7

Holant Dichotomy for Symmetric Constraints

Most of the results in previous chapters can be viewed as preparatory for the dichotomy theorems to be achieved in this and the next two chapters. In this chapter we prove a dichotomy theorem valid for all Holant problems on an arbitrary set of complex-valued *symmetric* constraint functions on Boolean variables, Theorem 7.19. This includes as special cases the dichotomy theorems for Holant* in Chapter 2 (when all unary functions are free), Holant^c in Chapter 6 (when the two unary functions Is-Zero $[1, 0]$ and Is-One $[0, 1]$ are free), and #CSP in Chapter 3 (when all Equalities are free) *restricted* to symmetric constraints.

For Holant problems without freely available auxiliary functions such as unary signatures, the construction of gadgets becomes more constrained, and thus more difficult. This difficulty has shown itself in Chapter 5. In particular, the only natural local operations available are to connect two edges, or to make a loop. If we want to prove hardness with some inductive argument based on the arity of constraint functions, there will be two base cases of arity 3 and 4. Note that signatures of arity ≤ 2 define tractable Holant problems. The arity 3 case is understood from Theorem 6.3 (and also Theorem 6.4), with the main work done in Chapter 5. Therefore it will be natural to study carefully arity 4 signatures.

However, before we can develop the theory satisfactorily for arity 4 signatures we must first deal with a phenomenon called *vanishing*.

7.1 Vanishing Signatures

Definition 7.1. A set of signatures $\mathcal{F}$ is called *vanishing* if the value Holant$(\Omega; \mathcal{F})$ is zero for every signature grid Ω. A signature f is called *vanishing* if the singleton set $\{f\}$ is vanishing.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

In this section, we characterize all sets of symmetric vanishing signatures. The following lemma makes the simple observation that vanishing signatures are closed under gadget constructions, and under linear combinations. Let $af + bg$ denote the entry-wise linear combination of two signatures f and g with the same arity n ($a, b \in \mathbb{C}$), i.e., $(af + bg)(x) = af(x) + bg(x)$ for all $x \in \{0, 1\}^n$.

Lemma 7.2. *Let $\mathcal{F}$ be a vanishing signature set. If a signature f can be realized by a gadget using signatures in $\mathcal{F}$, then $\mathcal{F} \cup \{f\}$ is also vanishing. If f and g are two signatures in $\mathcal{F}$ of the same arity, and $a, b \in \mathbb{C}$, then $\mathcal{F} \cup \{af + bg\}$ is vanishing as well.*

Proof. That $\mathcal{F} \cup \{f\}$ is vanishing is obvious, since any vertex in a signature grid assigned the function f can be replaced by the $\mathcal{F}$-gate defining it. If $f, g \in \mathcal{F}$, then the Holant value on any signature grid using $\mathcal{F} \cup \{af + bg\}$ can be expressed as a linear combination of Holant values using $\mathcal{F}$, one term for a choice of f or g at every vertex labeled $af + bg$. □

Obviously, the identically 0 signature, in which all entries are 0, is vanishing. This is trivial. However, there are less trivial examples of vanishing signatures. Notice that the unary signature $[1, \mathrm{i}]$ when connected to another $[1, \mathrm{i}]$ has a Holant value 0. Consider a signature set $\mathcal{F}$ where every signature is degenerate, and moreover, as a tensor product of unary signatures, more than half of them are $[1, \mathrm{i}]$. For any signature grid Ω with signatures from $\mathcal{F}$, each signature is decomposed into unary signatures, and Ω is decomposed into pairs of vertices, one for each edge. The total Holant value is the product of the Holant on each pair. Since more than half of the unaries in each signature are $[1, \mathrm{i}]$, more than half of the unaries in Ω are $[1, \mathrm{i}]$. Then two $[1, \mathrm{i}]$'s must be paired up and hence $\mathrm{Holant}_\Omega = 0$. Thus, all such signatures form a vanishing set. We also observe that this argument holds when $[1, \mathrm{i}]$ is replaced by $[1, -\mathrm{i}]$.

These signatures described above are generally not symmetric and our present aim is to characterize symmetric vanishing signatures. To this end, we define the following symmetrization operation.

Definition 7.3. Let S_n be the symmetric group of degree $n \geq 1$. Then for any integer t ($0 \leq t \leq n$) and unary signatures $v, v_1, \ldots, v_{n-t}$, we define

$$\mathrm{Sym}_n^t(v; v_1, \ldots, v_{n-t}) = \sum_{\pi \in S_n} \bigotimes_{k=1}^{n} u_{\pi(k)}, \tag{7.1}$$

where the ordered sequence $(u_1, u_2, \ldots, u_n) = (\underbrace{v, \ldots, v}_{t \text{ copies}}, v_1, \ldots, v_{n-t})$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Note that we include redundant permutations of v in the definition. Identical v_i's also introduce redundant permutations. These redundant permutations simply introduce a nonzero constant factor. However, the allowance of redundant permutations may simplify calculations. If $t = 1$, then $\mathrm{Sym}_n^1(u_1; u_2, \ldots, u_n)$ is a sum of $n!$ tensor products by permuting $u_1, \ldots, u_n \in \mathbb{C}^2$ in all possible ways, and the n unary signatures are treated completely symmetrically.

An illustrative example of Definition 7.3 is

$$\begin{aligned}\mathrm{Sym}_3^2([1, \mathrm{i}]; [a, b]) &= 2[a, b] \otimes [1, \mathrm{i}] \otimes [1, \mathrm{i}] + 2[1, \mathrm{i}] \otimes [a, b] \otimes [1, \mathrm{i}] \\ &\quad + 2[1, \mathrm{i}] \otimes [1, \mathrm{i}] \otimes [a, b] \\ &= 2[3a, 2\mathrm{i}a + b, -a + 2\mathrm{i}b, -3b].\end{aligned}$$

Definition 7.4. A nonzero symmetric signature f of arity n has *positive* $(+)$ *vanishing degree* $k \geq 0$, denoted by $\mathrm{vd}^+(f) = k$, if $k \leq n$ is the largest integer such that there exist $n - k$ unary signatures $v_1, \ldots, v_{n-k}$ satisfying

$$f = \mathrm{Sym}_n^k([1, \mathrm{i}]; v_1, \ldots, v_{n-k}).$$

If f is the all-zero signature, we define $\mathrm{vd}^+(f) = n + 1$.

We define *negative* $(-)$ *vanishing degree* vd^- similarly, using $-\mathrm{i}$ instead of i.

We will prove in Section 7.1.1 (Lemma 7.8) that any nonzero symmetric signature has an expression as a symmetrization of unary signatures as in Definition 7.4. Thus $\mathrm{vd}^\sigma(f)$ is well defined.

Note that $\mathrm{vd}^+(f) = 0$ iff f cannot be expressed as a symmetrization of n unary signatures consisting of at least one $[1, \mathrm{i}]$ (but an expression as in Definition 7.4 for n unary signatures $v_1, \ldots, v_n$ still exists). Also $\mathrm{vd}^+(f) = n$ iff f is the degenerate signature $[1, \mathrm{i}]^{\otimes n}$ (recall that we treat two signatures as the same if their ratio is a nonzero constant).

In Chapter 2, we showed in (2.10) to (2.12) that any generalized Fibonacci gate f with parameter $\lambda = 2\mathrm{i}$, i.e., $f_k = ck\mathrm{i}^{k-1} + d\mathrm{i}^k$ (for some $c, d \in \mathbb{C}$, and for all $0 \leq k \leq n$), has the form $f = Z^{\otimes n}[a, b, 0, \ldots, 0]$, for some $a, b \in \mathbb{C}$, where $Z = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$. A similar expression is true if the parameter $\lambda = -2\mathrm{i}$; in that case $f = Z^{\otimes n}[0, \ldots, 0, b, a]$. The expression in (2.12) shows that $\mathrm{vd}^+(Z^{\otimes n}[1, 0, 0, \ldots, 0]) = n$ and $\mathrm{vd}^+(Z^{\otimes n}[0, 1, 0, \ldots, 0]) = n - 1$. For the linear combination $f = Z^{\otimes n}[a, b, 0, \ldots, 0] = aZ^{\otimes n}[1, 0, 0, \ldots, 0] + bZ^{\otimes n}[0, 1, 0, \ldots, 0]$, Lemma 7.8 in Section 7.1.1 will show that $\mathrm{vd}^+(f) = n - 1$ if $b \neq 0$. Similarly for $f = Z^{\otimes n}[0, \ldots, 0, b, a]$, $\mathrm{vd}^-(f) = n - 1$ if $b \neq 0$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Definition 7.5. If f is a symmetric signature and $\sigma \in \{+, -\}$, then we define its recurrence degree $\mathrm{rd}^{\sigma}(f)$ to be

$$\mathrm{rd}^{\sigma}(f) = \mathrm{arity}(f) - \mathrm{vd}^{\sigma}(f).$$

For generalized Fibonacci gates f of arity n and parameter 2i, $\mathrm{vd}^{+}(f) \geq n - 1$ and $\mathrm{rd}^{\sigma}(f) \leq 1$, and equalities hold iff f is nondegenerate. All nonzero entries in the expression $(Z^{-1})^{\otimes n} f = [a, b, 0, \ldots, 0]$ have indices of Hamming weight ≤ 1, and there are $\mathrm{vd}^{+}(f) \geq n - 1$ many trailing zeros.

Notice that it is possible for a signature f to have both $\mathrm{vd}^{+}(f)$ and $\mathrm{vd}^{-}(f)$ nonzero. For example, $f = [1, 0, 1]$ has $\mathrm{vd}^{+}(f) = \mathrm{vd}^{-}(f) = 1$.

By Lemma 7.2 and the discussion after Lemma 7.2, we know that for a signature f of arity n, if $\mathrm{vd}^{\sigma}(f) > \frac{n}{2}$ for some $\sigma \in \{+, -\}$, then f is a vanishing signature. In Section 7.1.1, Lemma 7.8 will give another simple explanation for this fact. In particular, all generalized Fibonacci gates of arity $n \geq 3$ with parameter $\lambda = \pm 2\mathrm{i}$ are vanishing signatures.

This is easily generalized to a set of signatures.

Definition 7.6. For $\sigma \in \{+, -\}$, we define $\mathscr{V}^{\sigma} = \{f | 2\,\mathrm{vd}^{\sigma}(f) > \mathrm{arity}(f)\}$.

Equivalently,

$$\mathscr{V}^{\sigma} = \{f | 2\,\mathrm{rd}^{\sigma}(f) < \mathrm{arity}(f)\}.$$

Lemma 7.7. *For a set of symmetric signatures $\mathcal{F}$, if $\mathcal{F} \subseteq \mathscr{V}^{+}$ or $\mathcal{F} \subseteq \mathscr{V}^{-}$, then $\mathcal{F}$ is vanishing.*

In Theorem 7.11, we show that these two sets capture all symmetric vanishing signature sets.

7.1.1 Characterization of Vanishing Signatures via Z

There is a simple characterization of the set of vanishing signatures in terms of the holographic transformation by $Z = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$. Suppose f has $\mathrm{arity}(f) = n$ and $\mathrm{rd}^{+}(f) = d$. Then $f \in \mathscr{V}^{+}$ iff $d < n/2$. We perform a holographic transformation with Z,

$$\begin{aligned} \mathrm{Holant}(=_2 \mid f) &\equiv_{\mathrm{T}} \mathrm{Holant}([1, 0, 1]Z^{\otimes 2} \mid (Z^{-1})^{\otimes n} f) \\ &\equiv_{\mathrm{T}} \mathrm{Holant}([0, 1, 0] \mid \hat{f}), \end{aligned}$$

where $\hat{f} = (Z^{-1})^{\otimes n} f$. We claim that $\hat{f}$ has the form $[\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0]$, and $\hat{f}_d \neq 0$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Lemma 7.8. *For any symmetric signature f, its vanishing degree $\mathrm{vd}^{\sigma}(f)$ and recurrence degree $\mathrm{rd}^{\sigma}(f)$ are well defined, for $\sigma \in \{+, -\}$.*

Let $\hat{f} = (Z^{-1})^{\otimes n} f$. If $\mathrm{rd}^{+}(f) = d$, then $\hat{f} = [\hat{f}_0, \ldots, \hat{f}_d, 0, \ldots, 0]$ and $\hat{f}_d \neq 0$. The number of trailing 0 entries after $\hat{f}_d$ is $\mathrm{vd}^{+}(f)$. In particular, $f \in \mathscr{V}^{+}$ iff all nonzero entries $\hat{f}_i$ occur in strictly the first half of $\hat{f}$, i.e., all have index $i < \frac{n}{2}$.

