^2c. Combining these, we have $g' = [0, b^2c, 0, b^3]$. Note that $|g\rangle$ and $|g'\rangle$ are by definition equivalent under SLOCC, i.e. g and g' are in the same entanglement class. We may therefore reason about g' instead of g.

By Lemma 37, g' has GHZ type if and only if

$$0 \neq (0 - 0)^2 - 4\left((b^2c)^2 - 0\right)\left(0 - b^2cb^3\right) = 4b^9c^3.$$

Yet $b, c \neq 0$, so g' always has GHZ type. The gadget is planar. $\qquad \square$

The following lemma shows an analogous result for the case where the ternary function $f \in K \circ \mathcal{M}$ or $f \in KX \circ \mathcal{M}$. In that case, some additional support from a binary function outside the respective set is needed; we will show in the next section that such a binary function can be realised unless all functions are in $\langle K \circ \mathcal{M} \rangle$ or $\langle KX \circ \mathcal{M} \rangle$.

Lemma 49. *Let $\mathcal{M}'$ be one of $K \circ \mathcal{M}$ and $KX \circ \mathcal{M}$. Suppose $f \in \Upsilon_3 \cap \mathcal{M}'$ and $h \in \Upsilon_2 \setminus \langle \mathcal{M}' \rangle$, then there exists a non-degenerate symmetric ternary function $g \in S(\{f, h\})$. In both cases, g has GHZ type and is realisable by a planar gadget.*

Proof. Note that by Lemma 41 and the property that holographic transformations do not affect the entanglement class, f has W type in both cases, i.e. $|f\rangle = (A \otimes B \otimes C)|W\rangle$ for some matrices $A, B, C \in \mathrm{GL}_2(\mathbb{A})$.

Now first suppose $f \in K \circ \mathcal{M}$ and $h \notin \langle K \circ \mathcal{M} \rangle$. Note that any binary function h with this property is non-degenerate (and thus non-decomposable), since $\Upsilon_1 \subseteq K \circ \mathcal{M}$. Let $f' := K^{-1} \circ f \in \mathcal{M}$ and $h' := K^{-1} \circ h \notin \mathcal{M}$, then we can write

$$f' = \begin{pmatrix} f_{000} & f_{001} & f_{010} & 0 \\ f_{100} & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad h' = \begin{pmatrix} h_{00} & h_{01} \\ h_{10} & h_{11} \end{pmatrix},$$

where $f_{001}f_{010}f_{100} \neq 0$ by non-decomposability of f', $h_{00}h_{11} - h_{01}h_{10} \neq 0$ by non-decomposability of h', and $h_{11} \neq 0$ because $h' \notin \mathcal{M}$.

Consider the gadget in Figure 5, where the second input of h is connected to the first input of f. The effective function associated with this gadget is

$$f''(x_1, x_2, x_3) = \sum_{y \in \{0,1\}} h(x_1, y)f(y, x_2, x_3)$$

$$= \sum_{y, z_1, \ldots, z_5 \in \{0,1\}} K_{x_1, z_1} K_{y, z_2} h'(z_1, z_2) K_{y, z_3} K_{x_2, z_4} K_{x_3, z_5} f'(z_3, z_4, z_5)$$

$$= \sum_{z_1, \ldots, z_5 \in \{0,1\}} K_{x_1, z_1} K_{x_2, z_4} K_{x_3, z_5} h'(z_1, z_2) f'(z_3, z_4, z_5) \sum_{y \in \{0,1\}} K_{y, z_2} K_{y, z_3}$$

$$= \sum_{z_1, z_4, z_5 \in \{0,1\}} K_{x_1, z_1} K_{x_2, z_4} K_{x_3, z_5} \sum_{z_2, z_3 \in \{0,1\}} 2 \cdot h'(z_1, z_2) f'(z_3, z_4, z_5) \mathrm{NEQ}(z_2, z_3).$$

Let $f''' := \frac{1}{2}(K^{-1} \circ f'')$ or equivalently $f'' = 2(K \circ f''')$, then

$$f''' = \begin{pmatrix} f_{000}h_{01} + f_{100}h_{00} & f_{001}h_{01} & f_{010}h_{01} & 0 \\ f_{000}h_{11} + f_{100}h_{10} & f_{001}h_{11} & f_{010}h_{11} & 0 \end{pmatrix}.$$

Note that $f'' \in K \circ \mathcal{M}$ if and only if $f''' \in \mathcal{M}$, and $f'' \in KX \circ \mathcal{M}$ if and only if $f''' \in X \circ \mathcal{M}$. We now show $f'' \notin (K \circ \mathcal{M}) \cup (KX \circ \mathcal{M})$.

Assume for a contradiction that $f'' \in K \circ \mathcal{M}$, i.e. $f''' \in \mathcal{M}$. This implies that $0 = f'''(1,0,1) = f_{001}h_{11}$ and $0 = f'''(1,1,0) = f_{010}h_{11}$. But, as stated above, f being non-decomposable and h not being in $K \circ \mathcal{M}$ imply that f_{001}, f_{010} and h_{11} must be non-zero. Therefore $f''' \notin \mathcal{M}$ and $f'' \notin K \circ \mathcal{M}$.

Similarly, assume for a contradiction that $f'' \in KX \circ \mathcal{M}$, i.e. $f''' \in X \circ \mathcal{M}$. This implies

$$0 = f'''(0,0,0) = f_{000}h_{01} + f_{100}h_{00}$$
$$0 = f'''(0,0,1) = f_{001}h_{01}$$
$$0 = f'''(0,1,0) = f_{010}h_{01}$$
$$0 = f'''(1,0,0) = f_{000}h_{11} + f_{100}h_{10}.$$

Now, the condition $f_{001}h_{01} = 0$ implies $h_{01} = 0$ since $f_{001} \neq 0$ by non-decomposability of f. Then the first equality reduces to $0 = f_{100}h_{00}$, which implies $h_{00} = 0$ since $f_{100} \neq 0$ by non-decomposability of f. Yet $h_{01} = h_{00} = 0$ would imply that h is degenerate, contradicting the assumption of the lemma. Thus $f''' \notin X \circ \mathcal{M}$ and $f'' \notin KX \circ \mathcal{M}$.

Hence, we may replace f by f'' and proceed as in Lemma 48. Since $f'' \in S(\{f,h\})$ and f'' is realisable by a planar gadget, the function $g \in S(\{f''\})$ that results from applying this lemma to f'' is in $S(\{f,h\})$ and g is again realisable by a planar gadget.

Now suppose $f \in KX \circ \mathcal{M}$ and $h \notin \langle KX \circ \mathcal{M} \rangle$. As before, let $f' := K^{-1} \circ f$ and $h' := K^{-1} \circ h$, then $f' \in X \circ \mathcal{M}$ and $h' \notin \langle X \circ \mathcal{M} \rangle$. The properties of these functions differ from the ones in the previous case by bit flips on all inputs. Otherwise the argument is the same. $\square$

5.3 Realising binary functions

We have shown in the previous section that it is possible to realise a non-decomposable ternary symmetric function from an arbitrary non-decomposable ternary function under some mild further assumptions. Now, we show that if the full set of functions $\mathcal{F}$ is not a subset of $\langle K \circ \mathcal{M} \rangle$, there exists a planar gadget in $S(\mathcal{F} \cup \{\delta_0, \delta_1, \delta_+, \delta_-\})$ whose effective function g is binary, symmetric, non-decomposable and satisfies $g \notin \langle K \circ \mathcal{M} \rangle$. An analogous result holds with KX instead of K.

Lemma 50. *Suppose $f \in \Upsilon_n$ satisfies $f(1,1,a_3,\ldots,a_n) \neq 0$ for some $a_3,\ldots,a_n \in \{0,1\}$. Suppose furthermore that there exist functions $u_3,\ldots,u_n \in \mathcal{U}$ such that setting*

$$f'(x_1,x_2) := \sum_{x_3,\ldots,x_n \in \{0,1\}} f(x_1,\ldots,x_n) \prod_{j=3}^{n} u_j(x_j)$$

implies $f'(0,0)f'(1,1) - f'(0,1)f'(1,0) \neq 0$. Then $f \notin \langle \mathcal{M} \rangle$. Conversely, if $f \in \Upsilon_n \setminus \langle \mathcal{M} \rangle$, then there exists some permutation $\rho : [n] \to [n]$ such that f_ρ satisfies the above properties.

Proof. Consider the first part of the lemma and assume for a contradiction that $f \in \langle \mathcal{M} \rangle$. Then there exists a decomposition of f where each factor is in $\mathcal{M}$. In particular, the first and second arguments of f must belong to different factors: if they belonged to the same factor, that factor would be a function which is non-zero on an input of Hamming weight at least 2, so it could not be in $\mathcal{M}$. Hence there exists some permutation $\rho : \{3,4,\ldots,n\} \to \{3,4,\ldots,n\}$ and $k \in \{3,4,\ldots,n\}$ such that

$$f_\rho(x_1,\ldots,x_n) = f_1(x_1,x_{k+1},\ldots,x_n)f_2(x_2,\ldots,x_k),$$

where f_1, f_2 may be further decomposable. But then

$$f'(x_1,x_2) = \sum_{x_3,\ldots,x_n \in \{0,1\}} f_1(x_1,x_{k+1},\ldots,x_n)f_2(x_2,\ldots,x_k) \prod_{j=3}^{n} u_j(x_j) = w_1(x_1)w_2(x_2),$$

35

where

$$w_1(x_1) := \sum_{x_{k+1},\ldots,x_n \in \{0,1\}} f_1(x_1, x_{k+1}, \ldots, x_n) \prod_{j=k+1}^{n} u_j(x_j)$$

$$w_2(x_2) := \sum_{x_3,\ldots,x_k \in \{0,1\}} f_2(x_2, \ldots, x_k) \prod_{j=3}^{k} u_j(x_j).$$

This implies that $f'(0,0)f'(1,1) - f'(0,1)f'(1,0) = 0$, contradicting the assumption in the lemma. Thus f cannot be in $\langle \mathcal{M} \rangle$.

For the second part of the lemma, assume $f \in \Upsilon_n \setminus \langle \mathcal{M} \rangle$. In particular, f is not the all-zero function, so it has a decomposition. If all the factors in the decomposition of f are in $\mathcal{M}$, then $f \in \langle \mathcal{M} \rangle$. Hence there exists a non-decomposable factor g of f which is not in $\mathcal{M}$. Without loss of generality, assume that g contains exactly the first k arguments of f for some $k \in [n]$, otherwise permute the arguments. As $\mathcal{M}$ contains all unaries, g must be a function of arity $k \geq 2$. Furthermore, there must exist some bit string $\mathbf{a} \in \{0,1\}^k$ of Hamming weight at least 2 such that $g(\mathbf{a}) \neq 0$. Without loss of generality, assume $a_1 = a_2 = 1$, otherwise permute the arguments. Then f satisfies the first property. If $f = g$, i.e. f itself is non-decomposable, the second property is satisfied by Proposition 44. Otherwise, by Lemma 9, there exist $u_{k+1}, \ldots, u_n \in \{\delta_0, \delta_1\}$ such that

$$g(x_1, \ldots, x_k) = \sum_{x_{k+1},\ldots,x_n \in \{0,1\}} f(x_1, \ldots, x_n) \prod_{j=k+1}^{n} u_j(x_j).$$

We then find that the second property is satisfied by combining this equation with an application of Proposition 44 to g. $\qquad \square$

Lemma 51. *Let $\mathcal{M}'$ be one of $K \circ \mathcal{M}$ and $KX \circ \mathcal{M}$. Suppose $f \in \Upsilon \setminus \langle \mathcal{M}' \rangle$. Then there exists $g \in S(\{f, \delta_0, \delta_1, \delta_+, \delta_-\})$ which is binary, non-decomposable, and satisfies $g \notin \langle \mathcal{M}' \rangle$. Furthermore, this function is realised by a planar gadget.*

Proof. First, suppose $\mathcal{M}' = K \circ \mathcal{M}$. Note that $\Upsilon_1 \subseteq K \circ \mathcal{M} \subseteq \langle K \circ \mathcal{M} \rangle$, so $f \notin \langle K \circ \mathcal{M} \rangle$ implies arity$(f) \geq 2$. Furthermore, $\langle \Upsilon_1 \rangle \subseteq \langle K \circ \mathcal{M} \rangle$, thus any binary function which is not in $\langle K \circ \mathcal{M} \rangle$ must be non-decomposable.

We will show that we can realise a gadget like that in Figure 1a, whose effective function g is non-decomposable and satisfies $g \notin \langle K \circ \mathcal{M} \rangle$. As the gadget consists of a number of unary functions connected to one larger-arity function, planarity is assured even if we arbitrarily permute the arguments of f at intermediate stages of the proof.

The result will be proved by induction on the arity of f.

The base case is arity$(f) = 2$, in which case we may take $g = f$. This function is binary, non-decomposable, a gadget in $S(\{f, \delta_0, \delta_1, \delta_+, \delta_-\})$ and it is not in $\langle K \circ \mathcal{M} \rangle$. Furthermore, the gadget is trivially planar, so g has all the required properties.

Now suppose the desired result has been proved for functions of arity at most n, and suppose arity$(f) = n + 1$. We distinguish cases according to the decomposability of f.

Case 1: Suppose f is decomposable and consider some decomposition of f. If all factors are in $\langle K \circ \mathcal{M} \rangle$, then f is in $\langle K \circ \mathcal{M} \rangle$, so there must be a factor $f' \notin \langle K \circ \mathcal{M} \rangle$. By Lemma 9, $f' \in S(\{f, \delta_0, \delta_1, \delta_+, \delta_-\})$. Thus, we are done by the inductive hypothesis.