Similarly, if $\mathrm{rd}^{-}(f) = d$, then $\hat{f} = [0, \ldots, 0, \hat{f}_{n-d}, \ldots, \hat{f}_n]$ and $\hat{f}_{n-d} \neq 0$. The number of leading 0 entries before $\hat{f}_{n-d}$ is $\mathrm{vd}^{-}(f)$. In particular, $f \in \mathscr{V}^{-}$ iff all nonzero entries $\hat{f}_i$ occur in strictly the second half of $\hat{f}$, i.e., all have index $i > \frac{n}{2}$.

Proof. Suppose a nonzero symmetric signature $f = Z^{\otimes n} \hat{f}$ has an expression

$$f = \mathrm{Sym}_n^{n-t}([1, \mathrm{i}]; v_1, \ldots, v_t), \tag{7.2}$$

for some $0 \leq t \leq n$ and $v_1, \ldots, v_t \in \mathbb{C}^2$. Since

$$Z = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}, Z^{-1} = \tfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{bmatrix},$$

we have $Z^{-1}[1, \mathrm{i}] = \sqrt{2}[1, 0]$. Thus, f having an expression of the form (7.2) is equivalent to $\hat{f}$ having an expression of the form

$$\hat{f} = \mathrm{Sym}_n^{n-t}([1, 0]; u_1, \ldots, u_t), \tag{7.3}$$

up to a nonzero scalar factor, where $u_i = Z^{-1} v_i \in \mathbb{C}^2$ for $1 \leq i \leq t$.

From the expression (7.3) for $\hat{f} = [\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_n]$, it is clear that all entries of Hamming weight greater than t in $\hat{f}$ are 0. Moreover, if t is the minimum integer ($0 \leq t \leq n$) such that (7.2) holds for some $v_1, \ldots, v_t \in \mathbb{C}^2$, then $\hat{f}_t \neq 0$. This is clear if $t = 0$. Suppose $t > 0$ and for a contradiction $\hat{f}_t = 0$. Then one of the u_i must be a multiple of $[1, 0]$. This would imply that one of the v_i is a multiple of $[1, \mathrm{i}]$ in (7.2), contradicting the minimality of t.

Now suppose

$$\hat{f} = [\hat{f}_0, \ldots, \hat{f}_t, 0, \ldots, 0],$$

where $0 \leq t \leq n$ and $\hat{f}_t \neq 0$. Such a t is uniquely defined for any nonzero symmetric signature $\hat{f}$, hence also for $f = Z^{\otimes n} \hat{f}$. We can normalize $\hat{f}_t = t!(n-t)!$ by a nonzero scalar multiple. We prove that there are $u_1 = [a_1, 1], \ldots, u_t = [a_t, 1] \in \mathbb{C}^2$ such that

$$\hat{f} = \mathrm{Sym}_n^{n-t}([1, 0]; u_1, \ldots, u_t).$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Let F denote the signature defined by the right-hand side (RHS), for some $a_1, \ldots, a_t \in \mathbb{C}$ to be determined. Then clearly $F_k = 0$ for $k > t$. Define

$$c_k = \frac{\hat{f}_k}{k!(n-k)!}, \quad \text{for } 0 \le k \le t.$$

Then $c_t = 1$, and over the algebraically closed field $\mathbb{C}$ we can factor the polynomial

$$\sum_{k=0}^{t} c_k X^k = (X + a_1)(X + a_2) \cdots (X + a_t)$$

for some $a_1, a_2, \ldots, a_t \in \mathbb{C}$. It follows that $c_{t-k} = p_k(a_1, a_2, \ldots, a_t)$, where p_k is the elementary symmetric polynomial of degree k on t variables.

For $0 \le k \le t$,

$$\begin{aligned} F_k &= \binom{n-k}{n-t}(n-t)! \sum_{S \subseteq [t]:|S|=t-k} (t-k)!k! \prod_{i \in S} a_i \\ &= (n-k)!k! p_{t-k}(b_1, a_2, \ldots, a_t) = \hat{f}_k, \end{aligned}$$

because on input $(\underbrace{0, \ldots, 0}_{n-k}, \underbrace{1, \ldots, 1}_{k})$ of Hamming weight k, a term with a nonzero contribution in the sum $\mathrm{Sym}_n^{n-t}([1, 0]; u_1, \ldots, u_t)$ defining F must have all $n - t$ copies of $[1, 0]$ occurring in the first $n - k$ positions, and there is a subset of $[a_i, 1]$'s to be in the remaining $(n - k) - (n - t) = t - k$ positions.

The proof for $\mathrm{vd}^-(f)$ and $\mathrm{rd}^-(f)$ is symmetric by replacing i by $-\mathrm{i}$. □

For future reference, we also note the following. If $f = g + h$ is of arity n, where $\mathrm{rd}^+(g) = d$, $\mathrm{rd}^-(h) = d'$, and $d + d' < n$, then after a holographic transformation by Z, $\hat{f} = (Z^{-1})^{\otimes n} f$ takes the form $[\widehat{g}_0, \ldots, \widehat{g}_d, \underbrace{0, \ldots, 0}_{n-d-d'-1}, \widehat{h}_{n-d'}, \ldots, \widehat{h}_n]$, where $n - d - d' - 1 \ge 0$ zeros are present.

Suppose $f \in \mathscr{V}^+$. The expression $f = Z^{\otimes n}[\hat{f}_0, \ldots, \hat{f}_d, 0, \ldots, 0]$ with $d < n/2$ also offers an alternative explanation for why f is vanishing. In any bipartite graph for $\mathrm{Holant}([0, 1, 0] \mid \hat{f})$, the binary $\textsc{Disequality}$ $(\neq_2) = [0, 1, 0]$ on the left imposes the condition that half of the edges must take the value 0 and the other half must take the value 1. On the right side we have $d < n/2$; thus $\hat{f}$ requires that less than half of the edges are assigned the value 1. Therefore the Holant is always 0.

Under the Z transformation, one can observe another interesting phenomenon. For any $a, b \in \mathbb{C}$,

$$\mathrm{Holant}([0, 1, 0] \mid [a, b, 1, 0, 0]) \quad \text{and} \quad \mathrm{Holant}([0, 1, 0] \mid [0, 0, 1, 0, 0])$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

take exactly the same value on every signature grid. This is because, to contribute a nonzero term in the Holant, exactly half of the edges must be assigned 1. Then for the first problem, the signature on the right can never contribute a nonzero value involving a or b. Thus the Holant values of these two problems on any signature grid are always the same. Nevertheless, for every $a, b \in \mathbb{C}$ not both 0, there is no holographic transformation between these two problems. This shows that the converse of the Holant Theorem (Theorem 1.5) does not hold.

We know that Holant([0, 1, 0] | [0, 0, 1, 0, 0]) counts the number of Eulerian orientations in a 4-regular graph. This problem is #P-hard (Theorem 6.32) and in fact its planar version Pl-Holant($\neq_2$ | [0, 0, 1, 0, 0]) is also #P-hard for 4-regular graph. This will play an important role in proofs later. Translating back to the standard setting, the problem of counting Eulerian orientations in a 4-regular graph is Holant([3, 0, 1, 0, 3]). The problem Holant([0, 1, 0] | $[a, b, 1, 0, 0]$) corresponds to a certain signature $f = Z^{\otimes 4}[a, b, 1, 0, 0]$ of arity 4 with recurrence degree 2. It has a different appearance but induces exactly the same Holant value as the signature for counting Eulerian orientations. Therefore, all such signatures are #P-hard as well. We use this fact later.

We remark that for $f = Z^{\otimes n}[\hat{f}_0, \ldots, \hat{f}_n]$, the signature obtained by a self-loop on f is simply $f' = 2Z^{\otimes n}[\hat{f}_1, \ldots, \hat{f}_{n-1}]$, by removing the extremal entries $\hat{f}_0$ and $\hat{f}_n$. This is easily seen by considering it as an $\mathcal{F}$-gate construction in Holant($\neq_2 \mid \hat{f}$), where two input bits of $\hat{f}$ on the loop are each forced to take exactly 0 and 1, or 1 and 0 respectively. This operation is very handy, and can be repeated.

7.1.2 Characterization of Vanishing Signature Sets

Now we show that $\mathscr{V}^+$ and $\mathscr{V}^-$ capture all symmetric vanishing signature sets. To begin, we show that a vanishing signature set cannot contain both types of nonzero vanishing signatures.

Lemma 7.9. *Let $f \in \mathscr{V}^+$ and $g \in \mathscr{V}^-$. If neither f nor g is the zero signature, then the signature set $\{f, g\}$ is not vanishing.*

Proof. Let $\text{arity}(f) = n$ and $\text{rd}^+(f) = d$, so $2d < n$. Since f is not identically 0, we have $d \geq 0$. By Lemma 7.8, $\hat{f} = (Z^{-1})^{\otimes n} f = [\hat{f}_0, \ldots, \hat{f}_d, 0, \ldots, 0]$, where $\hat{f}_d \neq 0$. By forming d self-loops on f we get a nonzero multiple of $Z^{\otimes(n-2d)}[\hat{f}_d, 0, \ldots, 0]$, which is a nonzero multiple of $h = Z^{\otimes(n-2d)}[1, 0]^{\otimes(n-2d)}$. Similarly, from g of arity n' and $\text{rd}^-(g) = d'$, where $2d' < n'$, we can get a nonzero multiple of $h' = Z^{\otimes(n'-2d')}[0, 1]^{\otimes(n'-2d')}$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Then form a bipartite signature grid with $n' - 2d'$ vertices on one side, each assigned h and $n - 2d$ vertices on the other side, each assigned h'. Connect edges between the two sides arbitrarily as long as it is a 1-1 correspondence. Under the Z-transformation this is equivalent to connecting $n' - 2d'$ copies of $[1, 0]^{\otimes(2(n-2d))}$ with $n - 2d$ copies of $[0, 1]^{\otimes(2(n'-2d'))}$, via $(\neq_2)$. The resulting Holant is clearly nonzero. □

Lemma 7.10. *Every symmetric vanishing signature is in $\mathscr{V}^+ \cup \mathscr{V}^-$.*

Proof. Let f be a symmetric vanishing signature. We prove the lemma by induction on n, the arity of f. For $n = 1$, by connecting $f = [f_0, f_1]$ to itself, we have $f_0^2 + f_1^2 = 0$. Then up to a constant factor, we have either $f = [1, \mathrm{i}]$ or $f = [1, -\mathrm{i}]$. The lemma holds.

For $n = 2$, first we do a self-loop. The Holant value is $f_0 + f_2$. Also, we can connect two copies of f, in which case the Holant value is $f_0^2 + 2f_1^2 + f_2^2$. Since f is vanishing, both $f_0 + f_2 = 0$ and $f_0^2 + 2f_1^2 + f_2^2 = 0$. Solving them, we get $f = [1, \mathrm{i}, -1] = [1, \mathrm{i}]^{\otimes 2}$ or $[1, -\mathrm{i}, -1] = [1, -\mathrm{i}]^{\otimes 2}$ up to a constant factor.

Now assume $n > 2$ and the lemma holds for any signature of arity $k < n$. Let $f = Z^{\otimes n}\hat{f} = Z^{\otimes n}[\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_n]$ be a vanishing signature. Let f' be the signature of arity $n - 2$ by taking a self-loop on f. This is equivalent to forming a self-loop via $(\neq_2)$ in the Z basis, namely $f' = Z^{\otimes(n-2)}[\hat{f}_1, \ldots, \hat{f}_{n-1}]$ up to a nonzero scalar.

If f' is identically 0, then $f = Z^{\otimes n}[\hat{f}_0, 0, \ldots, 0, \hat{f}_n]$. We form a graph with two vertices, both assigned f, and n edges between them. The value of this Holant problem can be calculated in Holant$(\neq_2 | \hat{f})$ and is clearly $2\hat{f}_0\hat{f}_n$. As f is vanishing, it must be 0. Hence either $\hat{f}_0 = 0$ in which case $f \in \mathscr{V}^-$, or $\hat{f}_n = 0$ in which case $f \in \mathscr{V}^+$.