Case 2: Suppose f is non-decomposable. We will show that there exists $u \in \{\delta_0, \delta_1, \delta_+, \delta_-\}$ and $k \in [n+1]$ such that $\sum_{x_k \in \{0,1\}} f(x_1, \ldots, x_{n+1})u(x_k)$ is not in $\langle K \circ \mathcal{M} \rangle$. It is simpler to work with $\langle \mathcal{M} \rangle$ than $\langle K \circ \mathcal{M} \rangle$, so let $h = K^{-1} \circ f$; then $h \notin \langle \mathcal{M} \rangle$. Thus, by Lemma 50, there exists a permutation $\rho : [n+1] \to [n+1]$ such that $h_\rho(1, 1, a_3, \ldots, a_{n+1}) \neq 0$ for some

$a_3, \ldots, a_{n+1} \in \{0, 1\}$ and there exist functions $u_3, \ldots, u_{n+1} \in \mathcal{U}$ such that letting

$$h'(x_1, x_2) := \sum_{x_3, \ldots, x_{n+1} \in \{0,1\}} h_\rho(x_1, \ldots, x_{n+1}) \prod_{j=3}^{n+1} u_j(x_j) \tag{29}$$

implies $h'(0,0)h'(1,1) - h'(0,1)h'(1,0) \neq 0$. Note that

$$h'(x_1, x_2) = \sum_{x_3, \ldots, x_{n+1} \in \{0,1\}} (K^{-1} \circ f_\rho)(x_1, \ldots, x_{n+1}) \prod_{j=3}^{n+1} u_j(x_j)$$

$$= K^{-1} \circ \left(\sum_{x_3, \ldots, x_{n+1} \in \{0,1\}} f_\rho(x_1, \ldots, x_{n+1}) \prod_{j=3}^{n+1} ((K^{-1})^T \circ u_j)(x_j) \right)$$

and holographic transformations do not affect whether a function is decomposable. Thus, by applying Proposition 44 to f_ρ, we find that it suffices to take

$$(K^{-1})^T \circ u_j \in \{\delta_0, \delta_1, \delta_+, \delta_-\} \quad \Leftrightarrow \quad u_j \in \{\delta_+, \delta_-, \delta_i, \delta_{-i}\},$$

where $\delta_i := [1, i]$ and $\delta_{-i} := [1, -i]$ and we have ignored some scalar factors which, by Lemma 3, do not affect the complexity.

For each $v \in \{\delta_+, \delta_-, \delta_i, \delta_{-i}\}$ define

$$h_v(x_1, \ldots, x_n) := \sum_{x_{n+1} \in \{0,1\}} h_\rho(x_1, \ldots, x_{n+1}) v(x_{n+1}).$$

We now argue that for at least one $v \in \{\delta_+, \delta_-, \delta_i, \delta_{-i}\}$, the function h_v is not in $\langle \mathcal{M} \rangle$. Write $v = [1, \alpha]$, where $\alpha \in \{1, -1, i, -i\}$.

First, consider the value $h_v(1, 1, a_3, \ldots, a_n)$, where $a_3, \ldots, a_n$ are the above values resulting from applying Lemma 50 to h. Then

$$h_v(1, 1, a_3, \ldots, a_n) = h_\rho(1, 1, a_3, \ldots, a_n, 0) + \alpha h_\rho(1, 1, a_3, \ldots, a_n, 1) \tag{30}$$

is a linear polynomial in the variable α. Furthermore, this polynomial is not identically zero since $h_\rho(1, 1, a_3, \ldots, a_{n+1}) \neq 0$. Hence, this expression vanishes for at most one value of α, i.e. one of the h_v.

Secondly, let

$$h'_u(x_1, x_2) := \sum_{x_3, \ldots, x_{n+1} \in \{0,1\}} h_\rho(x_1, \ldots, x_{n+1}) v(x_{n+1}) \prod_{j=3}^{n} u_j(x_j),$$

where $u_j \in \{\delta_+, \delta_-, \delta_i, \delta_{-i}\}$ are the unary functions determined by applying Lemma 50 to h as in (29). Then

$$h'_v(0,0)h'_v(1,1) - h'_v(0,1)h'_v(1,0) \tag{31}$$

is a quadratic polynomial in α which is not identically zero. Thus this polynomial vanishes for at most two distinct values of α, i.e. two of the h_v.

Therefore, there must be at least one h_v such that both (30) and (31) are non-zero. By Lemma 50, this h_v is not in $\langle \mathcal{M} \rangle$. Furthermore,

$$(K \circ h_v)(x_1, \ldots, x_n) = \sum_{x_{n+1} \in \{0,1\}} f_\rho(x_1, \ldots, x_{n+1}) v'(x_{n+1}),$$

where $v' = (K^{-1})^T \circ v \in \{\delta_0, \delta_1, \delta_+, \delta_-\}$. Thus, $K \circ h_v \in S(\{f, \delta_0, \delta_1, \delta_+, \delta_-\}) \setminus \langle K \circ \mathcal{M} \rangle$ and we are done by the inductive hypothesis.

Now suppose $\mathcal{M}' = KX \circ \mathcal{M}$. The argument of Case 1 is the same as before, with $KX \circ \mathcal{M}$ instead of $K \circ \mathcal{M}$. In Case 2, again define $h := K^{-1} \circ f$, then the properties of this function differ from the ones written out above by a bit flip on all inputs. Note that, up to scalar factor, the set $\{\delta_+, \delta_-, \delta_i, \delta_{-i}\}$ is invariant under bit flip. Thus, the argument is analogous to before. $\quad \square$

5.4 Interreducing planar holant problems and planar counting CSPs

The interreducibility of certain bipartite holant problems and counting CSPs, as in Theorem 28, will be a crucial ingredient in our hardness proof. We therefore need to ensure this result holds in the planar case.

Recall that $\#\mathsf{R_3\text{-}CSP}\,(\mathcal{F}) \equiv_T \mathsf{Holant}\,(\mathcal{F} \mid \{\mathrm{EQ}_1, \mathrm{EQ}_2, \mathrm{EQ}_3\})$. The following lemma will be useful.

Lemma 52 ([18, Lemma 6.1]). *Let $g \in \Upsilon_2$ be a non-degenerate binary function. Then, for any finite $\mathcal{F} \subseteq \Upsilon$ containing g, we have $\#\mathsf{R_3\text{-}CSP}\,(\mathcal{F} \cup \{\mathrm{EQ}_2\}) \leq_T \#\mathsf{R_3\text{-}CSP}\,(\mathcal{F})$.*

In fact, this result can straightfordwardly be extended to the following.

Lemma 53. *Let $g \in \Upsilon_2$ be a non-degenerate binary function. Suppose $\mathcal{G}_1, \mathcal{G}_2 \subseteq \Upsilon$ are finite, then*

$$\mathsf{Pl\text{-}Holant}\,(\mathcal{G}_1 \cup \{g, \mathrm{EQ}_2\} \mid \mathcal{G}_2 \cup \{\mathrm{EQ}_1, \mathrm{EQ}_2, \mathrm{EQ}_3\})$$
$$\leq_T \mathsf{Pl\text{-}Holant}\,(\mathcal{G}_1 \cup \{g\} \mid \mathcal{G}_2 \cup \{\mathrm{EQ}_1, \mathrm{EQ}_2, \mathrm{EQ}_3\})\,.$$

Proof. The proof is analogous to that of [18, Lemma 6.1], noting that the constructions used in gadgets and in polynomial interpolation are planar, and that adding more functions on the RHS does not affect the argument. □

The following result about a polynomial-time reduction from any counting CSP to some problem in $\#\mathsf{P}$ is known, see e.g. [13, p. 212]. Nevertheless, we have not been able to find an explicit proof, so we give one here for completeness. This proof is based on a similar reduction for graph homomorphism problems [10, Lemma 7.1].

First, we define the field within which we will be working, and the computational problem to which the counting CSP will be reduced. For any finite $\mathcal{F} \subseteq \Upsilon$, let $A_{\mathcal{F}}$ be the set of algebraic complex numbers that appear as a value of some function in $\mathcal{F}$:

$$A_{\mathcal{F}} := \left\{ z \in \mathbb{A} \,\middle|\, \exists f \in \mathcal{F} \text{ and } \mathbf{x} \in \{0,1\}^{\mathrm{arity}(f)} \text{ such that } f(\mathbf{x}) = z \right\}. \tag{32}$$

Since $\mathcal{F}$ is fixed and finite, the set $A_{\mathcal{F}}$ is also fixed and finite. Let $\mathbb{Q}(A_{\mathcal{F}})$ be the algebraic extension of the field of rational numbers by the numbers in $A_{\mathcal{F}}$. Note that, given an instance (V, C) of $\#\mathsf{CSP}(\mathcal{F})$, the weight $\mathrm{wt}_{(V,C)}(\xi)$ associated with any assignment $\xi : V \to \{0,1\}$ is in $\mathbb{Q}(A_{\mathcal{F}})$, as is the total weight $Z_{(V,C)}$. We may thus define the following counting problem:

Name $\mathsf{COUNT}(\mathcal{F})$
Instance A tuple $((V, C), x)$, where V is a finite set of variables, C is a finite set of constraints over $\mathcal{F}$, and $x \in \mathbb{Q}(A_{\mathcal{F}})$.
Output $\#_{\mathcal{F}}((V, C), x) = \big| \{\text{assignments } \xi : V \to \{0,1\} \mid \mathrm{wt}_{(V,C)}(\xi) = x\} \big|$.

$\mathsf{COUNT}(\mathcal{F})$ is the problem of counting the number of accepting paths of a Turing machine that accepts an input $((V, C), x, \xi)$ if and only if $\mathrm{wt}_{(V,C)}(\xi) = x$. Here, $\xi : V \to \{0,1\}$ has to be an assignment of values to the variables V, otherwise the input is rejected. Given an assignment ξ, computing the associated weight in the fixed algebraic extension field $\mathbb{Q}(A_{\mathcal{F}})$ can be done in time polynomial in the size of (V, C). Furthermore, numbers within $\mathbb{Q}(A_{\mathcal{F}})$ can be compared efficiently. Therefore the Turing machine runs in time polynomial in the size of its input, and $\mathsf{COUNT}(\mathcal{F})$ is in $\#\mathsf{P}$.

Lemma 54. *Let $\mathcal{F} \subseteq \Upsilon$ be finite. Then $\#\mathsf{CSP}(\mathcal{F}) \leq_T \mathsf{COUNT}(\mathcal{F})$.*

Proof. Consider an instance (V, C) of $\#\mathsf{CSP}(\mathcal{F})$, where V is a finite set of variables and C is a finite set of constraints over $\mathcal{F}$. Since $\mathcal{F}$ is a fixed finite set, its elements can be enumerated in some order $f_1, \ldots, f_m$, where $m := |\mathcal{F}|$ is a constant. For each $j \in [m]$, define $a_j := \mathrm{arity}(f_j)$ as a short-hand. Let $n := |C|$ be the number of constraints, and define the following set of algebraic complex numbers:

$$\mathcal{X} = \left\{ \prod_{j \in [m]} \prod_{\mathbf{x}_j \in \{0,1\}^{a_j}} (f_j(\mathbf{x}_j))^{k_{j,\mathbf{x}_j}} \;\middle|\; k_{j,\mathbf{x}_j} \in \mathbb{Z}_{\geq 0} \text{ and } \sum_{j \in [m]} \sum_{\mathbf{x}_j \in \{0,1\}^{a_j}} k_{j,\mathbf{x}_j} = n \right\}.$$

Each element of $\mathcal{X}$ is uniquely determined by the integers $k_{j,\mathbf{x}_j}$, and there are $M := \sum_{j \in [m]} 2^{a_j}$ such integers. Thus, the elements of $\mathcal{X}$ are in bijection with the M-tuples of non-negative integers satisfying the property that the sum of all elements of the tuple is n. The set of all such M-tuples is exactly the support set of a multinomial distribution with n trials and M possible outcomes for each trial; therefore

$$|\mathcal{X}| = \binom{n + M - 1}{M - 1} \leq \frac{(n + M - 1)^{M-1}}{(M - 1)!},$$

where the inequality uses the straightforward-to-derive bound $\binom{n}{k} \leq \frac{n^k}{k!}$ on binomial coefficients. Now the parameter $(M - 1)$ is constant (it depends only on the properties of the fixed finite set $\mathcal{F}$), so $|\mathcal{X}|$ is polynomial in n and thus polynomial in the instance size.

Note that $\mathcal{X} \subseteq \mathbb{Q}(A_{\mathcal{F}})$. Consider an element of $\mathcal{X}$ of the form

$$\prod_{j \in [m]} \prod_{\mathbf{x}_j \in \{0,1\}^{a_j}} (f_j(\mathbf{x}_j))^{k_{j,\mathbf{x}_j}}.$$

The condition on the sum of the integers $k_{j,\mathbf{x}_j}$, together with non-negativity, implies that at most n of these integers are non-zero. Thus, each element of $\mathcal{X}$ can be computed in time polynomial in n. Since the size of $\mathcal{X}$ is also polynomial in n, this means the elements of $\mathcal{X}$ can be enumerated in time polynomial in n.

Recall from Section 3.2 that, for any assignment $\xi : V \to \{0,1\}$, we have $\mathrm{wt}_{(V,C)}(\xi) = \prod_{c \in C} f_c(\xi|_c)$, where f_c is the function associated with the constraint c and $\xi|_c$ is the restriction of the assignment ξ to the scope of c. Now, for any $j \in [m]$ and $\mathbf{x}_j \in \{0,1\}^{a_j}$, define $\kappa_{j,\mathbf{x}_j} := |\{c \in C \text{ such that } f_c = f_j \text{ and } \xi|_c = \mathbf{x}_j\}|$, then

$$\mathrm{wt}_{(V,C)}(\xi) = \prod_{c \in C} f_c(\xi|_c) = \prod_{j \in [m]} \prod_{\mathbf{x}_j \in \{0,1\}^{a_j}} (f_j(\mathbf{x}_j))^{\kappa_{j,\mathbf{x}_j}}$$

and

$$\sum_{j \in [m]} \sum_{\mathbf{x}_j \in \{0,1\}^{a_j}} \kappa_{j,\mathbf{x}_j} = |C| = n.$$

Hence $\mathrm{wt}_{(V,C)}(\xi) \in \mathcal{X}$. This in turn implies that

$$Z_{\mathcal{F}}(V, C) = \sum_{x \in \mathcal{X}} x \cdot \#_{\mathcal{F}}((V, C), x). \tag{33}$$

Recall that $|\mathcal{X}|$ is polynomial in the instance size, and that the elements of $\mathcal{X}$ are generated by a straightforward procedure. Therefore, the elements of $\mathcal{X}$ can be enumerated in polynomial time. Multiplication and addition within $\mathbb{Q}(A)$ are also efficient, hence (33) gives the desired reduction $\#\mathsf{CSP}(\mathcal{F}) \leq_T \mathsf{COUNT}(\mathcal{F})$. $\qquad\square$

We are now ready to prove the planar version of Theorem 28.

Theorem 55. *Let $\mathcal{G}_1, \mathcal{G}_2 \subseteq \Upsilon$ be finite. Let $[y_0, y_1, y_2] \in \Upsilon_2$ be an ω-normalised and non-degenerate function. In the case $y_0 = y_2 = 0$, further assume that $\mathcal{G}_1$ contains a unary function $[a, b]$ which is ω-normalised and satisfies $ab \neq 0$. Then:*

$$\mathsf{Pl\text{-}Holant}\left(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \mid \{\mathrm{EQ}_3\} \cup \mathcal{G}_2\right) \equiv_T \mathsf{Pl\text{-}\#CSP}(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \cup \mathcal{G}_2).$$

Proof. First, consider the reduction from the holant problem to the counting CSP. We have

$$\mathsf{Pl\text{-}Holant}\left(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \mid \{\mathrm{EQ}_3\} \cup \mathcal{G}_2\right) \leq_T \mathsf{Pl\text{-}Holant}\left(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \cup \{\mathrm{EQ}_3\} \cup \mathcal{G}_2\right)$$
$$\leq_T \mathsf{Pl\text{-}\#CSP}(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \cup \mathcal{G}_2),$$

where the first step is by forgetting the bipartition and the second step is by Definition 23.

The reduction from the counting CSP to the holant problem is more complicated and separates into multiple cases.

Case 1: Assume $\mathsf{Pl\text{-}Holant}\left(\{[y_0, y_1, y_2]\} \mid \{\mathrm{EQ}_3\}\right)$ is #P-hard.

Then the more general counting problem $\mathsf{Pl\text{-}Holant}\left(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \mid \{\mathrm{EQ}_3\} \cup \mathcal{G}_2\right)$ is also #P-hard. The set $\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \cup \mathcal{G}_2$ is finite, hence $\mathsf{COUNT}(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \cup \mathcal{G}_2)$ is in #P. Therefore,

$$\mathsf{COUNT}(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \cup \mathcal{G}_2) \leq_T \mathsf{Pl\text{-}Holant}\left(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \mid \{\mathrm{EQ}_3\} \cup \mathcal{G}_2\right).$$

Furthermore, by Lemma 54 with $\mathcal{F} = \{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \cup \mathcal{G}_2$,

$$\mathsf{Pl\text{-}\#CSP}(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \cup \mathcal{G}_2) \leq_T \mathsf{COUNT}(\{[y_0, y_1, y_2]\} \cup \mathcal{G}_1 \cup \mathcal{G}_2).$$

Combining the two reductions yields the desired result.

Case 2: Assume $\mathsf{Pl\text{-}Holant}\left(\{[y_0, y_1, y_2]\} \mid \{\mathrm{EQ}_3\}\right)$ is not #P-hard. By Corollary 31, this implies at least one of the following properties holds:

1. $[y_0, y_1, y_2] \in \langle \mathcal{E} \rangle$, or

2. $[y_0, y_1, y_2] \in \mathcal{A}$, or

3. $y_0, y_1, y_2 \neq 0$ and $y_0^3 = y_2^3$.

We have dropped the holographic transformation from Subcase 2 because $[y_0, y_1, y_2]$ is required to be ω-normalised, which forces the holographic transformation to be trivial.

For Properties 1 and 2, the desired reduction follows from Lemmas 2–4 of [17] since all the gadget constructions in those proofs are planar.

For Property 3, note that the equation $y_0^3 = y_2^3$ implies that $y_2 = e^{2ik\pi/3}y_0$ for some $k \in \{0, 1, 2\}$. Since $[y_0, y_1, y_2]$ is ω-normalised, we must have $k = 0$ and thus $y_2 = y_0$. We will now prove the following chain of reductions, where $g := [y_0, y_1, y_0]$:

$$\mathsf{Pl\text{-}\#CSP}\left(\{g\} \cup \mathcal{G}_1 \cup \mathcal{G}_2\right) \leq_T \mathsf{Pl\text{-}Holant}\left(\{g, \mathrm{EQ}_3\} \cup \mathcal{G}_1 \cup \mathcal{G}_2\right)$$
$$\leq_T \mathsf{Pl\text{-}Holant}\left(\{g, \mathrm{EQ}_2\} \cup \mathcal{G}_1 \mid \{\mathrm{EQ}_2, \mathrm{EQ}_3\} \cup \mathcal{G}_2\right)$$
$$\leq_T \mathsf{Pl\text{-}Holant}\left(\{g, \mathrm{EQ}_2\} \cup \mathcal{G}_1 \mid \{\mathrm{EQ}_1, \mathrm{EQ}_2, \mathrm{EQ}_3\} \cup \mathcal{G}_2\right)$$
$$\leq_T \mathsf{Pl\text{-}Holant}\left(\{g\} \cup \mathcal{G}_1 \mid \{\mathrm{EQ}_1, \mathrm{EQ}_2, \mathrm{EQ}_3\} \cup \mathcal{G}_2\right)$$
$$\leq_T \mathsf{Pl\text{-}Holant}\left(\{g\} \cup \mathcal{G}_1 \mid \{\mathrm{EQ}_3\} \cup \mathcal{G}_2\right).$$

The first reduction is the definition of $\mathsf{Pl\text{-}\#CSP}$, the second step is by Proposition 5, and the third step is because adding an additional function on the RHS cannot make the problem easier. The fourth reduction step is by Lemma 53.

The first three reduction steps do not use any of the specific properties of g, and the fourth step only uses its property of being non-degenerate. It is only the fifth (and last) reduction step – which we will now prove – that uses the specific symmetry properties of g.

40

Figure 6: (a) A gadget for $y_0 \cdot \text{EQ}_1$ and (b) a gadget for $y_0(y_0 + y_1) \cdot \text{EQ}_2$, where each black degree-2 vertex is assigned $[y_0, y_1, y_0]$ and each white degree-3 vertex is assigned EQ_3.

Consider the gadgets in Figure 6, which can both be used on the RHS of the problem Pl-Holant $(\{g\} \cup \mathcal{G}_1 \mid \{\text{EQ}_3\} \cup \mathcal{G}_2)$. The first gadget has effective function $y_0 \cdot \text{EQ}_1$ and the second gadget has effective function $y_0(y_0 + y_1) \cdot \text{EQ}_2$. Recall that, $y_0 \neq 0$ by the assumption of the subcase and $y_0^2 = y_0 y_2 \neq y_1^2$ by non-degeneracy of g. The latter implies that $y_0 + y_1 \neq 0$.

We thus have non-zero scalings of EQ_1 and EQ_2 on the RHS. Therefore, by Lemmas 3 and 8,

$$\text{Pl-Holant}\left(\{g\} \cup \mathcal{G}_1 \mid \{\text{EQ}_1, \text{EQ}_2, \text{EQ}_3\} \cup \mathcal{G}_2\right) \leq_T \text{Pl-Holant}\left(\{g\} \cup \mathcal{G}_1 \mid \{\text{EQ}_3\} \cup \mathcal{G}_2\right).$$

This establishes the desired result. $\qquad\qquad\square$

5.5 Proof of the Holant$^+$ dichotomy theorem

Theorem 56. *Let $\mathcal{F} \subseteq \Upsilon$ be finite.* Holant$^+$ $(\mathcal{F})$ *can be computed in polynomial time if $\mathcal{F}$ satisfies one of the following conditions:*

- $\mathcal{F} \subseteq \langle \mathcal{T} \rangle$, *or*

- *there exists $O \in \mathcal{O}$ such that $\mathcal{F} \subseteq \langle O \circ \mathcal{E} \rangle$, or*

- $\mathcal{F} \subseteq \langle K \circ \mathcal{E} \rangle = \langle KX \circ \mathcal{E} \rangle$, *or*

- $\mathcal{F} \subseteq \langle K \circ \mathcal{M} \rangle$ *or* $\mathcal{F} \subseteq \langle KX \circ \mathcal{M} \rangle$, *or*

- $\mathcal{F} \subseteq \mathcal{A}$.

In all other cases, the problem is #P-hard. The same dichotomy holds for Pl-Holant$^+(\mathcal{F})$.

Proof. Define $\mathcal{F}' := \mathcal{F} \cup \{\delta_0, \delta_1, \delta_+, \delta_-\}$. The tractability part of the theorem follows by reduction to a conservative holant problem or to #CSP, respectively: if $\mathcal{F}$ is a subset of one of the tractable sets of Theorem 15, then $\mathcal{F}'$ is also a subset of one of the tractable sets of Theorem 15, and thus Holant$^+$ $(\mathcal{F})$ can be solved in polynomial time. Similarly, if $\mathcal{F} \subseteq \mathcal{A}$, then $\mathcal{F}' \subseteq \mathcal{A}$. Furthermore,

$$\text{Holant}^+ (\mathcal{F}) = \text{Holant}\left(\mathcal{F}'\right) \leq_T \text{Holant}\left(\mathcal{F}' \cup \{\text{EQ}_3\}\right) \leq_T \#\text{CSP}(\mathcal{F}'),$$

where the first reduction holds because adding a function cannot make the problem easier, and the second reduction is Proposition 16. Now, by Theorem 20, $\mathcal{F}' \subseteq \mathcal{A}$ implies that $\#\text{CSP}(\mathcal{F}')$ can be solved in polynomial time. Thus, by the above reduction, Holant$^+$ $(\mathcal{F})$ can be solved in polynomial time.

Hence from now on we may assume that we are not in one of the known tractable cases. We will then prove the hardness of Holant$^+$ $(\mathcal{F})$ via Theorem 28 (or Theorem 55 in the planar case), which requires ternary and binary symmetric non-decomposable functions.

Not being in one of the known tractable cases implies in particular that $\mathcal{F} \not\subseteq \langle \mathcal{T} \rangle$, i.e. there is some function $f \in \mathcal{F}$ having at least one factor which is a non-decomposable function of arity ≥ 3. Thus, we can apply Theorem 45 to realise a non-decomposable ternary function $f' \in S(\{f, \delta_0, \delta_1, \delta_+, \delta_-\})$ via a planar gadget. This function has either W or GHZ type, we distinguish cases accordingly.

1. Suppose f' has W type. There are several subcases.

- If $f' \notin (K \circ \mathcal{M}) \cup (KX \circ \mathcal{M})$, then there exists a non-decomposable symmetric ternary function $h \in S(\{f'\})$ by Lemma 48.

- If $f' \in K \circ \mathcal{M}$, since $\mathcal{F} \nsubseteq \langle K \circ \mathcal{M} \rangle$, there exists $g \in \mathcal{F} \setminus \langle K \circ \mathcal{M} \rangle$. We can realise a non-decomposable binary function $g' \in S(\{g, \delta_0, \delta_1, \delta_+, \delta_-\}) \setminus \langle K \circ \mathcal{M} \rangle$ via a planar gadget by Lemma 51. Then Lemma 49 can be applied, yielding a non-decomposable symmetric ternary function $h \in S(\{f', g'\})$.

- If $f' \in KX \circ \mathcal{M}$, the process is analogous to the subcase $f' \in K \circ \mathcal{M}$.

In each subcase, by Lemma 48 or Lemma 49, the gadget for h is planar, the non-decomposable symmetric ternary function h is in $S(\mathcal{F}')$, and h has GHZ type.

2. Suppose f' has GHZ type. Again, there are several subcases.

- If f' is already symmetric, let $h := f'$.
- If f' is not symmetric, we can realise a non-decomposable symmetric ternary function $f'' \in S(\{f', \delta_0, \delta_1, \delta_+, \delta_-\})$ by Lemma 47. The gadget for f'' is planar.
    - If f'' has GHZ type, let $h := f''$.
    - If f'' has W type, go back to Case 1 with f'' in place of f' and apply the symmetrisation procedure for W-type functions to get a symmetric GHZ-type function.

To summarise, if $\mathcal{F}$ is not one of the tractable sets, then there exists a non-decomposable symmetric ternary function $h \in S(\mathcal{F}')$ which can be realised via a planar gadget and which has GHZ type.