Now suppose f' is not identically 0. Since f is vanishing, f' is vanishing as well. By induction, $f' \in \mathscr{V}^+ \cup \mathscr{V}^-$, and without loss of generality we assume $f' \in \mathscr{V}^+$. By Lemma 7.8, $f' = Z^{\otimes(n-2)}[\hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0]$, where $0 \le d - 1 = \mathrm{rd}^+(f') < (n-2)/2$, and $\hat{f}_d \neq 0$. Then $f = Z^{\otimes n}[\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0, \hat{f}_n]$, and we only need to prove $\hat{f}_n = 0$. Suppose otherwise.

By forming $d - 1$ self-loops on f', we get $Z^{\otimes(n-2d)}[1, 0]^{\otimes(n-2d)}$, or $[1, \mathrm{i}]^{\otimes t}$, up to a nonzero constant, where $t = n - 2d$. This construction is valid because $n > 2d$ and $d - 1 \ge 0$.

Define $g = Z^{\otimes(n-1)}[\hat{f}_1, \ldots, \hat{f}_n]$, which is the signature of connecting f with a single unary signature $[1, \mathrm{i}]$. Note that $[1, \mathrm{i}] = Z[1, 0]$ connected via $(\neq_2)$ to $f = Z^{\otimes n}\hat{f}$ simply removes the $\hat{f}_0$ term in the Z-form. We do not claim that g is constructible from f; however, we show that $\{f, g\}$ is vanishing, in particular $\{f', g\}$ is also vanishing. Let Ω be any signature grid on a graph $G = (V, E)$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

using $\{f, g\}$. We show that the Holant value $\text{Holant}_\Omega = 0$. Suppose g occurs m times in Ω. Define a new signature grid Ω' on a graph G' having $t|V| + m$ vertices as follows. For each i $(1 \le i \le t)$ we make a copy G_i of G on $V_i = \{v_i \mid v \in V\}$, i.e., for each edge $(u, v) \in E$, include the edge (u_i, v_i) in the new graph. Every vertex in $\bigcup_{i=1}^{t} V_i$ is assigned the same signature f in Ω'. Next, for each vertex $v \in V$ assigned g, we introduce an additional vertex v', and assign to v' the degenerate signature $[1, \mathrm{i}]^{\otimes t}$ that we just constructed. Then connect v' to each of $v_1, \ldots, v_t$. This effectively changes the signature at each of these v_i to g. Thus $\text{Holant}_{\Omega'} = (\text{Holant}_\Omega)^t$. Since Ω' is a signature grid on $\{f\}$ and f is vanishing, we have $\text{Holant}_{\Omega'} = 0$. Hence $\text{Holant}_\Omega = 0$.

We have shown that $\{f', g\}$ is vanishing; in particular g is vanishing. By induction $g \in \mathscr{V}^+ \cup \mathscr{V}^-$. We had assumed $\hat{f}_n \neq 0$, and so both f' and g are not identically 0. Lemma 7.9 applies, and we get $g \in \mathscr{V}^+$, since $f' \in \mathscr{V}^+$. Hence $\hat{f}_n = 0$ by the Z-form of g, and thus $f \in \mathscr{V}^+$ Lemma 7.8. □

Combining Lemmas 7.7, 7.9, and 7.10, we obtain the following theorem that characterizes all symmetric vanishing signature sets.

Theorem 7.11. *Let $\mathcal{F}$ be a set of symmetric signatures. Then $\mathcal{F}$ is vanishing if and only if $\mathcal{F} \subseteq \mathscr{V}^+$ or $\mathcal{F} \subseteq \mathscr{V}^-$.*

To finish this section, we prove some useful properties regarding vanishing degrees and recurrence degrees. For two symmetric signatures f and g such that $\text{arity}(f) \ge \text{arity}(g)$, let

$$\langle f, g\rangle = \langle g, f\rangle$$

denote the signature that results after connecting all edges of g to f. (If $\text{arity}(f) = \text{arity}(g)$, then $\langle f, g\rangle$ is a constant, which is considered as a signature of arity 0.)

Lemma 7.12. *For $\sigma \in \{+, -\}$, suppose symmetric signatures f and g satisfy $\text{vd}^\sigma(g) = 0$ and $\text{arity}(f) - \text{arity}(g) \ge \text{rd}^\sigma(f)$. Then $\text{rd}^\sigma(\langle f, g\rangle) = \text{rd}^\sigma(f)$.*

Proof. We consider $\sigma = +$ since the case $\sigma = -$ is similar. Let $\text{arity}(f) = n$, $\text{arity}(g) = m$, and $\text{rd}^+(f) = d$. Denote the signature $\langle f, g\rangle$ by h.

If $d = -1$, then f is identically 0 and so is h. Hence $\text{rd}^+(h) = -1$.

Suppose $d \ge 0$. By Lemma 7.8, $\text{vd}^\sigma(g) = 0$ is equivalent to $\hat{g}_m \neq 0$ in $g = Z^{\otimes m}\hat{g} = Z^{\otimes m}[\hat{g}_0, \ldots, \hat{g}_m]$. We also have $f = Z^{\otimes n}[\hat{f}_0, \ldots, \hat{f}_d, 0, \ldots, 0]$, with $\hat{f}_d \neq 0$. Then h has the form $Z^{\otimes(n-m)}[\hat{h}_0, \ldots, \hat{h}_{n-m}]$ of arity $n - m \ge d$, where

$$\hat{h}_i = \sum_{j=0}^{m} \binom{m}{j} \hat{f}_{i+j}\hat{g}_{m-j}, \quad (0 \le i \le n - m).$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

For $i > d$, all terms in the sum are 0 because $\hat{f}_{i+j} = 0$. For $i = d$, this is a nonzero sum because for $j = 0$, $f_{i+j} = f_d \neq 0$ exists, and so does $\hat{g}_m \neq 0$. However, all other terms for $j > 0$ in the sum for $\hat{h}_d$ are 0. Hence $\mathrm{rd}^+(h) = d$, by Lemma 7.8. □

Lemma 7.13. *For $\sigma \in \{+, -\}$, let f be a nonzero symmetric signature and suppose that f' is obtained from f by a self-loop. If $\mathrm{vd}^\sigma(f) > 0$, then $\mathrm{vd}^\sigma(f) - \mathrm{vd}^\sigma(f') = \mathrm{rd}^\sigma(f) - \mathrm{rd}^\sigma(f') = 1$.*

Proof. By Lemma 7.8, $\mathrm{vd}^+(f) > 0$ iff we have $\hat{f}_n = 0$ in the expression $f = Z^{\otimes n}[\hat{f}_0, \ldots, \hat{f}_n]$. The corresponding expression for f' is $Z^{\otimes n}[\hat{f}_1, \ldots, \hat{f}_{n-1}]$, which has one fewer trailing 0 after the last nonzero entry $\hat{f}_d$, where $d = \mathrm{rd}^+(f)$. Hence $\mathrm{vd}^+(f) - \mathrm{vd}^+(f') = 1$. Since $\mathrm{rd}^+(f) = n - \mathrm{vd}^+(f)$ and $\mathrm{rd}^+(f') = (n-2) - \mathrm{vd}^+(f')$, it follows that $\mathrm{rd}^+(f) - \mathrm{rd}^+(f') = 1$. □

Moreover, the set of vanishing signatures is closed under orthogonal transformations. This is because under any orthogonal transformation, the unary signatures $[1, \mathrm{i}]$ and $[1, -\mathrm{i}]$ are either invariant or transformed into each other, up to a nonzero multiple. Then considering the symmetrized form of any signature, we have the following lemma.

Lemma 7.14. *For a symmetric signature f of arity n, $\sigma \in \{+, -\}$, and an orthogonal matrix $T \in \mathbb{C}^{2\times 2}$, either $\mathrm{vd}^\sigma(f) = \mathrm{vd}^\sigma(T^{\otimes n} f)$ or $\mathrm{vd}^\sigma(f) = \mathrm{vd}^{-\sigma}(T^{\otimes n} f)$.*

7.1.3 Characterization by Recurrence Relations

Now we give another characterization of vanishing signatures.

Definition 7.15. An arity n symmetric signature of the form $f = [f_0, f_1, \ldots, f_n]$ is in $\mathscr{R}_t^+$ for a nonnegative integer $t \geq 0$ if

1. $t > n$; or
2. For any $0 \leq k \leq n - t$, $f_k, \ldots, f_{k+t}$ satisfy the recurrence relation of order t

$$\binom{t}{t}\mathrm{i}^t f_{k+t} + \binom{t}{t-1}\mathrm{i}^{t-1} f_{k+t-1} + \cdots + \binom{t}{0}\mathrm{i}^0 f_k = 0. \tag{7.4}$$

We define $\mathscr{R}_t^-$ similarly but with $-\mathrm{i}$ in place of i in (7.4).

We note that some particular categories of tractable cases in previous dichotomies are in $\mathscr{R}_2^\pm$. This includes case 3 of Theorem 6.3, the case of $\mathrm{I}(1, \pm 2\mathrm{i})$ in Class $\mathfrak{B}$ of Theorem 2.12, and the subcase of type $\mathrm{I}(1, \pm 2\mathrm{i})$ in case 1 of Theorem 6.12.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

It is easy to see that $\mathscr{R}_0^+ = \mathscr{R}_0^-$ is the set of all zero signatures (of any arity). Also, for $\sigma \in \{+, -\}$, we have $\mathscr{R}_t^\sigma \subseteq \mathscr{R}_{t'}^\sigma$ when $t \le t'$. By definition, if $\text{arity}(f) = n$ then $f \in \mathscr{R}_{n+1}^\sigma$.

Let $f = [f_0, f_1, \ldots, f_n] \in \mathscr{R}_t^+$ with $0 < t \le n$. Then the characteristic polynomial of its recurrence relation is $(1 + x\mathrm{i})^t$. Thus there exists a polynomial $p(x)$ of degree $\le t - 1$ such that $f_k = \mathrm{i}^k p(k)$, for $0 \le k \le n$. This statement extends to $\mathscr{R}_{n+1}^+$ since a polynomial of degree n can interpolate any set of $n + 1$ values. Furthermore, such an expression is unique. If there are two polynomials $p(x)$ and $q(x)$, both of degree $\le n$, such that $f_k = \mathrm{i}^k p(k) = \mathrm{i}^k q(k)$ for $0 \le k \le n$, then $p(x)$ and $q(x)$ must be the same polynomial. Now suppose $f_k = \mathrm{i}^k p(k)$ $(0 \le k \le n)$ for some polynomial p of degree $\le t - 1$, where $0 < t \le n$. Then f satisfies the recurrence (7.4) of order t. Hence $f \in \mathscr{R}_t^+$.

Thus $f \in \mathscr{R}_t^+$ iff there exists a polynomial $p(x)$ of degree $\le t - 1$ such that $f_k = \mathrm{i}^k p(k)$ $(0 \le k \le n)$, for all $1 \le t \le n + 1$. For $\mathscr{R}_t^-$, just replace i by $-\mathrm{i}$.

It follows that the set of all symmetric signatures of arity n in $\mathscr{R}_t^\sigma$ forms a vector space over $\mathbb{C}$ of dimension t, where $\sigma \in \{+, -\}$.

Lemma 7.16. *Let $f = [f_0, \ldots, f_n]$ be a symmetric signature of arity n, not identically 0. Then for any nonnegative integer $0 \le d < n$ and $\sigma \in \{+, -\}$, the following are equivalent:*

1. $\mathrm{rd}^\sigma(f) \le d$
2. $f \in \mathscr{R}_{d+1}^\sigma$.

Proof. We prove the case $\sigma = +$; the case $\sigma = -$ is symmetric.

By Lemma 7.8, the set V_n^d of symmetric signatures of arity n with $\mathrm{rd}^+(f) \le d$ consists of all expressions of the form $f = Z^{\otimes n}[\hat{f}_0, \ldots, \hat{f}_d, 0, \ldots, 0]$. Hence it forms a vector space over $\mathbb{C}$.

Let $\mathrm{Ex}_n^t = [0, \ldots, 0, 1, 0, \ldots, 0]$ be the Exact-t function of arity n, where $0 \le t \le n$, and the unique nonzero entry 1 appears at weight t. For $0 \le t \le d$, these are linearly independent, and since the linear transformation on the space of symmetric signatures of arity n induced by $Z^{\otimes n}$ is invertible, the image $\{Z^{\otimes n}\mathrm{Ex}_n^t \mid 0 \le t \le d\}$ forms a basis of V_n^d. Thus V_n^d has dimension $d + 1$, the same as that of $\mathscr{R}_{d+1}^+$. We only need to prove that each $Z^{\otimes n}\mathrm{Ex}_n^t \in \mathscr{R}_{d+1}^+$, for $0 \le t \le d$.