Recall from Section 3.4 that this means there exists $M \in \mathrm{GL}_2(\mathbb{A})$ such that $h = M \circ \mathrm{EQ}_3$ and either $M^T \circ \mathrm{EQ}_2$ is ω-normalised, or $M^T \circ \mathrm{EQ}_2 = c \cdot \mathrm{NEQ}$ for some $c \in \mathbb{A} \setminus \{0\}$. In the latter case, since $\left(\begin{smallmatrix} 1 & 0 \\ 0 & \lambda \end{smallmatrix}\right) \circ \mathrm{NEQ} = \lambda \cdot \mathrm{NEQ}$ for any λ, recall that we can choose M such that $M^T \circ \delta_0$ is ω-normalised. We may thus apply the following chain of interreductions:

$$
\begin{aligned}
\mathsf{Holant}^+ (\mathcal{F}) &= \mathsf{Holant} \left(\mathcal{F}' \right) \\
&\equiv_T \mathsf{Holant} \left(\mathcal{F}' \cup \{h\} \right) \\
&\equiv_T \mathsf{Holant} \left(\mathcal{F}' \cup \{\mathrm{EQ}_2\} \mid \{h, \mathrm{EQ}_2\} \right) \\
&\equiv_T \mathsf{Holant} \left(M^T \circ (\mathcal{F}' \cup \{\mathrm{EQ}_2\}) \mid \{\mathrm{EQ}_3, M^{-1} \circ \mathrm{EQ}_2\} \right) \\
&\equiv_T \#\mathsf{CSP} \left(M^T \circ (\mathcal{F}' \cup \{\mathrm{EQ}_2\}) \cup \{M^{-1} \circ \mathrm{EQ}_2\} \right)
\end{aligned}
$$

where the first step is the definition of Holant^+, the second step is by Lemma 8, the third step is by Proposition 5 with $\mathcal{G}_1 = \mathcal{F}'$ and $\mathcal{G}_2 = \{h\}$, the fourth step is by Theorem 6, and the last step is by Theorem 28. To prove $\mathsf{Holant}^+ (\mathcal{F})$ is hard, it therefore suffices to show that the counting CSP is hard whenever $\mathcal{F}$ is not one of the tractable families of Theorem 56.

We show the contrapositive: if the counting CSP is polynomial-time computable according to Theorem 20, then $\mathcal{F}$ is one of the tractable families of Theorem 56. The argument is split into cases according to the tractable cases of Theorem 20.

- Suppose $M^T \circ (\mathcal{F} \cup \{\mathrm{EQ}_2, \delta_0, \delta_1, \delta_+, \delta_-\}) \cup \{M^{-1} \circ \mathrm{EQ}_2\} \subseteq \mathcal{A}$.

    The condition $M^T \circ \{\mathrm{EQ}_2, \delta_0, \delta_1\} \subseteq \mathcal{A}$ is equivalent to $M \in \mathcal{B}$ by (8). The remaining conditions of the case are $M^T \circ \delta_+, M^T \circ \delta_-, M^{-1} \circ \mathrm{EQ}_2 \in \mathcal{A}$. Denote by f_L the binary function corresponding to a matrix $L \in \mathrm{GL}_2(\mathbb{A})$. By Lemma 40, $M \in \mathcal{B}$ and $M^T \circ \delta_+, M^T \circ \delta_- \in \mathcal{A}$ together imply that $M \in \mathcal{B}_{\mathcal{A}} := \{L \in \mathrm{GL}_2(\mathbb{A}) \mid f_L \in \mathcal{A}\}$, i.e. M is a matrix corresponding to a binary function in $\mathcal{A}$. Furthermore, by Lemma 39, $\mathcal{B}_{\mathcal{A}}$ is a group, so $M^{-1} \in \mathcal{B}_{\mathcal{A}}$. Now, by Lemma 38, $\mathcal{A}$ is closed under taking gadgets, so by Lemma 14, $M^{-1} \circ \mathrm{EQ}_2 \in \mathcal{A}$. Transposition of a matrix permutes the inputs of the corresponding function, so $M \in \mathcal{B}_{\mathcal{A}}$ also implies $M^T, (M^T)^{-1} \in \mathcal{B}_{\mathcal{A}}$. Thus, $M^T \circ \mathcal{F} \subseteq \mathcal{A}$ implies that $\mathcal{F} \subseteq (M^T)^{-1} \circ \mathcal{A} \subseteq \mathcal{A}$, one of the known tractable cases.

42

- Suppose $M^T \circ (\mathcal{F} \cup \{\text{EQ}_2, \delta_0, \delta_1, \delta_+, \delta_-\}) \cup \{M^{-1} \circ \text{EQ}_2\} \subseteq \langle \mathcal{E} \rangle$.

  Now, $M^T \circ \text{EQ}_2$ and $M^{-1} \circ \text{EQ}_2$ are non-decomposable symmetric binary functions with matrices $M^T M$ and $M^{-1}(M^{-1})^T$. All non-decomposable symmetric binary functions in $\langle \mathcal{E} \rangle$ take the form $[\lambda, 0, \mu]$ or $[0, \lambda, 0]$ for some $\lambda, \mu \in \mathbb{A} \setminus \{0\}$.

  - If $M^T M = \left(\begin{smallmatrix} \lambda & 0 \\ 0 & \mu \end{smallmatrix} \right)$, then $M = QD$ for some $Q \in \mathcal{O}$ and some invertible diagonal matrix D by Lemma 12. Now $(QD)^T \circ \mathcal{F} \subseteq \langle \mathcal{E} \rangle$ implies $\mathcal{F} \subseteq \left\langle (QD^{-1}) \circ \mathcal{E} \right\rangle = \langle Q \circ \mathcal{E} \rangle$, which is one of the known tractable families.

  - Similarly, if $M^T M = \lambda X$, then $M = KD$ or $M = KXD$ for some invertible diagonal matrix D by Lemma 13. Now $K^T \doteq XK^{-1}$, so $(KD)^T \circ \mathcal{F} \subseteq \langle \mathcal{E} \rangle$ implies $\mathcal{F} \subseteq \left\langle (KXD^{-1}) \circ \mathcal{E} \right\rangle = \langle K \circ \mathcal{E} \rangle$, which is another of the known tractable families.

    An analogous argument holds for KX instead of K.

We have shown that if $\#\mathsf{CSP} \left(M^T \circ (\mathcal{F}' \cup \{\text{EQ}_2\}) \cup \{M^{-1} \circ \text{EQ}_2\} \right)$ can be solved in polynomial time, this implies that $\mathcal{F}$ is one of the tractable families listed in the Theorem 56. By Theorem 20, the counting CSP is $\#$P-hard in all other cases. Thus, if $\mathcal{F}$ is not one of the tractable families listed in the theorem, then $\mathsf{Holant}^+ (\mathcal{F})$ is $\#$P-hard. This completes the proof of the theorem for the non-planar case.

As all gadgets used in this proof are planar, the above constructions also work in the planar case. The only difference is that, when considering planar holant problems, we need to use Theorem 55 instead of Theorem 28 and apply the planar $\#\mathsf{CSP}$ dichotomy from Theorem 25 instead of Theorem 20. In addition to the tractable cases from the general $\#\mathsf{CSP}$ dichotomy, which we have already excluded, Theorem 25 contains one additional tractable family: $\mathsf{Pl}\text{-}\#\mathsf{CSP} \left(M^T \circ (\mathcal{F}' \cup \{\text{EQ}_2\}) \cup \{M^{-1} \circ \text{EQ}_2\} \right)$ can be solved in polynomial time if $M^T \circ (\mathcal{F}' \cup \{\text{EQ}_2\}) \cup \{M^{-1} \circ \text{EQ}_2\} \subseteq \left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix} \right) \circ \mathcal{H}$, where $\mathcal{H}$ is the set of matchgate functions. All other cases remain $\#$P-hard. By Lemma 24, the only unary matchgate functions are scalings of the pinning functions δ_0 and δ_1, so the only unary functions in $\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix} \right) \circ \mathcal{H}$ are $\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix} \right) \circ \delta_0 \doteq \delta_+$ and $\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix} \right) \circ \delta_1 \doteq \delta_-$ (up to scaling). Yet $M^T \circ \mathcal{F}'$ contains at least the four unary functions $M^T \circ \{\delta_0, \delta_1, \delta_+, \delta_-\}$ and, by invertibility of M, these four functions have to be mapped to four pairwise linearly-independent functions. Thus, $M^T \circ (\mathcal{F}' \cup \{\text{EQ}_2\}) \cup \{M^{-1} \circ \text{EQ}_2\}$ cannot be a subset of $\left(\begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix} \right) \circ \mathcal{H}$. Therefore, the Holant^+ dichotomy remains unchanged when restricted to planar instances. $\qquad\square$

6   The full Holant^c dichotomy

We now combine the techniques developed in deriving the Holant^+ dichotomy with techniques from the real-valued Holant^c dichotomy [19] to get a full complexity classification for complex-valued Holant^c.

As in the Holant^+ case, the general proof strategy is to realise a non-decomposable ternary function and then a non-decomposable symmetric ternary function. Without δ_+ and δ_-, we can no longer use Theorem 45. Instead, we employ a technique using δ_0, δ_1 and self-loops from the proof of Theorem 5.1 in [19] with some slight modifications. Self-loops reduce the arity of a function in steps of 2, so sometimes this technique fails to yield a non-decomposable ternary function. When not yielding a ternary function, the process instead yields a non-decomposable arity-4 function with specific properties. The complexity classification for the latter case was resolved in [19] even for complex values.

The symmetrisation constructions for binary and ternary functions, as well as the subsequent hardness proofs, occasionally require a little extra work in the Holant^c setting as compared to the Holant^+ setting; we deal with those issues before proving the main theorem.

6.1 Hardness proofs involving a non-decomposable ternary function

First, we prove several lemmas that give a complexity classification for Holant^c problems in the presence of a non-decomposable ternary function. These results adapt techniques used in the Holant^+ complexity classification to the Holant^c setting. They also replace Lemmas 5.1, 5.3, and 5.5–5.7 of [19]. Whereas the last three of those only apply to real-valued functions, our new results work for complex values.

Lemma 57. *Suppose $\mathcal{F} \subseteq \Upsilon$ is finite and $f \in \mathcal{F}$ is a non-decomposable ternary function. If $f \in K \circ \mathcal{M}$ and $\mathcal{F} \nsubseteq \langle K \circ \mathcal{M} \rangle$, then $\mathsf{Holant}^c(\mathcal{F})$ is #P-hard. Similarly, if $f \in KX \circ \mathcal{M}$ and $\mathcal{F} \nsubseteq \langle KX \circ \mathcal{M} \rangle$, then $\mathsf{Holant}^c(\mathcal{F})$ is #P-hard.*

Proof. We consider the case $f \in K \circ \mathcal{M}$ and $\mathcal{F} \nsubseteq \langle K \circ \mathcal{M} \rangle$, the proof for the second case is analogous since $X \circ \mathcal{M}$ differs from $\mathcal{M}$ only by a bit flip on all function inputs.

As $\mathcal{F} \nsubseteq \langle K \circ \mathcal{M} \rangle$, we can find $h \in \mathcal{F} \setminus \langle K \circ \mathcal{M} \rangle$. Then h has arity at least 2 and is non-degenerate because all unary functions are in $K \circ \mathcal{M}$. We distinguish cases according to the arity and decomposability properties of h.

Case 1: Suppose h has arity 2, then non-degeneracy implies h is non-decomposable. Thus, by Lemma 49, there exists a non-decomposable symmetric ternary function $g \in S(\{f, h\})$. This function g is guaranteed to have GHZ type by the same lemma. We thus have

$$\mathsf{Holant}^c(\mathcal{F}) = \mathsf{Holant}(\mathcal{F} \cup \{\delta_0, \delta_1\}) \equiv_T \mathsf{Holant}(\mathcal{F} \cup \{\delta_0, \delta_1, g\})$$
$$\equiv_T \mathsf{Holant}(\mathcal{F} \cup \{\delta_0, \delta_1, \mathrm{EQ}_2\} \mid \{g, \mathrm{EQ}_2\})$$

where the equality is the definition of Holant^c, the first reduction is by Lemma 8, and the second reduction is by Proposition 5. Furthermore, there exists $M \in \mathrm{GL}_2(\mathbb{A})$ such that $g = M \circ \mathrm{EQ}_3$ and either $M^T \circ \mathrm{EQ}_2$ is ω-normalised or $M^T \circ \mathrm{EQ}_2 = c \cdot \mathrm{NEQ}$ for some $c \in \mathbb{A} \setminus \{0\}$. In the latter case, we can choose M such that $M^T \circ \delta_0$ is ω-normalised. Therefore,

$$\mathsf{Holant}^c(\mathcal{F}) \equiv_T \mathsf{Holant}(\mathcal{F} \cup \{\delta_0, \delta_1, \mathrm{EQ}_2\} \mid \{g, \mathrm{EQ}_2\})$$
$$\equiv_T \mathsf{Holant}\left(M^T \circ (\mathcal{F} \cup \{\delta_0, \delta_1, \mathrm{EQ}_2\}) \mid \{\mathrm{EQ}_3, M^{-1} \circ \mathrm{EQ}_2\}\right)$$
$$\equiv_T \#\mathsf{CSP}\left(M^T \circ (\mathcal{F} \cup \{\delta_0, \delta_1, \mathrm{EQ}_2\}) \cup \{M^{-1} \circ \mathrm{EQ}_2\}\right)$$

where the first step is from above, the second step is Theorem 6, and the third step is by Theorem 28. Now, $f \in \mathcal{F} \cap K \circ \mathcal{M}$ is a non-decomposable ternary function. By Lemma 41, any non-decomposable ternary function in $\mathcal{M}$ has W type, and holographic transformations by definition do not affect the entanglement class. Thus, $M^T \circ f$ has W type. But Lemma 41 also shows that any non-decomposable ternary function in $\langle \mathcal{E} \rangle$ or in $\mathcal{A}$ has GHZ type. Therefore $M^T \circ f \notin \langle \mathcal{E} \rangle$ and $M^T \circ f \notin \mathcal{A}$, hence the counting CSP is #P-hard.

Case 2: Suppose h is an n-ary function with $n > 2$, and h is non-decomposable.

Write the ternary function f as $K \circ f'$, where $f' \in \mathcal{M}$ means that it takes the form $f' = (a, b, c, 0, d, 0, 0, 0)$ for some $a, b, c, d \in \mathbb{A}$. Non-decomposability of f implies $bcd \neq 0$. Consider the three different gadgets that consist of a vertex assigned function f with a self-loop (where the three gadgets differ in which argument of f corresponds to the dangling edge). The gadget where the first edge is dangling has effective function

$$\sum_{y \in \{0,1\}} f(x, y, y) = \sum_{y, z_1, z_2, z_3 \in \{0,1\}} K_{x z_1} K_{y z_2} K_{y z_3} f'(z_1, z_2, z_3)$$
$$= 2 \sum_{z_1, z_2, z_3 \in \{0,1\}} K_{x z_1} \mathrm{NEQ}(z_2, z_3) f'(z_1, z_2, z_3)$$
$$= 2 \sum_{z_1 \in \{0,1\}} K_{x z_1} \left(f'(z_1, 0, 1) + f'(z_1, 1, 0)\right)$$
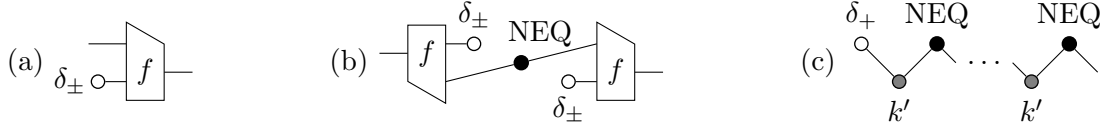$$= 2(b + c)(K \circ \delta_0)(x)$$

Figure 7: (a) The gadget for k, where we denote f by a non-symmetric box to indicate that it is not generally a symmetric function. (b) The gadget for k', which is symmetric. (c) A family of gadgets for producing unary functions.

Using a self loop on a vertex assigned function f, we can therefore realise the unary function $2(b + c)(K \circ \delta_0) = 2(b + c)[1, i]$, where i is the imaginary unit. The other two gadgets similarly yield $2(b + d)[1, i]$ and $2(c + d)[1, i]$. Since $bcd \neq 0$, at least one of those gadgets is non-zero. Thus we can realise $\delta_i = [1, i]$ up to irrelevant scaling.

We can now prove the following chain of interreductions:

$$\begin{aligned}
\mathsf{Holant}^c\,(\mathcal{F}) &= \mathsf{Holant}\,(\mathcal{F} \cup \{\delta_0, \delta_1\}) \\
&\equiv_T \mathsf{Holant}\,(\mathcal{F} \cup \{\delta_0, \delta_1, \delta_i, f, h\}) \\
&\equiv_T \mathsf{Holant}\,(\mathcal{F} \cup \{f, h, \mathrm{EQ}_2\} \mid \{\delta_0, \delta_1, \delta_i, \mathrm{EQ}_2\}) \\
&\equiv_T \mathsf{Holant}\,\left(K^{-1} \circ (\mathcal{F} \cup \{f, h, \mathrm{EQ}_2\}) \mid K^T \circ \{\delta_0, \delta_1, \delta_i, \mathrm{EQ}_2\}\right)
\end{aligned}$$

Here, the equality is the definition of Holant^c. The first reduction is by Lemma 8 and the above gadget constructions. The second reduction is by Proposition 5 and the third reduction is by Theorem 6.

Recall that $K = \left(\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right)$, so $K^T \circ \delta_0 \doteq \delta_+$, $K^T \circ \delta_1 \doteq \delta_-$, $K^T \circ \delta_i \doteq \delta_1$ and $K^T \circ \mathrm{EQ}_2 \doteq \mathrm{NEQ}$. Let $h' := K^{-1} \circ h$ and recall from above that $f' = K^{-1} \circ f = (a, b, c, 0, d, 0, 0, 0)$. Thus, the effective function of the gadget in Figure 7a is $k(x, y) = \sum_{z \in \{0,1\}} f'(x, y, z)\delta_\pm(z)$. This is a LHS gadget and $k = (a \pm b, c, d, 0)$. Since $b \neq 0$, there is a choice of sign such that $k(0, 0) \neq 0$.

Then the gadget in Figure 7b has effective function

$$k'(x, y) = \sum_{z_1, z_2 \in \{0,1\}} k(x, z_1)\mathrm{NEQ}(z_1, z_2)k(y, z_2).$$

It is a symmetric LHS gadget, and $k' = (2d(a \pm b), cd, cd, 0) \doteq [\frac{2}{c}(a \pm b), 1, 0]$. Note that with the above choice of sign, $z := k'(0, 0) \neq 0$.

A chain of ℓ of these symmetric gadgets, connected to $\delta_\pm$ at one end and connected by copies of NEQ, as shown in Figure 7c, gives a RHS gadget with function $[1, \ell z \pm 1]$. Thus, since $z \neq 0$, we can realise polynomially many different unary functions on the RHS.

Since $h' = K^{-1} \circ h \notin \mathcal{M}$, there exists a bit string $\mathbf{a} \in \{0, 1\}^n$ of Hamming weight at least 2 such that $h(\mathbf{a}) \neq 0$. Without loss of generality, assume $a_1 = a_2 = 1$. Otherwise permute the argument of h, the resulting function is in $S(\{h\})$ so it can be added to the LHS of our holant problem without affecting the complexity. Let $v_m = \delta_{a_m}$ for $m \in \{3, 4, \ldots, n\}$ and define

$$g_1(x_1, x_2) := \sum_{x_3, \ldots, x_n \in \{0,1\}} h'(x_1, \ldots, x_n) \prod_{m=3}^{n} v_m(x_m).$$

Then

$$g_1(1, 1) \neq 0. \tag{34}$$

Furthermore, by Proposition 44, we know that there exist $u_m \in \{\delta_0, \delta_1, \delta_+, \delta_-\}$ for all $m \in \{3, 4, \ldots, n\}$ such that the following function is non-decomposable:

$$g_2(x_1, x_2) := \sum_{x_3, \ldots, x_n \in \{0,1\}} h'(x_1, \ldots, x_n) \prod_{m=3}^{n} u_m(x_m).$$

45

The non-decomposability condition for binary functions is

$$g_2(0,0)g_2(1,1) - g_2(0,1)g_2(1,0) \neq 0. \tag{35}$$

Now consider a third function

$$g_3(x_1, x_2) := \sum_{x_3, \ldots, x_n \in \{0,1\}} h'(x_1, \ldots, x_n) \prod_{m=3}^{n} w_m(x_m)$$

where $w_m = [1, 1 + \ell_m z]$ for each $m \in \{3, 4, \ldots, n\}$, with the ℓ_m being integer variables whose values are yet to be determined. Define

$$p(\ell_3,, \ldots, \ell_m) := g_3(1,1)\big(g_3(0,0)g_3(1,1) - g_3(0,1)g_3(1,0)\big).$$

Then p is a multivariate polynomial where the maximum exponent of any variable is 3. By (34) and (35), this polynomial is not identically zero.

Now since the variable ℓ_3 has degree at most 3 in p, there exists a value $\lambda_3 \in \{0, 1, 2, 3\}$ such that $p(\lambda_3, \ell_4, \ldots, \ell_m)$ is not identically zero. We may repeat this argument for $\ell_4, \ldots, \ell_m$ until we have found values $\lambda_4, \ldots, \lambda_m$ for all the variables such that $p(\lambda_3, \ldots, \lambda_m) \neq 0$. Each resulting w_m is realisable by a RHS gadget. Thus, g_3 is realisable by a LHS gadget. The function g_3 is binary, non-decomposable, and not in $K \circ \mathcal{M}$. Therefore we can proceed as in the case where h is binary.

Case 3: Suppose h is an n-ary function with $n > 2$, and h is decomposable. Since $h \notin \langle K \circ \mathcal{M} \rangle$, in any decomposition of h there must be one factor h' which is not in $\langle K \circ \mathcal{M} \rangle$. This factor has arity at least 2 since all unary functions are in $\langle K \circ \mathcal{M} \rangle$, and its arity is strictly smaller than that of h. Furthermore, by Lemma 9, $h' \in S(\{h, \delta_0, \delta_1\})$. We may thus apply the argument to h' instead of h.

If h satisfies the conditions of Case 1, we are immediately done. Case 2 straightforwardly reduces to Case 1. Finally, Case 3 yields a function of smaller arity than the original one, to which the case distinction can then be applied. Because the arity decreases every time we hit Case 3, the argument terminates. Thus the proof is complete. $\qquad\square$

Lemma 58. *Suppose $\mathcal{F} \subseteq \Upsilon$ is finite and contains a non-decomposable ternary function f. Then $\mathsf{Holant}^c(\mathcal{F})$ is #P-hard unless:*

1. *There exists $O \in \mathcal{O}$ such that $\mathcal{F} \subseteq \langle O \circ \mathcal{E} \rangle$, or*

2. *$\mathcal{F} \subseteq \langle K \circ \mathcal{E} \rangle = \langle KX \circ \mathcal{E} \rangle$, or*

3. *$\mathcal{F} \subseteq \langle K \circ \mathcal{M} \rangle$ or $\mathcal{F} \subseteq \langle KX \circ \mathcal{M} \rangle$, or*

4. *$\mathcal{F} \subseteq M \circ \mathcal{A}$ for some $M \in \mathcal{B}$, as defined in (8).*

In all the exceptional cases, the problem $\mathsf{Holant}^c(\mathcal{F})$ can be solved in polynomial time.

Remark. The case $\mathcal{F} \subseteq \langle \mathcal{T} \rangle$ does not appear here because f is a non-decomposable ternary function and $f \in \mathcal{F}$, which implies $\mathcal{F} \not\subseteq \langle \mathcal{T} \rangle$.

Proof. We distinguish two cases according to whether the ternary function f is symmetric.
Case 1: Suppose f is symmetric. By the entanglement classification (cf. Section 4.1), f has either GHZ type or W type. We treat these two subcases separately.
Subcase a: Suppose f is of GHZ type. Given an instance of $\mathsf{Holant}^c(\mathcal{F})$, by Proposition 4,

$$\mathsf{Holant}^c(\mathcal{F}) = \mathsf{Holant}(\mathcal{F} \cup \{\delta_0, \delta_1\}) \equiv_T \mathsf{Holant}(\mathcal{F} \cup \{\delta_0, \delta_1\} \mid \{\mathrm{EQ}_2\}).$$

Furthermore, we have

$$\mathsf{Holant}\left(\mathcal{F} \cup \{\delta_0, \delta_1\} \mid \{\mathrm{EQ}_2\}\right) \equiv_T \mathsf{Holant}\left(\mathcal{F} \cup \{\delta_0, \delta_1\} \mid \{\mathrm{EQ}_2, \delta_0, \delta_1\}\right).$$

The $\leq_T$ direction is immediate; for the other direction, note that any occurrence of δ_0 or δ_1 on the RHS can be replaced by a gadget consisting of a LHS copy of the unary function connected to EQ_2.

By Theorem 6, for any $M \in \mathrm{GL}_2(\mathbb{A})$,

$$\mathsf{Holant}\left(\mathcal{F} \cup \{\delta_0, \delta_1\} \mid \{\mathrm{EQ}_2, \delta_0, \delta_1\}\right)$$
$$\equiv_T \mathsf{Holant}\left(M^{-1} \circ (\mathcal{F} \cup \{\delta_0, \delta_1\}) \,\middle|\, M^T \circ \{\mathrm{EQ}_2, \delta_0, \delta_1\}\right). \quad (36)$$

In the following, we aim to identify a matrix M such that Theorem 28 can be applied to show that the RHS of (36) is equivalent to a counting CSP. To apply the theorem, the following three properties must hold for M and $\mathcal{F}$:

- $\mathrm{EQ}_3 \in M^{-1} \circ \mathcal{F}$,

- $M^T \circ \mathrm{EQ}_2$ is ω-normalised, and

- if $M^T \circ \mathrm{EQ}_2 = [0, \lambda, 0]$ for some $\lambda \in \mathbb{A} \setminus \{0\}$, then there must exist an ω-normalised function in $M^T \circ \{\delta_0, \delta_1\}$ which is not a pinning function

We now show that it is always possible to choose M so these conditions are satisfied.

As f has GHZ type, there exists $A \in \mathrm{GL}_2(\mathbb{A})$ such that $f = A \circ \mathrm{EQ}_3$. The matrix A is not unique (cf. Section 3.4); for now we pick an arbitrary one among all matrices that satisfy $f = A \circ \mathrm{EQ}_3$.

- Suppose $A^T \circ \mathrm{EQ}_2 \neq \lambda \cdot \mathrm{NEQ}$ for any $\lambda \in \mathbb{A}$. If $A^T \circ \mathrm{EQ}_2$ is ω-normalised, let $M := A$. Otherwise, by the argument in Section 6.1, there exists $D_\omega := \left(\begin{smallmatrix} 1 & 0 \\ 0 & \omega \end{smallmatrix}\right)$ with $\omega \in \{e^{2i\pi/3}, e^{4i\pi/3}\}$ such that $(D_\omega A^T) \circ \mathrm{EQ}_2$ is ω-normalised. Let $M := AD_\omega$, then $f = M \circ \mathrm{EQ}_3$. Now, since $\mathrm{EQ}_3 \in M^{-1} \circ \mathcal{F}$ and $M^T \circ \mathrm{EQ}_2$ is an ω-normalised symmetric binary function, Theorem 28 can be applied.