We prove this by showing that $Z^{\otimes n}\mathrm{Ex}_n^t$ can be expressed as $p(k)\mathrm{i}^k$, where p is a polynomial of degree $\le t$.

We can express Ex_n^t as a sum of $\binom{n}{t}$ terms of the form $v_1 \otimes v_2 \otimes \ldots \otimes v_n$ where $v_i \in \mathbb{C}^2$ $(1 \le i \le n)$ and exactly t of them are $[0, 1]$ and $n - t$ of them are $[1, 0]$. Then up to a nonzero scalar, $Z^{\otimes n}\mathrm{Ex}_n^t$ is a sum of $\binom{n}{t}$ terms of the form $u_1 \otimes u_2 \otimes \ldots \otimes u_n$ where $u_i \in \mathbb{C}^2$ and exactly t of them are $[1, -\mathrm{i}]$ and $n - t$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

of them are [1, i]. The signature entry of $Z^{\otimes n}\mathrm{Ex}_n^t$ at weight k ($0 \le k \le n$) is

$$\sum_s \binom{k}{s}\binom{n-k}{t-s}(-\mathrm{i})^s \mathrm{i}^{k-s}. \tag{7.5}$$

This is because the $\binom{n}{t}$ terms of $Z^{\otimes n}\mathrm{Ex}_n^t$ are each uniquely determined by the locations of the t factors $[1, -\mathrm{i}]$. On input $(\underbrace{1, \ldots, 1}_{k}, \underbrace{0, \ldots, 0}_{n-k})$, a term with some s of the first k bit positions and $t-s$ of the last $n-k$ bit positions occupied by $[1, -\mathrm{i}]$ contributes $(-\mathrm{i})^s \mathrm{i}^{k-s}$. The sum in (7.5) is over the range of s satisfying both $0 \le s \le k$ and $0 \le t-s \le n-k$, i.e., $\max\{0, k+t-n\} \le s \le \min\{k, t\}$. This is justified by the combinatorial argument above. However, we will rewrite the sum (7.5) as over the range $0 \le s \le t$, and note that if $s > k$ then the first factor $\binom{k}{s} = 0$, and if $s < k+t-n$ then the second factor $\binom{n-k}{t-s} = 0$. Hence (7.5) can be written as

$$\mathrm{i}^k \sum_{0 \le s \le t} \binom{k}{s}\binom{n-k}{t-s}(-1)^s. \tag{7.6}$$

It is easy to see that, for $0 \le s \le t$, each $\binom{k}{s}\binom{n-k}{t-s}$ is a polynomial in k of degree t. The conclusion follows. □

Note a slight subtlety in the proof of Lemma 7.16. The combinatorial justification for the sum over s results in a sum whose range depends on k. In order to be rigorous in arguing that the sum over s is a polynomial in k of degree $\le t$, we must replace the range of this sum to be independent of k.

Corollary 7.17. *For a nonzero symmetric signature f of arity n, $\mathrm{rd}^+(f) = t$, if and only if $f \in \mathscr{R}_{t+1}^+ - \mathscr{R}_t^+$. Similarly $\mathrm{rd}^-(f) = t$ if and only if $f \in \mathscr{R}_{t+1}^- - \mathscr{R}_t^-$.*

Remark: Membership in $\mathscr{R}_t^\pm$ is easily decidable by (7.4). Therefore by Lemma 7.16, for any symmetric signature f, its recurrence degree $\mathrm{rd}^\pm(f)$ can be computed in polynomial time in the size of its symmetric signature. This is one advantage of using $\mathscr{R}_t^\pm$ compared to the concept of vanishing degree $\mathrm{vd}^\pm$, while conceptually the latter relates more directly to the notion of vanishing. Lemmas 7.8 and 7.16 show that the two concepts are equivalent. By Lemma 7.16 membership in $\mathscr{V}^\pm$ is decidable in polynomial time in the size of a symmetric signature. This also follows from Lemma 7.8 because the Z-transformation is also computable for a symmetric signature in polynomial time in the size of a symmetric signature. However, checking for concrete signatures by the recurrence (7.4) in Definition 7.15 is more straightforward.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

When the results in this chapter were first obtained, we relied heavily in the concept of recurrence degree, which was defined not by Definition 7.5, but in terms of (7.4) and Corollary 7.17 (made into a definition). The reorganization of the proofs in this chapter emphasizes the Z-form as the primary tool. This in hindsight simplifies many proof details. We thank Artem Govorov and Heng Guo especially for discussions that led to some new and simplified proofs; in particular the simplified proof of Lemma 7.16 is due to Govorov, which we present here with his permission.

Note that although we call it the recurrence degree, it refers to a particular recurrence relation. For any nonzero symmetric signature f, by the uniqueness of the representing polynomial $p(x)$, it follows that $\mathrm{rd}^{\sigma}(f) = t$ iff $\deg(p) = t$, where $0 \le t \le n$. We remark that $\mathrm{rd}^{\sigma}(f)$ is the maximum integer t such that f does *not* belong to $\mathscr{R}_t^{\sigma}$. Also, for an arity n signature f, $\mathrm{rd}^{\sigma}(f) = n$ if and only if f does not satisfy any recurrence relation (7.4) of order $t \le n$ for $\sigma \in \{+, -\}$.

7.2 Theorem Statement and Proof of Tractability

For a class $\mathscr{C}$ of signatures we define the notion of $\mathscr{C}$-transformable in the context of Holant problems. We will particularly use this notion for $\mathscr{C} = \mathscr{A}$ or $\mathscr{C} = \mathscr{P}$.

Definition 7.18. A signature f (respectively a signature set $\mathcal{F}$) is $\mathscr{C}$-transformable if there exists a holographic transformation T such that $f \in T\mathscr{C}$ (respectively the set $\mathcal{F} \subseteq T\mathscr{C}$) and $[1, 0, 1]T^{\otimes 2} \in \mathscr{C}$.

Now we can formally state the dichotomy theorem for Holant problems on an arbitrary set of symmetric signatures.

Theorem 7.19. *Let $\mathcal{F}$ be any set of symmetric, complex-valued signatures in Boolean variables. Then* Holant($\mathcal{F}$) *is* #P*-hard unless $\mathcal{F}$ satisfies one of the following conditions, in which case the problem is in P:*

1. *All nondegenerate signatures in $\mathcal{F}$ are of arity ≤ 2.*
2. *$\mathcal{F}$ is $\mathscr{A}$-transformable.*
3. *$\mathcal{F}$ is $\mathscr{P}$-transformable.*
4. *$\mathcal{F} \subseteq \mathscr{V}^{\sigma} \cup \{f \in \mathscr{R}_2^{\sigma} \mid \mathrm{arity}(f) = 2\}$ for some $\sigma \in \{+, -\}$.*
5. *All nondegenerate signatures in $\mathcal{F}$ are in $\mathscr{R}_2^{\sigma}$ for some $\sigma \in \{+, -\}$.*

Note that any signature in $\mathscr{R}_2^{\sigma}$ having arity ≥ 3 is a vanishing signature. Thus all nondegenerate signatures of arity ≥ 3 in case 5 are vanishing. Although both

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

cases 4 and 5 are largely concerned with vanishing signatures, these two cases differ. In case 4, all signatures in $\mathcal{F}$ including unary signatures, but excluding binary signatures, must be vanishing of a single type σ; the binary signatures are only required to be in $\mathscr{R}_2^\sigma$. In contrast, case 5 has no requirement placed on degenerate signatures which include all unary signatures. Instead case 5 requires that all nondegenerate signatures of arity ≥ 2 belong to $\mathscr{R}_2^\sigma$. Note that $f \in \mathscr{R}_2^\sigma$ is equivalent to $\mathrm{rd}^\sigma \leq 1$, and therefore for any signature f of arity ≥ 3 this implies that $f \in \mathscr{V}^\sigma$. For large arity $n \geq 3$, this is a strong form of vanishing, meaning that it has a large vanishing degree $\mathrm{vd}^\sigma \geq n - 1$.

Case 5 is tractable due to Fibonacci gates. In fact every nondegenerate signature in $\mathscr{R}_2^\sigma$ is a generalized Fibonacci signature with the parameter $\lambda = \sigma 2\mathrm{i}$ in equation (2.5), and after replacing all degenerate signatures with the corresponding unary signatures we can invoke Holant* tractability. However, in the following we present a unified proof of tractability based on vanishing signatures, which leads to an alternative algorithm for this case.

Proof of Tractability of Theorem 7.19. For any signature grid Ω, Holant_Ω is the product of the Holant on each connected component, so we only need to compute over connected components.

For case 1, after decomposing all degenerate signatures into unary ones, a connected component of the graph is either a path or a cycle and the Holant can be computed using matrix product and trace. Cases 2 and 3 are tractable because, after a particular holographic transformation, their instances are tractable instances of #CSP($\mathcal{F}$), by the dichotomy theorem for #CSP, Theorem 3.7. For case 4, any binary signature $g \in \mathscr{R}_2^\sigma$ has $\mathrm{rd}^\sigma(g) \leq 1$, and thus $\mathrm{vd}^\sigma(g) \geq 1 = \mathrm{arity}(g)/2$. Any signature $f \in \mathscr{V}^\sigma$ has $\mathrm{vd}^\sigma(f) > \mathrm{arity}(f)/2$. If $\mathcal{F}$ contains a signature f of arity ≥ 3, then it must belong to $\mathscr{V}^\sigma$. Then by the combinatorial view, more than half of the unary signatures are $[1, \sigma\mathrm{i}]$, so Holant_Ω vanishes. On the other hand, if the arity of every signature in $\mathcal{F}$ is ≤ 2, then we have reduced to tractable case 1.

Now consider case 5. First, replace each degenerate signature by its unary signature tensor factors. Any remaining signature of arity ≥ 2 is nondegenerate and thus belongs to $\mathscr{R}_2^\sigma$. Then recursively absorb any unary signature into its neighboring signature. If it is connected to another unary signature, then this produces a global constant factor. If it is connected to a binary signature, then this creates another unary signature. We observe that if $f \in \mathscr{R}_2^\sigma$ has $\mathrm{arity}(f) \geq 2$, then for any unary signature u, after connecting f to u, the signature $\langle f, u\rangle$ still belongs to $\mathscr{R}_2^\sigma$. Hence after recursively absorbing all unary signatures in the above process, we have a signature grid where all signatures belong to $\mathscr{R}_2^\sigma$. Any

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

signature f that remains and has arity ≥ 3 belongs to $\mathscr{V}^{\sigma}$ since $\mathrm{rd}^{\sigma}(f) \leq 1$ and thus $\mathrm{vd}^{\sigma}(f) \geq \mathrm{arity}(f) - 1 > \mathrm{arity}(f)/2$. Thus we have reduced to tractable case 4. □

7.3 A Sample of Problems

We illustrate the scope of Theorem 7.19 by several concrete problems. Note that some problems are naturally expressed with real weights, but they are linked to other problems that use complex weights. Sometimes the inherent link between two real-weighted problems is provided by a transformation through $\mathbb{C}$.

Problem: #VertexCover

Input: An undirected graph G

Output: The number of vertex covers in G

This classic problem is most naturally expressed as a spin system with edge function being the Boolean Or function, or as a counting CSP problem #CSP([0, 1, 1]). As a real-weighted bipartite Holant problem it is Holant([0, 1, 1] | $\mathcal{EQ}$), where $\mathcal{EQ} = \{(=_k) \mid k \geq 1\}$ is the set of all Equalities. A vertex assigned an equality signature forces all its incident edges to be assigned the same value; this is equivalent to assigning a value to the vertex. The degree 2 vertices assigned the binary Or = [0, 1, 1] should be thought of as an edge between its neighboring vertices. These edge-like vertices force at least one of its neighbors to be selected. The number of assignments satisfying these requirements is exactly the number of vertex covers.