- Suppose $A^T \circ \mathrm{EQ}_2 = \lambda \cdot \mathrm{NEQ}$ for some $\lambda \in \mathbb{A}$, then $A^T A \doteq X$. Thus, by Lemma 13, $A = KD$ or $A = KXD$ for some invertible diagonal matrix D. In either of these cases, all entries of A are non-zero, so $A^T \circ \delta_0 = [a, b]$ for some $a, b, \in \mathbb{A} \setminus \{0\}$. Thus there exists $M := AD_\omega$ for some $\omega^3 = 1$ such that $f = M \circ \mathrm{EQ}_3$ and $M^T \circ \delta_0$ is ω-normalised, i.e. the conditions of Theorem 28 are satisfied.

In either case, Theorem 28 yields

$$\mathsf{Holant}^c\left(\mathcal{F}\right) \equiv_T \#\mathsf{CSP}\left(M^{-1} \circ (\mathcal{F} \cup \{\delta_0, \delta_1\}) \cup M^T \circ \{\mathrm{EQ}_2, \delta_0, \delta_1\}\right), \quad (37)$$

where $\mathrm{EQ}_3 \in M^{-1} \circ \mathcal{F}$. The matrix M may still not be uniquely defined, but the remaining ambiguity does not affect any of the subsequent arguments.

By (37) and Theorem 20, $\mathsf{Holant}^c\left(\mathcal{F}\right)$ is #P-hard unless

$$\mathcal{F}' := M^{-1} \circ (\mathcal{F} \cup \{\delta_0, \delta_1\}) \cup M^T \circ \{\mathrm{EQ}_2, \delta_0, \delta_1\}$$

is a subset of either $\langle \mathcal{E} \rangle$ or $\mathcal{A}$. Again, we treat these two cases separately.

- Suppose $\mathcal{F}' \subseteq \langle \mathcal{E} \rangle$. All non-decomposable binary functions in $\langle \mathcal{E} \rangle$ are of the form

$$\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix} \qquad \text{or} \qquad \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}$$

for some $\alpha, \beta \in \mathbb{A} \setminus \{0\}$. Note that $M^T \circ \mathrm{EQ}_2$ corresponds to the matrix $M^T M$. As $M^T \circ \mathrm{EQ}_2 \in \mathcal{F}'$, there are two subcases.

– Suppose $M^T M = \left(\begin{smallmatrix} \alpha & 0 \\ 0 & \beta \end{smallmatrix} \right)$. Then by Lemma 12, $M = QD$ for some orthogonal 2 by 2 matrix Q and some invertible diagonal matrix D. Now, $\mathcal{F}' \subseteq \langle \mathcal{E} \rangle$ implies $M^{-1} \circ \mathcal{F} \subseteq \langle \mathcal{E} \rangle$, which is equivalent to $\mathcal{F} \subseteq \langle M \circ \mathcal{E} \rangle$ since holographic transformations and closure under tensor products commute. But $D \circ \mathcal{E} = \mathcal{E}$ for any invertible diagonal matrix D. Hence $\mathcal{F} \subseteq \langle Q \circ \mathcal{E} \rangle$, where Q is orthogonal, i.e. $\mathcal{F}$ satisfies Item 1 of this lemma.

– Suppose $M^T M = \left(\begin{smallmatrix} 0 & \alpha \\ \beta & 0 \end{smallmatrix} \right)$ Then by Lemma 13, $M = KD$ or $M = KXD$ for some invertible diagonal matrix D. As in the previous case, $\mathcal{F}' \subseteq \langle \mathcal{E} \rangle$ implies $\mathcal{F} \subseteq \langle M \circ \mathcal{E} \rangle$. Since $\mathcal{E}$ is invariant under holographic transformations by diagonal matrices and under bit flips, this is the same as $\mathcal{F} \subseteq \langle K \circ \mathcal{E} \rangle$. Hence $\mathcal{F}$ satisfies Item 2 of this lemma.

This completes the analysis of the subcase $\mathcal{F}' \subseteq \langle \mathcal{E} \rangle$.

- Suppose $\mathcal{F}' \subseteq \mathcal{A}$. This implies in particular

$$M^T \circ \{\mathrm{EQ}_2, \delta_0, \delta_1\} \subseteq \mathcal{A}, \tag{38}$$

which is the definition of $M \in \mathcal{B}$, cf. (8).

The assumption $\mathcal{F}' \in \mathcal{A}$ also implies $M^{-1} \circ \{\delta_0, \delta_1\} \subseteq \mathcal{A}$. This does not yield any further restrictions on $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right)$. To see this, note we already know that $M^T \circ \{\delta_0, \delta_1\} \subseteq \mathcal{A}$, i.e. $[a, b], [c, d] \in \mathcal{A}$. Now, up to an irrelevant scalar factor, $M^{-1} \circ \{\delta_0, \delta_1\}$ is $\{[d, -c], [-b, a]\}$. Note that $[-b, a] = -\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right) \circ [a, b]$ and $[d, -c] = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right) \circ [c, d]$. The function corresponding to $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix} \right)$ is affine and $\mathcal{A}$ is closed under taking gadgets (cf. Lemma 38) and under scalings. Hence, by Lemma 14, $[a, b], [c, d] \in \mathcal{A}$ implies $[d, -c], [-b, a] \in \mathcal{A}$.

Thus $\mathcal{F}' \subseteq \mathcal{A}$ implies $M \in \mathcal{B}$. Furthermore, $\mathcal{F}' \subseteq \mathcal{A}$ implies $M^{-1} \circ \mathcal{F} \subseteq \mathcal{A}$, or equivalently $\mathcal{F} \subseteq M \circ \mathcal{A}$, i.e. $\mathcal{F}$ satisfies Item 4 of this lemma.

To summarise, if f has GHZ type, the problem is tractable if $\mathcal{F} \subseteq \langle Q \circ \mathcal{E} \rangle$ for some orthogonal 2 by 2 matrix Q, if $\mathcal{F} \subseteq \langle K \circ \mathcal{E} \rangle$, or if there exists $M \in \mathcal{B}$ such that $\mathcal{F} \subseteq M \circ \mathcal{A}$. In all other cases, the problem is #P-hard by reduction from #CSP.

Subcase b: Suppose f is of W type, then:

- If $f \notin (K \circ \mathcal{M}) \cup (KX \circ \mathcal{M})$, Holant (f) is #P-hard by Theorem 29.

- If $\mathcal{F} \subseteq \langle K \circ \mathcal{M} \rangle$ or $\mathcal{F} \subseteq \langle KX \circ \mathcal{M} \rangle$, the problem is polynomial-time computable by Theorem 15; this is Item 3 of this lemma.

- If $f \in K \circ \mathcal{M}$ but $\mathcal{F} \nsubseteq \langle K \circ \mathcal{M} \rangle$, the problem is #P-hard by Lemma 57, and analogously with KX instead of K.

Case 2: Suppose f is not symmetric.

- If $f \notin (K \circ \mathcal{M}) \cup (KX \circ \mathcal{M})$, we can realise a non-decomposable symmetric ternary function by Lemmas 47 and 48 and then proceed as in Case 1.

- If $\mathcal{F} \subseteq \langle K \circ \mathcal{M} \rangle$ or $\mathcal{F} \subseteq \langle KX \circ \mathcal{M} \rangle$, the problem is polynomial-time computable by Theorem 15; this is Item 3 of the lemma.

- Finally, if $f \in K \circ \mathcal{M}$ but $\mathcal{F} \nsubseteq \langle K \circ \mathcal{M} \rangle$, or $f \in KX \circ \mathcal{M}$ but $\mathcal{F} \nsubseteq \langle KX \circ \mathcal{M} \rangle$, the problem is #P-hard by Lemma 57.

This covers all cases. □

6.2   Main theorem

We now have all the components required to prove the main dichotomy for Holant^c. The following theorem generalises Theorem 5.1 of [19], which applies only to real-valued functions. The proof follows the one in [19, Theorem 5.1] fairly closely.

Theorem 59. *Let* $\mathcal{F} \subseteq \Upsilon$ *be finite. Then* $\mathsf{Holant}^c(\mathcal{F})$ *is* #P-hard *unless:*

- $\mathcal{F} \subseteq \langle \mathcal{T} \rangle$, *or*

- *there exists* $O \in \mathcal{O}$ *such that* $\mathcal{F} \subseteq \langle O \circ \mathcal{E} \rangle$, *or*

- $\mathcal{F} \subseteq \langle K \circ \mathcal{E} \rangle = \langle KX \circ \mathcal{E} \rangle$, *or*

- $\mathcal{F} \subseteq \langle K \circ \mathcal{M} \rangle$ *or* $\mathcal{F} \subseteq \langle KX \circ \mathcal{M} \rangle$, *or*

- *there exists* $B \in \mathcal{B}$ *such that* $\mathcal{F} \subseteq B \circ \mathcal{A}$, *or*

- $\mathcal{F} \subseteq \mathcal{L}$.

In all of the exceptional cases, $\mathsf{Holant}^c(\mathcal{F})$ *is polynomial-time computable.*

Proof. If $\mathcal{F}$ is in one of the first four exceptional cases, then polynomial-time computability follows from Theorem 15. In the penultimate exceptional case, let $\mathcal{F}' := B^{-1} \circ \mathcal{F} \subseteq \mathcal{A}$, then

$$
\begin{aligned}
\mathsf{Holant}^c(\mathcal{F}) &\leq_T \mathsf{Holant}^c(\mathcal{F} \mid \{\mathrm{EQ}_2\}) \\
&\leq_T \mathsf{Holant}^c\left(\mathcal{F}' \mid \{B^T \circ \mathrm{EQ}_2\}\right) \\
&\leq_T \mathsf{Holant}^c\left(\mathcal{F}' \cup \{B^T \circ \mathrm{EQ}_2\}\right) \\
&\leq_T \mathsf{Holant}^c\left(\mathcal{F}' \cup \{B^T \circ \mathrm{EQ}_2, \mathrm{EQ}_3\}\right) \\
&\leq_T \#\mathsf{CSP}(\mathcal{F}' \cup \{B^T \circ \mathrm{EQ}_2\})
\end{aligned}
$$

where the first reduction is Proposition 4 and the second reduction is Theorem 6 with $M = B^{-1}$. The third and fourth reductions hold because dropping the bipartite restriction and adding a function cannot make the problem easier. The final reduction combines Proposition 16 and Lemma 21. Now, $B^T \circ \mathrm{EQ}_2 \in \mathcal{A}$ by (8). A counting CSP with affine constraint functions is polynomial-time computable by Theorem 20.

In the final exceptional case, we have

$$
\mathsf{Holant}^c(\mathcal{F}) \leq_T \mathsf{Holant}^c(\mathcal{F} \cup \{\mathrm{EQ}_4\}) \leq_T \#\mathsf{CSP}_2^c(\mathcal{F})
$$

where the final reduction can be proved analogously to Proposition 16. Polynomial-time computability thus follows from Theorem 22.

From now on, assume $\mathcal{F}$ does not satisfy any of the tractability conditions. In particular, this implies that $\mathcal{F} \not\subseteq \langle \mathcal{T} \rangle$, i.e. $\mathcal{F}$ has multipartite entanglement.

By Lemma 9, without loss of generality, we may focus on non-decomposable functions. So assume that there is some non-decomposable function $f \in S(\mathcal{F} \cup \{\delta_0, \delta_1\})$ of arity $n \geq 3$. If the function has arity 3, we are done by Lemma 58. Hence assume $n \geq 4$.

Now if there was a bit string $\mathbf{a} \in \{0,1\}^n$ such that $f(\mathbf{x}) = 0$ for all $\mathbf{x} \in \{0,1\}^n \setminus \{\mathbf{a}\}$, then f would be a scaling of some function in $\langle \{\delta_0, \delta_1\} \rangle$, so it would be degenerate and thus decomposable. Hence, f being non-decomposable implies that there exist two distinct n-bit strings $\mathbf{a}, \mathbf{b} \in \{0,1\}^n$ such that $f(\mathbf{a})f(\mathbf{b}) \neq 0$.

As in the proof of [19, Theorem 5.1], let:

$$
D_0 = \min\{d(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \neq \mathbf{y}, f(\mathbf{x}) \neq 0, f(\mathbf{y}) \neq 0\}, \tag{39}
$$

49

where $d(\cdot, \cdot)$ is the Hamming distance. By the above argument, D_0 exists. We now distinguish cases according to the values of D_0.

Case $D_0 \geq 4$ and D_0 is even: Pick a pair of bit strings $\mathbf{a}, \mathbf{b}$ such that $f(\mathbf{a})f(\mathbf{b}) \neq 0$ with minimum Hamming distance $d(\mathbf{a}, \mathbf{b})$. Pin all inputs where the two bit strings agree (without loss of generality, we always assume bit strings agree on the last $n - D_0$ bits; otherwise permute the arguments). This realises a function $f' \in \Upsilon_{D_0}$ of the form

$$f'(\mathbf{x}) = \begin{cases} \alpha & \text{if } x_j = a_j \text{ for all } j \in [D_0] \\ \beta & \text{if } x_j = \bar{a}_j \text{ for all } j \in [D_0] \\ 0 & \text{otherwise}, \end{cases} \tag{40}$$

where $\alpha\beta \neq 0$. If $D_0 = 4$, this is a generalised equality of arity 4. If $D_0 > 4$, there must exist $j, k \in [D_0]$ with $j \neq k$ but $a_j = a_k$. Via a self-loop, we can realise the function $\sum_{x_j, x_k \in \{0,1\}} \mathrm{EQ}_2(x_j, x_k) f(\mathbf{x})$, which has arity $(D_0 - 2)$ and still satisfies a property like (40) for the remaining arguments. This process can be repeated, reducing the arity in steps of 2, to finally realise an arity-4 generalised equality function. Then #P-hardness follows by Lemma 33.