To apply Theorem 7.19, we perform a holographic transformation by $T = \begin{bmatrix} 0 & -\mathrm{i} \\ 1 & \mathrm{i} \end{bmatrix}$. Note that $T^{-1} = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & 0 \end{bmatrix}$, and

$$\begin{aligned}[0, 1, 1] = (0\ 1\ 1\ 1) &= [1, 1]^{\otimes 2} + [\mathrm{i}, 0]^{\otimes 2} = \left\{[1, 0]^{\otimes 2} + [0, 1]^{\otimes 2}\right\} (T^{-1})^{\otimes 2} \\ &= (1\ 0\ 0\ 1)(T^{-1})^{\otimes 2} = (=_2)(T^{-1})^{\otimes 2}.\end{aligned}$$

Thus, a holographic transformation by T yields

$$\begin{aligned}\mathrm{Holant}([0, 1, 1] \mid \mathcal{EQ}) &\equiv_{\mathrm{T}} \mathrm{Holant}([0, 1, 1]T^{\otimes 2} \mid T^{-1}\mathcal{EQ}) \\ &\equiv_{\mathrm{T}} \mathrm{Holant}(=_2 \mid T^{-1}\mathcal{EQ}) \\ &\equiv_{\mathrm{T}} \mathrm{Holant}(T^{-1}\mathcal{EQ}).\end{aligned}$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

The EQUALITY signature $(=_k)$ of arity k in $\mathcal{EQ}$ is transformed by T^{-1} to

$$\begin{aligned} f_{(k)} &= (T^{-1})^{\otimes k}(=_k) \\ &= \begin{bmatrix} 1 & 1 \\ \mathrm{i} & 0 \end{bmatrix}^{\otimes k} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes k} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}^{\otimes k} \right\} \\ &= \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}^{\otimes k} + \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes k} \\ &= [2, \mathrm{i}, -1, -\mathrm{i}, 1, \mathrm{i}, -1, -\mathrm{i}, 1, \mathrm{i}, \ldots] \end{aligned}$$

of arity k. By Theorem 7.19, Holant$(T^{-1}\mathcal{EQ})$ is #P-hard. Indeed, even Holant$(f_{(k)})$, the restriction of this problem to k-regular graphs is #P-hard, for $k \geq 3$. This can also be obtained more directly by Theorem 7.52, a dichotomy for a single signature, which is a special case of Theorem 7.19.

Problem: #λ-VERTEXCOVER

Input: An undirected graph G

Output: $\sum_{C \in \mathcal{C}(G)} \lambda^{e(C)}$, where $\mathcal{C}(G)$ denotes the set of all vertex covers of G, and $e(C)$ is the number of edges with both endpoints in the vertex cover C

Our dichotomy also easily handles this edge-weighted vertex cover problem that is denoted by Holant$([0, 1, \lambda] \mid \mathcal{EQ})$. Suppose $\lambda \neq 0$. On regular graphs, this problem is equivalent to the so-called *hardcore gas model*, which is the vertex-weighted problem denoted by Holant$([1, 1, 0] \mid \mathcal{F})$, where $\mathcal{F}$ consists of signatures of the form $[1, 0, \ldots, 0, \mu]$. By flipping 0 and 1, this is the same as Holant$([0, 1, 1] \mid \mathcal{F}')$ with $\mathcal{F}'$ containing $[\mu, 0, \ldots, 0, 1]$. For k-regular graphs, the following diagonal transformation $T = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\lambda} \end{bmatrix}$ shows this equivalence, where $\lambda = 1/\mu^{1/k}$;

$$\begin{aligned} \text{Holant}([0, 1, \lambda] \mid =_k) &\equiv_{\mathrm{T}} \text{Holant}([0, 1, \lambda]T^{\otimes 2} \mid (T^{-1})^{\otimes k}(=_k)) \\ &\equiv_{\mathrm{T}} \text{Holant}(\tfrac{1}{\lambda}[0, 1, 1] \mid [1, 0, \ldots, 0, \lambda^k]) \\ &\equiv_{\mathrm{T}} \text{Holant}([0, 1, 1] \mid [\mu, 0, \ldots, 0, 1]). \end{aligned}$$

This problem, denoted by #k-λ-VERTEXCOVER, is #P-hard for $k \geq 3$. To see this, apply the holographic transformation $T = \begin{bmatrix} 0 & -\mathrm{i}\lambda \\ 1 & \mathrm{i} \end{bmatrix}$ to the edge-weighted form of the problem Holant$([0, 1, \lambda] \mid =_k)$. Then $[0, 1, \lambda]$ is transformed to $[0, 1, \lambda]T^{\otimes 2} = \lambda(=_2)$ and $(=_k)$ is transformed to

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

$(T^{-1})^{\otimes k}(=_k) = g_{(\lambda,k)} = \frac{1}{\lambda^k}[\lambda^k + 1, \mathrm{i}, -1, -\mathrm{i}, 1, \dots]$, where $T^{-1} = \frac{1}{\lambda}\begin{bmatrix} 1 & \lambda \\ \mathrm{i} & 0 \end{bmatrix}$. Since Holant$(g_{(\lambda,k)})$ is #P-hard by Theorem 7.19, or more directly by Theorem 7.52, we conclude that #k-λ-VertexCover is also #P-hard.

If $\lambda = 0$, then the above problem is Holant$([0, 1, 0] \mid \mathcal{EQ})$, which is tractable. However, the transformation T above is singular in this case. We can in fact apply another transformation $T' = \begin{bmatrix} 1 - \frac{\lambda}{2} & -\left(1 + \frac{\lambda}{2}\right)\mathrm{i} \\ 1 & \mathrm{i} \end{bmatrix}$ such that it transforms the problem Holant$([0, 1, \lambda] \mid =_k)$ into Holant$(h_{(\lambda,k)})$ for some $h_{(\lambda,k)}$ regardless of whether $\lambda = 0$ or not. Then by applying Theorem 7.52, we reach the same conclusion that #λ-VertexCover is #P-hard on k-regular graphs when $\lambda \neq 0$. We note that when $\lambda = 0$, $T' = \begin{bmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{bmatrix} = \sqrt{2}Z^{-1}$, where $Z = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$.

We now consider some orientation problems.

Problem: #NoSinkOrientation

Input: An undirected graph G

Output: The number of orientations of G such that each vertex has at least one outgoing edge.

This problem is denoted by Holant$([0, 1, 0] \mid \mathcal{F})$, where $\mathcal{F}$ consists of $f_{(k)} = [0, 1, \dots, 1, 1]$ for any arity k. Each degree 2 vertex on the left side of the bipartite graph must have different values assigned to its incident edges. We associate an oriented edge between the neighbors of such vertices with the head on the side assigned 0 and the tail on the side assigned 1. This problem is #P-hard even over k-regular graphs provided $k \geq 3$. Just as with the bipartite form of the vertex cover problem, we do a holographic transformation to apply Theorem 7.19. This time, we pick $T = \frac{1}{2}\begin{bmatrix} 1 & -\mathrm{i} \\ 1 & \mathrm{i} \end{bmatrix} = \frac{1}{\sqrt{2}}Z^{-1}$, with $T^{-1} = \sqrt{2}Z = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$ and get

$$\begin{aligned} \text{Holant}([0, 1, 0] \mid f_{(k)}) &\equiv_{\mathrm{T}} \text{Holant}([0, 1, 0]T^{\otimes 2} \mid (T^{-1})^{\otimes k} f_{(k)}) \\ &\equiv_{\mathrm{T}} \text{Holant}(\tfrac{1}{2}[1, 0, 1] \mid \hat{f}_{(k)}) \\ &\equiv_{\mathrm{T}} \text{Holant}(\hat{f}_{(k)}), \end{aligned}$$

where $\hat{f}_{(k)} = [2^k - 1, -\mathrm{i}, 1, \mathrm{i}, -1, \dots]$. The signature $-\hat{f}_{(k)} = [1 - 2^k, \mathrm{i}, -1, -\mathrm{i}, 1, \dots]$ is actually a special case of the #k-λ-VertexCover problem with $\lambda = 2e^{\pi \mathrm{i}/k}$. Therefore, this problem is #P-hard. However, if we consider

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

this problem modulo 2^k, $\hat{f}_{(k)}$ becomes $[-1, -\mathrm{i}, 1, \mathrm{i}, -1, \dots]$, and it belongs to one of the tractable cases in Theorem 7.19. Thus, #NoSinkOrientation is tractable modulo 2^t, where t is the minimum degree of the input graph.

Problem: #NoSinkNoSourceOrientation

Input: An undirected graph G

Output: The number of orientations of G such that each vertex has at least one incoming and one outgoing edge

This problem is denoted by Holant($[0, 1, 0] \mid \mathcal{F}$), where $\mathcal{F}$ consists of $f_{(k)} = [0, 1, \dots, 1, 0]$ for any arity k. This problem is also #P-hard on k-regular graphs for $k \geq 3$. We pick the same T as in the previous problem and get

$$\begin{aligned}\text{Holant}([0, 1, 0] \mid f_{(k)}) &\equiv_{\mathrm{T}} \text{Holant}([0, 1, 0]T^{\otimes 2} \mid (T^{-1})^{\otimes k} f_{(k)}) \\ &\equiv_{\mathrm{T}} \text{Holant}(\tfrac{1}{2}[1, 0, 1] \mid \hat{f}_{(k)}) \\ &\equiv_{\mathrm{T}} \text{Holant}(\hat{f}_{(k)}),\end{aligned}$$

where $\hat{f}_{(k)} = [2^k - 2, 0, 2, 0, -2, \dots]$. Here we transform from one real-weighted Holant problem to another real-weighted Holant problem via a complex-weighted transformation. The #P-hardness follows from Theorem 7.52. Like the previous problem, this problem is tractable modulo 2^t, where t is the minimum degree of the input graph.

Theorem 7.19 is applicable to a set of signatures when different vertices may have different constraints.

Problem: #One-In-Or-One-Out-Orientation

Input: An undirected graph G with each vertex labeled "One-In" or "One-Out"

Output: The number of orientations of G such that each vertex has exactly one incoming or exactly one outgoing edge as specified by its label

This problem is denoted by Holant($[0, 1, 0] \mid \mathcal{F}$), where the set $\mathcal{F}$ consists of signatures of the form $f = [0, 1, 0, \dots, 0]$ and $g = [0, \dots, 0, 1, 0]$. Once again, it is #P-hard on k-regular graphs for $k \geq 3$. We apply the same transformation as in the above two orientation problems. The result is Holant($\{\hat{f}, \hat{g}\}$), where $\hat{f} = [k, (k-2)\mathrm{i}, -(k-4), \dots]$ and $\hat{g} = [k, -(k-2)\mathrm{i}, -(k-4), \dots]$ of arity k. In fact, the entries of $\hat{f}$ satisfy a second-order recurrence relation with characteristic polynomial $(x - \mathrm{i})^2$ while the entries of $\hat{g}$ satisfy one with characteristic polynomial $(x + \mathrm{i})^2$. Thus $\hat{f} \in \mathscr{R}_2^+$ and $\hat{g} \in \mathscr{R}_2^-$. The hardness follows from Theorem 7.19. See also Lemma 7.30. However, the restriction of this problem to planar graphs is tractable since these signatures are all matchgate

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

signatures, by Theorem 4.11. If we consider only one signature, then both Holant($\hat{f}$) and Holant($\hat{g}$) are tractable. This is also true for a set of signatures with different arities but of one type, either of the type $\hat{f}$ or of the type $\hat{g}$. This is case 5 of Theorem 7.19. One can also observe that the problem Holant($\hat{f}$) is equivalent to the problem Holant([0, 1, 0] | [0, 1, 0, . . . , 0]), which is always 0 provided $k \geq 3$ by a simple counting argument. Similarly for Holant($\hat{g}$). For $k \geq 3$, $\hat{f} \in \mathscr{V}^+$ and $\hat{g} \in \mathscr{V}^-$ are vanishing signatures. However, combining two such vanishing signatures of the opposite type $\mathscr{V}^+$ and $\mathscr{V}^-$, the problem Holant($\{\hat{f}, \hat{g}\}$) is #P-hard.

One sufficient condition for a signature of arity ≥ 3 to be vanishing is that its entries satisfy a second-order recurrence relation with characteristic polynomial $(x \pm \mathrm{i})^2$. If the entries of a nondegenerate signature f satisfy a second-order recurrence relation with characteristic polynomial $(x - \mu)^2$ for $\mu \neq \pm\mathrm{i}$, then there exists an orthogonal holographic transformation such that f is transformed into a weighted matching signature.

Problem: #λ-WeightedMatching

Input: An undirected graph G

Output: $\sum_{M \in \mathcal{M}(G)} \lambda^{v(M)}$, where $\mathcal{M}(G)$ is the set of all matchings in G and $v(M)$ is the number of unmatched vertices in the matching M

The Holant expression of this problem is Holant($\mathcal{F}$), where $\mathcal{F}$ consists of signatures of the form $[\lambda, 1, 0, \ldots, 0]$. When $\lambda = 0$, this problem counts perfect matchings, which is #P-hard even for bipartite graphs [Val79a] but tractable over planar graphs by Kasteleyn's algorithms (Theorem 4.7). When $\lambda = 1$, this problem counts general matchings. Vadhan [Vad01] proved that counting general matchings is #P-hard over k-regular graphs for $k \geq 5$, but left open the question for $k = 4$. Theorem 7.19, or more directly Theorem 7.52, shows that #λ-WeightedMatching is #P-hard, for any weight λ and on any k-regular graphs for $k \geq 3$.

7.4 Outline of Hardness Proof for Theorem 7.19

The hardness proof of Theorem 7.19 will take up the remainder of this chapter. Before proving Theorem 7.19 which deals with a set of signatures $\mathcal{F}$, we first prove a special case Theorem 7.52, where $\mathcal{F}$ consists of a single signature. The proof is by induction on the arity of the signature. The induction is done by taking a self-loop, which causes the arity to go down by 2. Thus, we need two base cases, a dichotomy for an arity 3 signature, Theorem 6.3, and a dichotomy

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

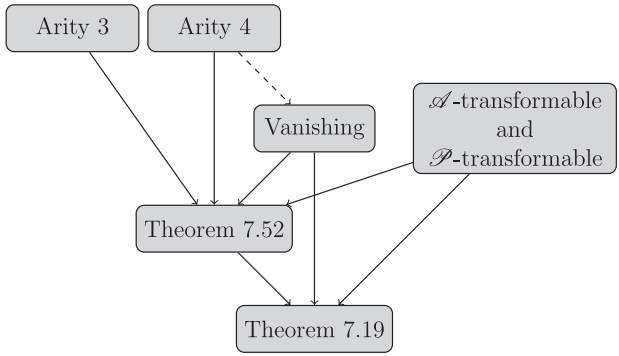


Figure 7.1. Dependency graph of key hardness results for our main dichotomy, Theorem 7.19. The dashed edge indicates a dependence in terms of techniques rather than the result itself. "Arity 3(4)" stands for the arity 3(4) single signature dichotomy. "Vanishing" ("$\mathscr{A}$-transformable and $\mathscr{P}$-transformable") stands for the lemmas regarding vanishing ($\mathscr{A}$-transformable and $\mathscr{P}$-transformable) signatures. Dependencies on previous dichotomy theorems are not shown.

for an arity 4 signature, Theorem 7.29, which is a crucial ingredient in our proof of Theorem 7.19. The arity 4 dichotomy not only is a base case of the single signature dichotomy Theorem 7.52, but also is utilized several times in the inductive step.

After obtaining the dichotomy for an arity 4 signature, the proof continues by revisiting the vanishing signatures to determine what signatures combine with them to give #P-hardness. When adding unary or binary signatures, the only possible combinations that maintain the tractability of the vanishing signatures are as described in cases 4 and 5 in Theorem 7.19. Moreover, combining two vanishing signatures of the opposite type of arity ≥ 3 implies #P-hardness. The proof of this last statement uses techniques that are similar to those in the proof of the arity 4 dichotomy.

Another ingredient of the proof is to understand the signatures that are $\mathscr{A}$-transformable or $\mathscr{P}$-transformable. We obtain explicit characterizations of these signatures. We use these characterizations to prove dichotomy theorems for any signature set containing an $\mathscr{A}$- or $\mathscr{P}$-transformable signature. Unless every signature in the set is $\mathscr{A}$- or $\mathscr{P}$-transformable, the problem is #P-hard. The proofs of these dichotomy theorems utilize the #CSP^d dichotomy, Theorem 6.16.

The main dichotomy, Theorem 7.19, depends on Theorem 7.52 and the results regarding vanishing signatures as well as $\mathscr{A}$- and $\mathscr{P}$-transformable signatures. Figure 7.1 summarizes the dependencies among these results. The material in the next four sections is technical. It may be helpful to take a glance

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

now of the *Proof of hardness in Theorem 7.19* at the end of Section 7.8. This may provide some understanding for the organization of the material.

7.5 Dichotomy for One Signature of Arity 4

Definition 7.20. A 4×4 matrix is *redundant* if its middle two rows and middle two columns are the same. Denote the set of all redundant 4×4 matrices over $\mathbb{C}$ by $\mathrm{RM}_4(\mathbb{C})$.

Recall the notion of a signature matrix in (6.19) of a signature of arity 4. As an example, the signature matrix in (6.20) of a symmetric arity 4 signature is redundant.

Define a function $\varphi : \mathbb{C}^{4\times 4} \to \mathbb{C}^{3\times 3}$ by

$$\varphi(M) = AMB,$$

where

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Intuitively, the operation φ replaces the middle two columns of M with their sum and then the middle two rows of M with their average. (These two steps commute.) In the other direction, we define the following function $\psi : \mathbb{C}^{3\times 3} \to \mathrm{RM}_4(\mathbb{C})$ by

$$\psi(N) = BNA.$$

Intuitively, the operation ψ duplicates the middle row of N and then splits the middle column evenly into two columns. Notice that $\varphi(\psi(N)) = N$. When restricted to $\mathrm{RM}_4(\mathbb{C})$, φ is an isomorphism between the semigroup of 4×4 redundant matrices and the semigroup of 3×3 matrices, under matrix multiplication, and ψ is its inverse. This can be seen by noticing that

$$AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad BA = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

are the identity elements of their respective semigroups.

Definition 7.21. If g is a signature of arity 4 with a redundant signature matrix M_g, then we define the *compressed signature matrix* of g as $\widetilde{M_g} = \varphi(M_g)$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

If all signatures in an $\mathcal{F}$-gate have even arity, then the $\mathcal{F}$-gate also has even arity. Knowing that binary signatures alone in Holant problems do not produce #P-hardness (see tractable case 1 of Theorem 7.19), with the above constraint in mind, we would like to interpolate other arity 4 signatures using a given arity 4 signature. We are particularly interested in the following signature g with signature matrix

$$M_g = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{7.7}$$

This is the identity element in the semigroup $\mathrm{RM}_4(\mathbb{C})$. Thus, $\widetilde{M_g} = I_3$. Lemma 7.24 will show that the Holant problem with this signature is #P-hard. In Lemma 7.23, we consider when we can interpolate it. This is a key lemma. There are three cases in Lemma 7.23 and one of them requires the following technical lemma.

Lemma 7.22. *Let $M = [B_0\, B_1\, \cdots\, B_t]$ be an $n \times n$ block matrix such that there exists $\lambda \in \mathbb{C}$, for all integers $0 \le k \le t$, block B_k is an $n \times c_k$ matrix for some integer $c_k \ge 0$, and the entry of a nonempty block B_k (for $c_k \ge 1$) at row r and column c is $(B_k)_{rc} = r^{c-1}\lambda^{kr}$, where $1 \le r \le n$, $1 \le c \le c_k$. If λ is nonzero and is not a root of unity, then M is nonsingular.*

Proof. We prove by induction on n. If $n = 1$, then the sole entry is λ^k for some nonnegative integer k. This is nonzero since $\lambda \neq 0$. Assume $n > 1$ and let the leftmost nonempty block be B_j. We divide row r by λ^{jr}, which is allowed since $\lambda \neq 0$. This effectively changes block B_ℓ into a block of the form $B_{\ell-j}$. Thus, we have another matrix of the same form as M but with a nonempty block B_0. To simplify notation, we denote this matrix again by M. The first column of B_0 is all 1's. We subtract row $r-1$ from row r, for r from n down to 2. This gives us a new matrix $M' = [B'_0\, B'_1\, \cdots\, B'_t]$, and $\det M = \det M'$. Then $\det M'$ is the determinant of the $(n-1) \times (n-1)$ submatrix M'' obtained from M' by removing the first row and column. Now we do column operations on M'' to return the blocks to the proper form so that we can apply induction.

For any nonempty block B'_k other than B'_0, we prove by induction on the number of columns in B'_k that B'_k can be repaired. In the base case, the rth element of the first column is $(B'_k)_{r1} = \lambda^{kr} - \lambda^{k(r-1)} = \lambda^{k(r-1)}(\lambda^k - 1)$ for $r \ge 2$. We divide this column by $\lambda^k - 1$ to obtain $\lambda^{k(r-1)}$, which is allowed since λ is not a root of unity and $k \neq 0$. This is now the correct form for the r'-th element of the first column of a block in M'', for the row index $r' = r - 1 \ge 1$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Now for the inductive step of B'_k, assume that the first $d-1$ columns of the block B'_k have been repaired to the correct form to be a block in M''. That is, for row index $r \geq 2$, which denotes the $(r-1)$-th row of M'', the rth element in the first $d-1$ columns of B'_k has the form $(B'_k)_{rc} = (r-1)^{c-1}\lambda^{k(r-1)}$. The rth element in column d of B'_k currently has the form $(B'_k)_{rd} = r^{d-1}\lambda^{kr} - (r-1)^{d-1}\lambda^{k(r-1)}$. Then we do column operations

$$\begin{aligned}(B'_k)_{rd} - \sum_{c=1}^{d-1}\binom{d-1}{c-1}(B'_k)_{rc} &= r^{d-1}\lambda^{kr} - (r-1)^{d-1}\lambda^{k(r-1)} \\ &\quad - \sum_{c=1}^{d-1}\binom{d-1}{c-1}(r-1)^{c-1}\lambda^{k(r-1)} \\ &= r^{d-1}\lambda^{kr} - r^{d-1}\lambda^{k(r-1)} \\ &= r^{d-1}\lambda^{k(r-1)}(\lambda^k - 1)\end{aligned}$$

and divide by $(\lambda^k - 1)$ to get $r^{d-1}\lambda^{k(r-1)}$. Once again, this is allowed since λ is not a root of unity and $k \neq 0$. Then more column operations of the same kind yield

$$\begin{aligned}&r^{d-1}\lambda^{k(r-1)} - \sum_{c=1}^{d-1}\binom{d-1}{c-1}(r-1)^{c-1}\lambda^{k(r-1)} \\ &\quad = \lambda^{k(r-1)}\left(r^{d-1} + (r-1)^{d-1} - \sum_{c=1}^{d}\binom{d-1}{c-1}(r-1)^{c-1}\right)\end{aligned}$$

and the expression inside the large parentheses on the RHS is precisely $(r-1)^{d-1}$. This gives the correct form for the rth element in column d of B'_k in M''.

Finally we repair the columns in B'_0, also by induction on the number of columns. In the base case, if B'_0 has only one column, then there is nothing to prove, since this block has disappeared in M''. Otherwise, $(B'_0)_{r2} = r - (r-1) = 1$, so the second column is already in the correct form to be the first column in M'', and there is still nothing to prove. For the inductive step, assume that columns 2 to $d-1$ are in the correct form to be the first block in M'' for $d \geq 3$. That is, the entry at row $r \geq 2$ and column c from 2 through $d-1$ has the form $(B'_0)_{rc} = (r-1)^{c-2}$. The rth element in column d currently has the form $(B'_0)_{rd} = r^{d-1} - (r-1)^{d-1}$. Then we do the column operations

$$\begin{aligned}(B'_0)_{rd} - \sum_{c=2}^{d-1}\binom{d-1}{c-2}(B'_0)_{rc} &= r^{d-1} - (r-1)^{d-1} - \sum_{c=2}^{d-1}\binom{d-1}{c-2}(r-1)^{c-2} \\ &= (d-1)(r-1)^{d-2}\end{aligned}$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

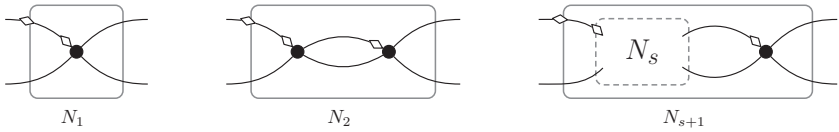


Figure 7.2. Recursive construction to interpolate g. The vertices are assigned f. For not necessarily symmetric signatures we place a diamond on the edge corresponding to the most significant index bit. The remaining index bits are in order of decreasing significance as one travels counterclockwise around the vertex.

and divide by $d-1$, which is nonzero, to get $(r-1)^{d-2}$. Recall that the block B'_0 in M'' is indexed by r and d starting at 2. This is the correct form for the rth element in column d of B'_0 in M''. Therefore, we invoke our original induction hypothesis that the $(n-1)\times(n-1)$ matrix M'' has a nonzero determinant, which completes the proof. □

Lemma 7.23. *Let g be the arity 4 signature with M_g given in Equation (7.7) and let f be an arity 4 signature with complex weights. If M_f is redundant and $\widetilde{M}_f$ is nonsingular, then for any set $\mathcal{F}$ containing f, we have*

$$\text{Holant}(\mathcal{F}\cup\{g\}) \leq_{\text{T}} \text{Holant}(\mathcal{F}).$$

The proof also works for planar graphs, thus,

$$\text{Pl-Holant}(\mathcal{F}\cup\{g\}) \leq_{\text{T}} \text{Pl-Holant}(\mathcal{F}).$$

Proof. Consider an instance Ω of $\text{Holant}(\mathcal{F}\cup\{g\})$. Suppose that g appears n times in Ω. We construct from Ω a sequence of instances Ω_s of $\text{Holant}(\mathcal{F})$ indexed by $s\geq 1$. We obtain Ω_s from Ω by replacing each occurrence of g with the gadget N_s in Figure 7.2 with f assigned to all vertices. In Ω_s, the edge corresponding to the ith significant index bit of N_s connects to the same location as the edge corresponding to the ith significant index bit of g in Ω. Since f may not be symmetric, in Figure 7.2 we place a diamond on the edge corresponding to the most significant index bit. The remaining index bits are in order of decreasing significance as one travels counterclockwise around the vertex.