Case $D_0 \geq 3$ and D_0 is odd: Pin and use self-loops analogously to the previous case to realise a function of the form (40) with arity 3. Then #P-hardness follows by Lemma 58.

Case $D_0 = 2$: We can realise by pinning either a function $g = [\alpha, 0, \beta]$ or a function $\tilde{g} = \left(\begin{smallmatrix} 0 & \alpha \\ \beta & 0 \end{smallmatrix}\right)$ (up to permutation of the arguments), i.e. a generalised equality or a generalised disequality. In the second subcase, suppose the indices of f that were not pinned in order to realise $\tilde{g}$ are j and k. Replace f by the gadget $\sum_{z \in \{0,1\}} \tilde{g}(x_j, z) f(x_1, \ldots, x_{j-1}, z, x_{j+1}, \ldots, x_n)$. The new function has the same entanglement and the same D_0 as the old one, but pinning all inputs except j and k now gives rise to a generalised equality function as in the first subcase. If necessary, redefine α, β according to this new function. In the following, we assume $j = 1$, $k = 2$; this is without loss of generality as permuting arguments is a gadget operation.

Following the proof of [19, Theorem 5.1], define $f_{\mathbf{x}}(z_1, z_2) := f(z_1, z_2, \mathbf{x})$ for any $\mathbf{x} \in \{0,1\}^{n-2}$ and let

$$A_1 := \{\mathbf{x} \in \{0,1\}^{n-2} \mid f_{\mathbf{x}}(z_1, z_2) = c[\alpha, 0, \beta] \text{ for some } c \in \mathbb{A} \setminus \{0\}\},$$
$$B_1 := \{\mathbf{x} \in \{0,1\}^{n-2} \mid f_{\mathbf{x}}(z_1, z_2) \neq c[\alpha, 0, \beta] \text{ for any } c \in \mathbb{A}\}.$$

Note that A_1 is non-empty since we can realise g by pinning. The set B_1 must then be non-empty since otherwise f would be decomposable as $f(z_1, z_2, \mathbf{x}) = g(z_1, z_2) f'(\mathbf{x})$ for some $f' \in \Upsilon_{n-2}$. Furthermore, note that $A_1 \cap B_1 = \emptyset$ and that if $f_{\mathbf{y}}$ is identically 0, then $\mathbf{y}$ is not in $A_1 \cup B_1$. Thus we can define:

$$D_1 = \min\{d(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in A_1, \mathbf{y} \in B_1\}.$$

Pick a pair $\mathbf{a} \in A_1, \mathbf{b} \in B_1$ with minimum Hamming distance and pin wherever they are equal, as in the cases where $D_0 \geq 3$. This realises a function

$$h(x_1, \ldots, x_{D_1+2}) = g(x_1, x_2) \prod_{k=1}^{D_1} \delta_{a_k}(x_{k+2}) + g'(x_1, x_2) \prod_{k=1}^{D_1} \delta_{\bar{a}_k}(x_{k+2}),$$

where $g'(x_1, x_2) := f_{\mathbf{b}}(x_1, x_2)$. Note that the assumption $D_0 = 2$ implies that either $f_{\mathbf{y}}(0,1) = f_{\mathbf{y}}(1,0) = 0$ or $f_{\mathbf{y}}(0,0) = f_{\mathbf{y}}(1,1) = 0$ for all $\mathbf{y} \in B_1$. (Assume otherwise, i.e. each of the sets $\{f_{\mathbf{y}}(0,1), f_{\mathbf{y}}(1,0)\}$ and $\{f_{\mathbf{y}}(0,0), f_{\mathbf{y}}(1,1)\}$ contains at least one non-zero value. Then it is straightforward to see that there are two inputs for which f is non-zero and whose Hamming distance is 1, so by (39) we should have $D_0 = 1$, a contradiction.) Thus, g' is either $\left(\begin{smallmatrix} \lambda & 0 \\ 0 & \mu \end{smallmatrix}\right)$ or $\left(\begin{smallmatrix} 0 & \lambda \\ \mu & 0 \end{smallmatrix}\right)$ for some $\lambda, \mu \in \mathbb{A}$ such that λ, μ are not both zero. If $g' = \left(\begin{smallmatrix} \lambda & 0 \\ 0 & \mu \end{smallmatrix}\right)$, then the definition of B_1 furthermore implies $\alpha\mu - \beta\lambda \neq 0$.

We now distinguish cases according to the values of D_1.

- If $D_1 \geq 3$, then $\mathrm{arity}(h) = D_1 + 2 \geq 5$. Distinguish cases according to g'.

  - Suppose $g' = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$. If $\lambda \neq 0$, we can pin the first two inputs of h to 00 to get a function $\alpha(\prod_{k=1}^{D_1} \delta_{a_k}(x_{k+2})) + \lambda(\prod_{k=1}^{D_1} \delta_{\bar{a}_k}(x_{k+2}))$. The resulting function still has arity at least 3, so we can proceed as in the cases $D_0 \geq 4$ or $D_0 \geq 3$. If $\lambda = 0$ then $\mu \neq 0$ and we can pin to 11 instead to get the function $\beta(\prod_{k=1}^{D_1} \delta_{a_k}(x_{k+2})) + \mu \prod_{k=1}^{D_1} \delta_{\bar{a}_k}(x_{k+2}))$, and proceed analogously.

  - Suppose $g' = \begin{pmatrix} 0 & \lambda \\ \mu & 0 \end{pmatrix}$. If $\lambda \neq 0$, we can pin the first input of h to 0 to realise:

  $$\alpha \delta_0(x_2) \prod_{k=1}^{D_1} \delta_{a_k}(x_{k+2}) + \lambda \delta_1(x_2) \prod_{k=1}^{D_1} \delta_{\bar{a}_k}(x_{k+2})$$

  at which point we again proceed as in the cases $D_0 \geq 4$ or $D_0 \geq 3$. If $\lambda = 0$ then $\mu \neq 0$ and we can pin the first input to 1 instead to get the function

  $$\beta \delta_1(x_2) \prod_{k=1}^{D_1} \delta_{a_k}(x_{k+2}) + \mu \delta_0(x_2) \prod_{k=1}^{D_1} \delta_{\bar{a}_k}(x_{k+2}),$$

  and proceed analogously.

- If $D_1 = 2$, there are eight subcases, depending on the form of g' and the value of $\mathbf{a}$. They can be considered pairwise, grouping $\mathbf{a}$ with $\bar{\mathbf{a}}$.

  - If $g' = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ and $\mathbf{a} = 00$, the function after pinning is

  $$\begin{pmatrix} \alpha & 0 & 0 & \lambda \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \beta & 0 & 0 & \mu \end{pmatrix}$$

  with $\alpha\beta \neq 0$ and $\alpha\mu - \beta\lambda \neq 0$. In this case, apply Lemma 32 to show

  $$\mathsf{Holant}^c(\mathcal{F} \cup \{\mathrm{EQ}_4\}) \leq_T \mathsf{Holant}^c(\mathcal{F}).$$

  Now by Lemma 33, $\#\mathsf{CSP}_2^c(\mathcal{F} \cup \{\mathrm{EQ}_4\}) \leq_T \mathsf{Holant}^c(\mathcal{F} \cup \{\mathrm{EQ}_4\})$. Therefore hardness follows by Theorem 22. Here, the original proof in [19] used a different technique requiring real values. If $\mathbf{a} = 11$, the argument is analogous with the first and last columns swapped.

  - If $g' = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$ and $\mathbf{a} = 01$, the function after pinning is

  $$\begin{pmatrix} 0 & \alpha & \lambda & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \beta & \mu & 0 \end{pmatrix}$$

  where again $\alpha\beta \neq 0$ and $\alpha\mu - \beta\lambda \neq 0$. Call this function h'. Then the gadget $\sum_{y,z \in \{0,1\}} h'(x_1, x_2, y, z) h'(x_3, x_4, y, z)$ corresponds to taking the product of the above matrix with its transpose; it takes values

  $$\begin{pmatrix} \alpha^2 + \lambda^2 & 0 & 0 & \alpha\beta + \lambda\mu \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \alpha\beta + \lambda\mu & 0 & 0 & \beta^2 + \mu^2 \end{pmatrix}$$

Now
$$\det \begin{pmatrix} \alpha^2 + \lambda^2 & \alpha\beta + \lambda\mu \\ \alpha\beta + \lambda\mu & \beta^2 + \mu^2 \end{pmatrix} = (\alpha\mu - \beta\lambda)^2 \neq 0,$$

which means Lemma 32 can be applied again. Thus hardness follows as in the first subcase.

If $\mathbf{a} = 10$, the middle two columns of the first matrix are swapped, but the argument is analogous.

   — If $g' = \begin{pmatrix} 0 & \lambda \\ \mu & 0 \end{pmatrix}$ and $\mathbf{a} = 00$, the function after pinning is

$$\begin{pmatrix} \alpha & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda \\ 0 & 0 & 0 & \mu \\ \beta & 0 & 0 & 0 \end{pmatrix}.$$

If $\lambda \neq 0$, pin the first input of h to 0 to get the function $[\alpha, 0, 0, \lambda]$. If $\lambda = 0$ then $\mu \neq 0$, so pin the first input of h to 1 to get the function $\begin{pmatrix} 0 & 0 & 0 & \mu \\ \beta & 0 & 0 & 0 \end{pmatrix}$.

If $\mathbf{a} = 11$, then the first and last columns are swapped, the argument is otherwise analogous. In each case, the resulting function has arity 3 and is non-decomposable (in fact, it is a generalised equality), so we can show hardness via Lemma 58.

   — If $g' = \begin{pmatrix} 0 & \lambda \\ \mu & 0 \end{pmatrix}$ and $\mathbf{a} = 01$, the function after pinning is

$$\begin{pmatrix} 0 & \alpha & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & \mu & 0 \\ 0 & \beta & 0 & 0 \end{pmatrix}$$

and we can pin as in the previous subcase. If $\mathbf{a} = 10$, the middle columns are swapped and the process is analogous. All resulting functions have arity 3 and are non-decomposable (in fact, they are generalised equalities), so we can show hardness via Lemma 58.

- If $D_1 = 1$, then $h(x_1, x_2, x_3) = g(x_1, x_2)\delta_{a_1}(x_3) + g'(x_1, x_2)\delta_{\bar{a}_1}(x_3)$ is a non-decomposable ternary function since g and g' are linearly independent. Thus, we can apply Lemma 58 to show hardness.

Case $D_0 = 1$: By pinning, we can realise $g = [\alpha, \beta]$ for some $\alpha, \beta \neq 0$. Without loss of generality, assume this function arises by pinning the last $(n-1)$ inputs of f; otherwise permute the arguments. Define $f_{\mathbf{x}}(z) := f(z, \mathbf{x})$ for any $\mathbf{x} \in \{0, 1\}^{n-1}$ and let

$$A_2 := \{\mathbf{x} \in \{0, 1\}^{n-1} \mid f_{\mathbf{x}}(z) = c[\alpha, \beta] \text{ for some } c \in \mathbb{A} \setminus \{0\}\},$$
$$B_2 := \{\mathbf{x} \in \{0, 1\}^{n-1} \mid f_{\mathbf{x}}(z) \neq c[\alpha, \beta] \text{ for any } c \in \mathbb{A}\}.$$

Analogously to A_1 and B_1, these two sets are non-empty, do not intersect, and do not contain any $\mathbf{x}$ such that $f_{\mathbf{x}}(z)$ is identically zero. Then let:

$$D_2 = \min\{d(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in A_2, \mathbf{y} \in B_2\},$$

which is well-defined. By pinning we can realise a function

$$h(x_1, \ldots, x_{D_2+1}) = g(x_1) \prod_{k=1}^{D_2} \delta_{a_k}(x_{k+1}) + g'(x_1) \prod_{k=1}^{D_2} \delta_{\bar{a}_k}(x_{k+1}),$$

where $g' = [\lambda, \mu]$ with $\alpha\mu - \beta\lambda \neq 0$.

- If $D_2 \geq 3$, then h has arity greater than 3. If $\lambda \neq 0$, pin the first input of h to 0, which yields a generalised equality function of arity at least 3, and then proceed as in the cases $D_0 \geq 4$ or $D_0 \geq 3$. If $\lambda = 0$ then $\mu \neq 0$, so we can pin to 1 instead for an analogous result.

- If $D_2 = 2$, h is a non-decomposable ternary function so we are done by Lemma 58. This is another change compared to the proof in [19], where hardness was only shown for a real-valued function of the given form.