Now to determine the relationship between Holant_Ω and Holant_{Ω_s}, we use the isomorphism between $\text{RM}_4(\mathbb{C})$ and $\mathbb{C}^{3\times 3}$. To compute Holant_{Ω_s}, we effectively replace M_g in Holant_Ω with $M_{N_s}=(M_f)^s$, the sth power of the signature matrix M_f. By the Jordan normal form of $\widetilde{M}_f$, there exists T and $\Lambda\in\mathbb{C}^{3\times 3}$ such that

$$\widetilde{M}_f = T\Lambda T^{-1} = T\begin{bmatrix}\lambda_1 & b_1 & 0\\ 0 & \lambda_2 & b_2\\ 0 & 0 & \lambda_3\end{bmatrix}T^{-1},$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

where $b_1, b_2 \in \{0, 1\}$. By assumption $\lambda_1\lambda_2\lambda_3 = \det(\widetilde{M}_f) \neq 0$. Also since $\widetilde{M}_g = \varphi(M_g) = I_3$, and $TI_3T^{-1} = I_3$, we have $\psi(T)M_g\psi(T^{-1}) = M_g$. We can view our construction of Ω_s as first replacing each M_g by $\psi(T)M_g\psi(T^{-1})$, which does not change the Holant value, and then replacing each new M_g, sandwiched between $\psi(T)$ and $\psi(T^{-1})$, by $\psi(\Lambda^s) = \psi(\Lambda)^s$ to obtain Ω_s. Observe that

$$\begin{aligned}\varphi(\psi(T)\psi(\Lambda^s)\psi(T^{-1})) &= T\Lambda^sT^{-1} = (\widetilde{M}_f)^s = (\varphi(M_f))^s \\ &= \varphi((M_f)^s) = \varphi(M_{N_s}),\end{aligned}$$

and both $\psi(T)\psi(\Lambda^s)\psi(T^{-1})$ and M_{N_s} are redundant, hence $\psi(T)\psi(\Lambda^s)\psi(T^{-1}) = M_{N_s}$. Since $M_g = \psi(T)\,M_g\psi(T^{-1})$ and $M_{N_s} = \psi(T)\psi(\Lambda^s)\psi(T^{-1})$, replacing each M_g, sandwiched between $\psi(T)$ and $\psi(T^{-1})$, by $\psi(\Lambda^s)$ indeed transforms Ω to Ω_s. (We note that, by the isomorphism, $\psi(T^{-1})$ *is* the multiplicative inverse of $\psi(T)$ within the semigroup $\mathrm{RM}_4(\mathbb{C})$; but we prefer not to write it as $\psi(T)^{-1}$ since it is not the usual matrix inverse as a 4×4 matrix. Indeed, $\psi(T)$ is not invertible as a 4×4 matrix.)

In the analysis that follows, we stratify assignments in Ω_s based on assignment values to $\psi(\Lambda^s)$. The inputs to $\psi(\Lambda^s)$ are from $\{0, 1\}^2 \times \{0, 1\}^2$. However, we can combine the inputs 01 and 10, since $\psi(\Lambda^s)$ is redundant. Thus we actually stratify assignments in Ω_s based on assignment values to Λ^s, which are from $\{0, 1, 2\} \times \{0, 1, 2\}$. In this compressed form, the row and column assignments to Λ^s are the Hamming weight of the two actual inputs from $\{0, 1\}^2$ to the uncompressed form $\psi(\Lambda^s)$.

Now we begin the analysis based on the values of b_1 and b_2, i.e., the shape Λ takes as the Jordan normal form of $\widetilde{M}_f$.

1. $b_1 = b_2 = 0$. In this case $\widetilde{M}_f$ is diagonalizable, we only need to consider the assignments to Λ^s that assign
   – $(0, 0)$ i many times,
   – $(1, 1)$ j many times, and
   – $(2, 2)$ k many times,
   such that $i + j + k = n$, the number of times g occurs in Ω, which is also the number of times the gadget N_s occurs in Ω_s, for all $s \geq 1$. Note that any other assignment contributes 0 to the Holant sum. Also note that

$$\psi(\Lambda^s) = \begin{bmatrix} \lambda_1^s & 0 & 0 & 0 \\ 0 & \lambda_2^s/2 & \lambda_2^s/2 & 0 \\ 0 & \lambda_2^s/2 & \lambda_2^s/2 & 0 \\ 0 & 0 & 0 & \lambda_3^s \end{bmatrix}.$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Let c_{ijk} be the sum over all such assignments of the products of evaluations (excluding Λ^s but including the contributions from T and T^{-1}) on Ω_s. Note that c_{ijk} is independent of s. Then by the form of M_g in (7.7), we have

$$\text{Holant}_\Omega = \sum_{i+j+k=n} \frac{c_{ijk}}{2^j}$$

and the value of the Holant on Ω_s, for $s \geq 1$, is

$$\text{Holant}_{\Omega_s} = \sum_{i+j+k=n} \left(\lambda_1^i \lambda_2^j \lambda_3^k\right)^s \left(\frac{c_{ijk}}{2^j}\right),$$

where the factor $\frac{1}{2}$ comes from the middle four entries in M_g in (7.7) and in $\psi(\Lambda^s)$. The coefficient matrix is Vandermonde, but it may not have full rank because it may be that $\lambda_1^i \lambda_2^j \lambda_3^k = \lambda_1^{i'} \lambda_2^{j'} \lambda_3^{k'}$ for some $(i, j, k) \neq (i', j', k')$. However, this is not a problem since we are interested only in the sum $\sum c_{ijk}/2^j$. If two coefficients are the same, we replace their corresponding unknowns $\frac{c_{ijk}}{2^j}$ and $\frac{c_{i'j'k'}}{2^{j'}}$ with their sum as a new variable. After all such combinations, we have a Vandermonde system of full rank. In particular, none of the entries are zero since $\lambda_1\lambda_2\lambda_3 = \det(\widetilde{M}_f) \neq 0$. Therefore, we can solve the linear system in polynomial time and obtain the value of Holant_Ω.

2. $b_1 \neq b_2$. We can assume that the Jordan blocks in Λ have been permuted so that $b_1 = 1$ and $b_2 = 0$. Then $\lambda_1 = \lambda_2$, and we denote it by λ. Denote λ_3 by μ. We have

$$\Lambda^s = \begin{bmatrix} \lambda^s & s\lambda^{s-1} & 0 \\ 0 & \lambda^s & 0 \\ 0 & 0 & \mu^s \end{bmatrix}, \text{ and } \psi(\Lambda^s) = \begin{bmatrix} \lambda^s & s\lambda^{s-1}/2 & s\lambda^{s-1}/2 & 0 \\ 0 & \lambda^s/2 & \lambda^s/2 & 0 \\ 0 & \lambda^s/2 & \lambda^s/2 & 0 \\ 0 & 0 & 0 & \mu^s \end{bmatrix}.$$

We only need to consider the assignments to Λ^s that assign
– $(0, 0)$ i many times,
– $(1, 1)$ j many times,
– $(2, 2)$ k many times, and
– $(0, 1)$ ℓ many times,
with $i + j + k + \ell = n$, since any other assignment contributes 0 to the Holant sum. Let $c_{ijk\ell}$ be the sum over all such assignments of the products of evaluations (excluding Λ^s but including the contributions from T and T^{-1}) on Ω_s; $c_{ijk\ell}$ is independent of s. Then

$$\text{Holant}_\Omega = \sum_{i+j+k=n} \frac{c_{ijk0}}{2^j} \tag{7.8}$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

and the value of the Holant on Ω_s, for $s \geq 1$, is

$$\begin{aligned}\text{Holant}_{\Omega_s} &= \sum_{i+j+k+\ell=n} \lambda^{(i+j)s} \mu^{ks} \left(s\lambda^{s-1}\right)^{\ell} \left(\frac{c_{ijk\ell}}{2^{j+\ell}}\right) \\ &= \lambda^{ns} \sum_{i+j+k+\ell=n} \left(\frac{\mu}{\lambda}\right)^{ks} s^{\ell} \left(\frac{c_{ijk\ell}}{\lambda^{\ell} 2^{j+\ell}}\right). \end{aligned} \tag{7.9}$$

If μ/λ is a root of unity, then take a t such that $(\mu/\lambda)^t = 1$. Then

$$\text{Holant}_{\Omega_{st}} = \lambda^{nst} \sum_{i+j+k+\ell=n} s^{\ell} \left(\frac{t^{\ell} c_{ijk\ell}}{\lambda^{\ell} 2^{j+\ell}}\right).$$

For $s \geq 1$, this gives a coefficient matrix $\left(\left(\frac{st}{2\lambda}\right)^{\ell}\right)$ (with rows indexed by s and columns indexed by the tuple (i, j, k, ℓ)) that is Vandermonde. Although this system is not full rank, we can replace all the unknowns $c_{ijk\ell}/2^j$ having $i + j + k = n - \ell$ by their sum to form a new unknown $c'_{\ell} = \sum_{i+j+k=n-\ell} c_{ijk\ell}/2^j$, where $0 \leq \ell \leq n$. The unknown c'_0 is the Holant of Ω that we seek. The resulting Vandermonde system

$$\text{Holant}_{\Omega_{st}} = \lambda^{nst} \sum_{\ell=0}^{n} \left(\frac{st}{2\lambda}\right)^{\ell} c'_{\ell}$$

has full rank, so we can solve for the unknowns in polynomial time and obtain the value of $c'_0 = \mathrm{e} \sum_{i+j+k=n} \frac{c_{ijk0}}{2^j}$.

If μ/λ is not a root of unity, then we replace all the unknowns $c_{ijk\ell}/(\lambda^{\ell} 2^{j+\ell})$ having $i + j = m$ with their sum to form new unknowns $c'_{mk\ell}$, for any $0 \leq m, k, \ell$ and $m + k + \ell = n$. Then by (7.8) and (7.9),

$$\text{Holant}_{\Omega} = \sum_{m+k=n} c'_{mk0} \quad \text{and} \quad \text{Holant}_{\Omega_s} = \lambda^{ns} \sum_{m+k+\ell=n} \left(\frac{\mu}{\lambda}\right)^{ks} s^{\ell} c'_{mk\ell}.$$

After a suitable ordering of the columns, the matrix of coefficients whose rows are indexed by s satisfies the hypothesis of Lemma 7.22. Therefore, the linear system has full rank. We can solve for the unknowns in polynomial time and obtain the value of Holant_{Ω}.