- If $D_2 = 1$, h is a non-decomposable binary function $\left(\begin{smallmatrix} \alpha & \beta \\ \lambda & \mu \end{smallmatrix} \right)$ (possibly up to a bit flip of the second argument, which does not affect the proof). Unlike in [19], we do not attempt to use this binary function for interpolation. Instead we immediately proceed to defining A_3, B_3, and D_3 analogous to before. Without loss of generality, assume that the two variables of h correspond to the first two variables of f (otherwise permute arguments). Let

$$A_3 := \{\mathbf{x} \in \{0,1\}^{n-2} \mid f_{\mathbf{x}}(z_1, z_2) = c \cdot h(z_1, z_2) \text{ for some } c \in \mathbb{A} \setminus \{0\}\},$$
$$B_3 := \{\mathbf{x} \in \{0,1\}^{n-2} \mid f_{\mathbf{x}}(z_1, z_2) \neq c \cdot h(z_1, z_2) \text{ for any } c \in \mathbb{A}\},$$

then A_3 and B_3 are non-empty, they do not intersect, and they do not contain any $f_{\mathbf{x}}$ which is identically zero. Thus we may define $D_3 = \min\{d(\mathbf{x}, \mathbf{y}) \mid \mathbf{x} \in A_3, \mathbf{y} \in B_3\}$. Let $\mathbf{a} \in A_3$ be such that there exists $\mathbf{b} \in B_3$ with $d(\mathbf{a}, \mathbf{b}) = D_3$. By pinning in all places that $\mathbf{a}$ and $\mathbf{b}$ agree, we realise a function

$$h(x_1, x_2) \prod_{k=1}^{D_3} \delta_{a_k}(x_{k+2}) + h'(x_1, x_2) \prod_{k=1}^{D_3} \delta_{\bar{a}_k}(x_{k+2}) \tag{41}$$

where $h'(x_1, x_2) := f_{\mathbf{b}}(x_1, x_2)$ is not a scaling of h. Suppose $h' = \left(\begin{smallmatrix} \alpha' & \beta' \\ \lambda' & \mu' \end{smallmatrix} \right)$. Distinguish cases according to D_3. It will be useful to consider these in ascending order since the case $D_3 = 2$ follows straightforwardly from $D_3 = 1$.

  - If $D_3 = 1$, the function in (41) is ternary. Note that the case $a_1 = 1$ differs from the case $a_1 = 0$ only by a bit flip on the third input, which is a SLOCC transformation by $(I \otimes I \otimes X)$ and thus does not affect entanglement. So it suffices to consider $a_1 = 0$; the function then takes values

$$\begin{pmatrix} \alpha & \alpha' & \beta & \beta' \\ \lambda & \lambda' & \mu & \mu' \end{pmatrix}.$$

Recall that $\alpha, \beta \neq 0$, the unprimed variables satisfy $\alpha\mu - \beta\lambda \neq 0$, and $(\alpha', \beta', \lambda', \mu')$ is not a scaling of $(\alpha, \beta, \lambda, \mu)$.

By Lemma 36, the ternary function is non-decomposable unless at least two of the following three expressions are false:

$$(\alpha\beta' \neq \beta\alpha') \vee (\mu\lambda' \neq \lambda\mu')$$
$$(\lambda\alpha' \neq \alpha\lambda') \vee (\mu\beta' \neq \beta\mu')$$
$$(\beta'\lambda' \neq \alpha'\mu') \vee (\beta\lambda \neq \alpha\mu)$$

Now, the third expression is always true since $\alpha\mu - \beta\lambda \neq 0$ implies $\beta\lambda \neq \alpha\mu$. So the function being decomposable would imply $\alpha\beta' = \beta\alpha'$, $\mu\lambda' = \lambda\mu'$, $\lambda\alpha' = \alpha\lambda'$ and $\mu\beta' = \beta\mu'$. But the first of these equations implies $\beta' = \beta\alpha'/\alpha$ since $\alpha \neq 0$. Similarly, the third equation implies $\lambda' = \lambda\alpha'/\alpha$. Also, since $\beta \neq 0$, the fourth equation implies $\mu' = \mu\beta'/\beta = \mu\alpha'/\alpha$. Thus,

$$h' = \begin{pmatrix} \alpha' & \beta' \\ \lambda' & \mu' \end{pmatrix} = \frac{\alpha'}{\alpha} \begin{pmatrix} \alpha & \beta \\ \lambda & \mu \end{pmatrix} = \frac{\alpha'}{\alpha} \cdot h,$$

which is a contradiction since h' is not a scaling of h. Thus, the ternary function must be non-decomposable, and hardness follows by Lemma 58.

This is a change compared to the proof in [19], where multiple cases were distinguished and the hardness lemmas only applied to real-valued functions.

– If $D_3 = 2$, the function in (41) has arity 4. By connecting $[\alpha, \beta]$ to the last input, we get one of the functions

$$\alpha \cdot h(x_1, x_2)\delta_{a_1}(x_3) + \beta \cdot h'(x_1, x_2)\delta_{\bar{a}_1}(x_3)$$
$$\beta \cdot h(x_1, x_2)\delta_{a_1}(x_3) + \alpha \cdot h'(x_1, x_2)\delta_{\bar{a}_1}(x_3)$$

depending on whether a_2 is 0 or 1. Note that these functions differ from the function considered in the previous subcase – $h(x_1, x_2)\delta_{a_1}(x_3) + h'(x_1, x_2)\delta_{\bar{a}_1}(x_3)$ – only by a SLOCC with $\begin{pmatrix} \alpha & 0 \\ 0 & \beta \end{pmatrix}$ or $\begin{pmatrix} \beta & 0 \\ 0 & \alpha \end{pmatrix}$ on the third input (depending on whether $a_1 = a_2$). Hence they must be in the same entanglement class and hardness follows as in the previous subcase.

This is another change compared to [19], where the hardness lemma only applied to real values and the construction did not employ the function $[\alpha, \beta]$.

– If $D_3 \geq 3$, we consider different cases according to the relationships between the values of h and h'.

If $\alpha\alpha' = \beta\beta' = \lambda\lambda' = \mu\mu' = 0$, then $\alpha' = \beta' = 0$ since α, β are non-zero by assumption. Thus, at least one of λ' and μ' must be non-zero because h' is not identically zero. Also, at least one of λ and μ must be non-zero because h is non-decomposable. Therefore we have either $\lambda = \mu' = 0$ and $\lambda', \mu \neq 0$, or $\lambda' = \mu = 0$ and $\lambda, \mu' \neq 0$. In the first case, $\lambda = \mu' = 0$ and $\lambda', \mu \neq 0$, pin the first input to 1 to get:

$$\mu\delta_1(x_2)\prod_{k=1}^{D_3}\delta_{a_k}(x_{k+2}) + \lambda'\delta_0(x_2)\prod_{k=1}^{D_3}\delta_{\bar{a}_k}(x_{k+2}).$$

This is a generalised equality function of arity at least 4, so we can proceed as before. In the second case, $\lambda' = \mu = 0$ and $\lambda\mu' \neq 0$, pinning the first input to 1 works again, albeit resulting in a different generalised equality function.

Otherwise, there exists a pair of primed and unprimed coefficients of the same label that are both non-zero. If these are α and α', pin the first two inputs to 00 to get a generalised equality of arity at least 3. If the non-zero pair are β and β', pin to 01, and so on. Given the generalised equality function, we may proceed as in the cases $D_0 \geq 4$ or $D_0 \geq 3$, depending on whether the arity of this function is even or odd.

We have covered all cases, hence the proof is complete. $\qquad\square$

Acknowledgements

References

[1] Miriam Backens. A New Holant Dichotomy Inspired by Quantum Computation. In Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl, editors, *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, volume 80 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 16:1–16:14, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[2] Miriam Backens. Number of superclasses of four-qubit entangled states under the inductive entanglement classification. *Phys. Rev. A*, 95(2):022329, February 2017.

[3] Miriam Backens. A complete dichotomy for complex-valued Holant^c. In Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella, editors, *45th International Colloquium on Automata, Languages, and Programming (ICALP 2018)*, volume 107 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 12:1–12:14, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[4] Miriam Backens and Leslie Ann Goldberg. Holant Clones and the Approximability of Conservative Holant Problems. *ACM Trans. Algorithms*, 16(2):23:1–23:55, March 2020.

[5] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Phys. Rev. Lett.*, 70(13):1895–1899, March 1993.

[6] Charles H. Bennett, Sandu Popescu, Daniel Rohrlich, John A. Smolin, and Ashish V. Thapliyal. Exact and asymptotic measures of multipartite pure-state entanglement. *Phys. Rev. A*, 63(1):012307, December 2000.

[7] P. Oscar Boykin, Tal Mor, Matthew Pulver, Vwani Roychowdhury, and Farrokh Vatan. On universal and fault-tolerant quantum computing: A novel basis and a new constructive proof of universality for Shor's basis. In *40th Annual Symposium on Foundations of Computer Science (Cat. No.99CB37039)*, pages 486–494, New York, October 1999. IEEE.

[8] Andrei Bulatov, Leslie Ann Goldberg, Mark Jerrum, David Richerby, and Stanislav Živný. Functional clones and expressibility of partition functions. *Theoret. Comput. Sci.*, 687:11–39, July 2017.

[9] Andrei A. Bulatov, Martin Dyer, Leslie Ann Goldberg, Mark Jerrum, and Colin McQuillan. The Expressibility of Functions on the Boolean Domain, with Applications to Counting CSPs. *J. ACM*, 60(5):32:1–32:36, October 2013.

[10] J. Cai, X. Chen, and P. Lu. Graph Homomorphisms with Complex Values: A Dichotomy Theorem. *SIAM J. Comput.*, 42(3):924–1029, January 2013.

[11] J. Cai, H. Guo, and T. Williams. A Complete Dichotomy Rises from the Capture of Vanishing Signatures. *SIAM J. Comput.*, 45(5):1671–1728, January 2016.

[12] J. Cai, P. Lu, and M. Xia. Dichotomy for Holant* Problems of Boolean Domain. In *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, Proceedings, pages 1714–1728. Society for Industrial and Applied Mathematics, January 2011.

[13] Jin-Yi Cai and Xi Chen. *Complexity Dichotomies for Counting Problems*, volume 1: *Boolean Domain*. Cambridge University Press, November 2017.

[14] Jin-Yi Cai and Vinay Choudhary. Valiants Holant Theorem and matchgate tensors. *Theoret. Comput. Sci.*, 384(1):22–32, September 2007.

[15] Jin-Yi Cai and Zhiguo Fu. Holographic Algorithm with Matchgates Is Universal for Planar #CSP over Boolean Domain. *SIAM J. Comput.*, pages STOC17-50–STOC17-151, November 2019.

[16] Jin-Yi Cai, Heng Guo, and Tyson Williams. Clifford gates in the Holant framework. *Theoret. Comput. Sci.*, 745:163–171, 2018.

[17] Jin-Yi Cai, Sangxia Huang, and Pinyan Lu. From Holant to #CSP and Back: Dichotomy for Holant^c Problems. *Algorithmica*, 64(3):511–533, March 2012.

[18] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. The complexity of complex weighted Boolean #CSP. *J. Comput. System Sci.*, 80(1):217–236, February 2014. First appeared as "Holant Problems and Counting CSP" in STOC '09.

[19] Jin-Yi Cai, Pinyan Lu, and Mingji Xia. Dichotomy for Real Holant^c Problems. In *Proceedings of the Twenty-Ninth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1802–1821. Society for Industrial and Applied Mathematics, 2018.

[20] Jeroen Dehaene and Bart De Moor. Clifford group, stabilizer states, and linear and quadratic operations over GF(2). *Phys. Rev. A*, 68(4):042318, October 2003.

[21] Simon J Devitt, William J Munro, and Kae Nemoto. Quantum error correction for beginners. *Rep. Progr. Phys.*, 76(7):076001, July 2013.

[22] W. Dür, G. Vidal, and J. I. Cirac. Three qubits can be entangled in two inequivalent ways. *Phys. Rev. A*, 62(6):062314, November 2000.

[23] M. Dyer, L. A. Goldberg, and M. Jerrum. The Complexity of Weighted Boolean CSP. *SIAM J. Comput.*, 38(5):1970–1986, January 2009.

[24] Artur K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67(6):661–663, August 1991.

[25] Mariami Gachechiladze and Otfried Gühne, February 2017. Personal communication.

[26] Mariami Gachechiladze and Otfried Gühne. Completing the proof of Generic quantum nonlocality. *Phys. Lett. A*, 381(15):1281–1285, April 2017.

[27] Brett Giles and Peter Selinger. Exact synthesis of multiqubit Clifford+T circuits. *Phys. Rev. A*, 87(3):032332, March 2013.

[28] Daniel Gottesman. The Heisenberg Representation of Quantum Computers. *arXiv:quant-ph/9807006*, July 1998. Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics, eds. S. P. Corney, R. Delbourgo, and P. D. Jarvis, pp. 32-43 (Cambridge, MA, International Press, 1999).

[29] Richard Jozsa and Noah Linden. On the role of entanglement in quantum-computational speed-up. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 459(2036):2011–2032, 2003.

[30] Michael Kowalczyk and Jin-Yi Cai. Holant Problems for 3-Regular Graphs with Complex Edge Functions. *Theory Comput. Syst.*, 59(1):133–158, July 2016.

[31] L. Lamata, J. Len, D. Salgado, and E. Solano. Inductive entanglement classification of four qubits under stochastic local operations and classical communication. *Phys. Rev. A*, 75(2):022318, February 2007.

[32] Dafa Li, Xiangrong Li, Hongtao Huang, and Xinxin Li. Simple criteria for the SLOCC classification. *Phys. Lett. A*, 359(5):428–437, December 2006.

[33] Jiabao Lin and Hanpin Wang. The Complexity of Holant Problems over Boolean Domain with Non-Negative Weights. In Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl, editors, *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, volume 80 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 29:1–29:14, Dagstuhl, Germany, 2017. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[34] Ashley Montanaro and Dan J. Shepherd. Hadamard gates and amplitudes of computational basis states, September 2006. Available at `http://www.maths.bris.ac.uk/~csxam/papers/hadamard.pdf`.

[35] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2010.

[36] Sandu Popescu and Daniel Rohrlich. Generic quantum nonlocality. *Phys. Lett. A*, 166(56):293–297, June 1992.

[37] L. Valiant. Holographic Algorithms. *SIAM J. Comput.*, 37(5):1565–1594, January 2008.

[38] F. Verstraete, J. Dehaene, B. De Moor, and H. Verschelde. Four qubits can be entangled in nine different ways. *Phys. Rev. A*, 65(5):052112, April 2002.