3. $b_1 = b_2 = 1$. In this case, we have $\lambda_1 = \lambda_2 = \lambda_3$, denoted by λ. We have

$$\Lambda^s = \begin{bmatrix} \lambda^s & s\lambda^{s-1} & \frac{s(s-1)}{2}\lambda^{s-2} \\ 0 & \lambda^s & s\lambda^{s-1} \\ 0 & 0 & \lambda^s \end{bmatrix},$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

and

$$\psi(\Lambda^s) = \begin{bmatrix} \lambda^s & \frac{s}{2}\lambda^{s-1} & \frac{s}{2}\lambda^{s-1} & \frac{s(s-1)}{2}\lambda^{s-2} \\ 0 & \frac{\lambda^s}{2} & \frac{\lambda^s}{2} & s\lambda^{s-1} \\ 0 & \frac{\lambda^s}{2} & \frac{\lambda^s}{2} & s\lambda^{s-1} \\ 0 & 0 & 0 & \lambda^s \end{bmatrix}.$$

We only need to consider the assignments to Λ^s that assign
– $(0, 0)$ or $(2, 2)$ i many times,
– $(1, 1)$ j many times,
– $(0, 1)$ k many times,
– $(1, 2)$ ℓ many times, and
– $(0, 2)$ m many times,
with $i + j + k + \ell + m = n$, since any other assignment contributes 0 to the Holant sum. Let $c_{ijk\ell m}$ be the sum over all such assignments of the products of evaluations (excluding Λ^s but including the contributions from T and T^{-1}) on Ω_s; $c_{ijk\ell m}$ is independent of s. Then

$$\text{Holant}_{\Omega} = \sum_{i+j=n} \frac{c_{ij000}}{2^j}$$

and the value of the Holant on Ω_s, for $s \geq 1$, is

$$\begin{aligned} \text{Holant}_{\Omega_s} &= \sum_{i+j+k+\ell+m=n} \lambda^{(i+j)s} \left(s\lambda^{s-1}\right)^{k+\ell} \left(s(s-1)\lambda^{s-2}\right)^m \left(\frac{c_{ijk\ell m}}{2^{j+k+m}}\right) \\ &= \lambda^{ns} \sum_{i+j+k+\ell+m=n} s^{k+\ell+m}(s-1)^m \left(\frac{c_{ijk\ell m}}{\lambda^{k+\ell+2m}2^{j+k+m}}\right). \end{aligned}$$

We replace all the unknowns $c_{ijk\ell m}/(\lambda^{k+\ell+2m}2^{j+k+m})$ having $i + j = p$ and $k + \ell = q$ with their sum to form new unknowns c'_{pqm}, for any $0 \leq p, q, m$ and $p + q + m = n$. The Holant of Ω is now c'_{n00}. This new linear system is

$$\text{Holant}_{\Omega_s} = \lambda^{ns} \sum_{p+q+m=n} s^{q+m}(s-1)^m c'_{pqm}.$$

But this linear system is still rank deficient. We now index the columns by the tuple (q, m), where $q \geq 0$, $m \geq 0$, and $q + m \leq n$. Correspondingly, we rename the variables $x_{q,m} = c'_{pqm}$. Note that $p = n - q - m$ is determined by (q, m). Observe that the column indexed by (q, m) is the sum of the columns indexed by $(q - 1, m)$ and $(q - 2, m + 1)$ provided $q - 2 \geq 0$. Namely, $s^{q+m}(s-1)^m = s^{q-1+m}(s-1)^m + s^{q-2+m+1}(s-1)^{m+1}$. Of course this is meaningful only if $q \geq 2$, $m \geq 0$ and $q + m \leq n$. Let

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

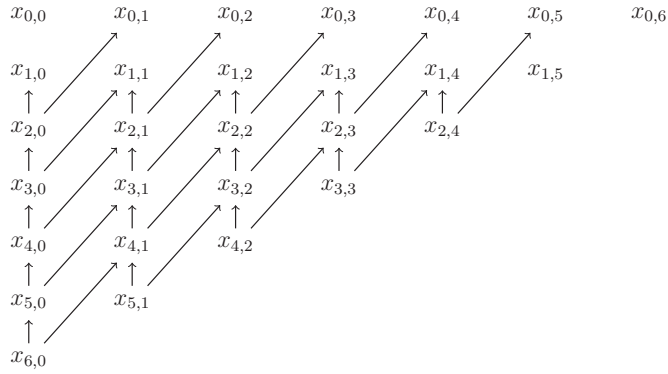


Figure 7.3. A triangular table of variables and their update patterns, for $n = 6$. Rows are indexed by q and columns are indexed by m.

$a_{s,(q,m)} = s^{q+m}(s-1)^m$. Then the linear system is

$$\sum_{q\geq 0,\ m\geq 0,\ q+m\leq n} a_{s,(q,m)} x_{q,m} = \frac{\text{Holant}_{\Omega_s}}{\lambda^{ns}}. \tag{7.10}$$

For any fixed $s \geq 1$, we write $\alpha_{q,m} = a_{s,(q,m)}$, then $\alpha_{q,m}x_{q,m} = (\alpha_{q-1,m} + \alpha_{q-2,m+1})x_{q,m}$. Therefore we can eliminate variable $x_{q,m}$ by adding its value to $x_{q-1,m}$ and to $x_{q-2,m+1}$. More precisely we define new variables

$$x_{q-1,m} \leftarrow x_{q-1,m} + x_{q,m}$$

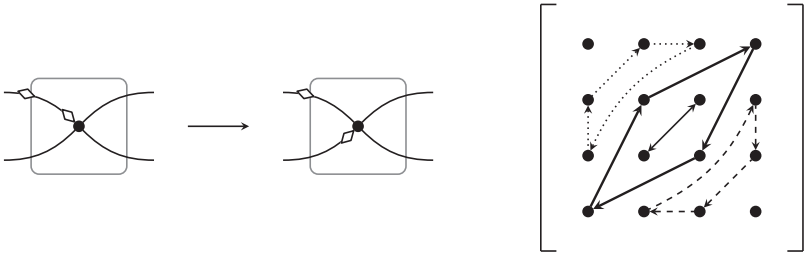
$$x_{q-2,m+1} \leftarrow x_{q-2,m+1} + x_{q,m}$$

for every $0 \leq m \leq n-2$ and for q from $n-m$ down to 2. The order of these updates is important. Once a variable $x_{q,m}$ has been added to $x_{q-1,m}$ and $x_{q-2,m+1}$, there should not be any update involving $x_{q,m}$, which has been eliminated. One valid sequence is a double nested loop, with the outer loop being "For $m = 0$ to $n-2$" and for each m in the range there is the inner loop "For $q = n-m$ down to 2". This is illustrated schematically for $n = 6$ in Figure 7.3.

Observe that in each update, some variable $x_{q,m}$ is added to a variable $x_{q',m'}$, where $q' < q$ (a strictly lower index value of q), and $m' \geq m$ (an index value that is $\geq m$). Thus in the order by the double nested loop, once $x_{q,m}$ has been added to $x_{q-1,m}$ and $x_{q-2,m+1}$, there will be no more updates involving $x_{q,m}$.

A more crucial observation is that the column in the linear system (7.10) indexed by the tuple $(q, m) = (0, 0)$ is never changed. This is because, in order to be an updated entry, there must be some $q \geq 2$ and $m \geq 0$ such that

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press



(a) A counterclockwise rotation.

(b) Movement of signature matrix entries under a counterclockwise rotation.

Figure 7.4. The movement of the entries in the signature matrix of an arity 4 signature under a counterclockwise rotation of the input edges. The Hamming weight 1 entries are in the dotted cycle, the Hamming weight 2 entries are in the two solid cycles (one has length 4 and the other one is a swap), and the entries of Hamming weight 3 are in the dashed cycle.

$(q-1, m) = (0, 0)$ or $(q-2, m+1) = (0, 0)$, which is clearly impossible. Hence $x_{0,0} = c'_{n00}$ is still the Holant value on Ω. The $2n+1$ unknowns that remain are

$$x_{0,0},\ x_{1,0},\ x_{0,1},\ x_{1,1},\ x_{0,2},\ x_{1,2}, \ldots,\ x_{0,n-1},\ x_{1,n-1},\ x_{0,n}$$

and their coefficients in row s are

$$1, s, s(s-1), s^2(s-1), s^2(s-1)^2, \ldots, s^{n-1}(s-1)^{n-1},$$
$$s^n(s-1)^{n-1}, s^n(s-1)^n.$$

For $0 \le \delta \le 2n$, it is clear that the δth entry in this row is a monic polynomial in s of degree δ, and thus s^δ is a linear combination of the first $\delta+1$ entries indexed by 0 to δ. It follows that the coefficient matrix is a product of the standard Vandermonde matrix multiplied to its right by an upper triangular matrix with all 1's on the diagonal. Hence the matrix is nonsingular, and we can solve the linear system in polynomial time and, in particular, compute c'_{n00}. □

For an asymmetric signature, we often want to reorder the input bits under a circular permutation. For a single counterclockwise rotation of 90°, the effect on the entries of the signature matrix of an arity 4 signature is given in Figure 7.4.

We will derive most of our #P-hardness results through Lemma 7.23 and the following Lemma 7.24. This is done by a reduction from the problem of

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

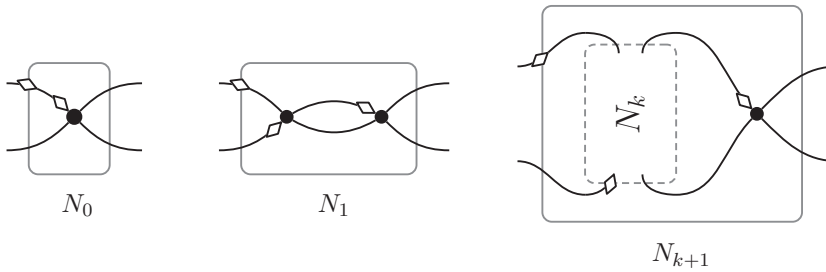


Figure 7.5. Recursive construction to approximate $[1, 0, \frac{1}{3}, 0, 1]$. The vertices are assigned g.

counting Eulerian orientations on 4-regular graphs, which is the Holant problem Holant$((\neq_2) \mid [0, 0, 1, 0, 0])$. In Theorem 6.32 we have shown that this problem #EO is #P-hard even for planar graphs. Recall that under a holographic transformation by $\begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$, this bipartite Holant problem becomes the Holant problem Holant$([1, 0, 1/3, 0, 1])$, up to a nonzero factor.

Lemma 7.24. *Let g be the arity 4 signature with M_g given in Equation (7.7) so that $\widetilde{M}_g = I_3$. Then* Holant(g) *is* #P*-hard. In fact* Pl-Holant(g) *is* #P*-hard.*

Proof. We reduce from the #EO problem Holant$(\mathscr{O})$, where $\mathscr{O} = [1, 0, 1/3, 0, 1]$, which is #P-hard by Theorem 6.32 even for 4-regular planar graphs. We achieve this via an arbitrarily close approximation using the recursive construction in Figure 7.5 with g assigned to every vertex.

We claim that the signature matrix M_{N_k} of gadget N_k is

$$M_{N_k} = \begin{bmatrix} 1 & 0 & 0 & a_k \\ 0 & a_{k+1} & a_{k+1} & 0 \\ 0 & a_{k+1} & a_{k+1} & 0 \\ a_k & 0 & 0 & 1 \end{bmatrix},$$

where $a_k = \frac{1}{3} - \frac{1}{3}\left(-\frac{1}{2}\right)^k$. This is true for N_0. Inductively assume M_{N_k} has this form. Then the rotated form of the signature matrix for N_k, as described in Figure 7.4, is

$$\begin{bmatrix} 1 & 0 & 0 & a_{k+1} \\ 0 & a_k & a_{k+1} & 0 \\ 0 & a_{k+1} & a_k & 0 \\ a_{k+1} & 0 & 0 & 1 \end{bmatrix}. \tag{7.11}$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

The action of g on the far right side of N_{k+1} in Figure 7.5 is to replace each of the middle two entries in the middle two rows of the matrix in (7.11) with their average, $(a_k + a_{k+1})/2 = a_{k+2}$. This gives $M_{N_{k+1}}$.

Let G be a graph with n vertices and $H_{\mathscr{O}}$ (resp. H_{N_k}) be the Holant value on G with all vertices assigned $\mathscr{O}$ (resp. N_k). Since each signature entry in $\mathscr{O}$ can be expressed as a rational number with denominator 3, each term in the sum of $H_{\mathscr{O}}$ can be expressed as a rational number with denominator 3^n, and $H_{\mathscr{O}}$ itself is a sum of 2^{2n} such terms, where $2n$ is the number of edges in G. If the error $|H_{N_k} - H_{\mathscr{O}}|$ is $\le 1/3^{n+1}$, then we can recover $H_{\mathscr{O}}$ from H_{N_k} by selecting the nearest rational number to H_{N_k} with denominator 3^n.

For each signature entry x in $M_{\mathscr{O}}$, its corresponding entry $\tilde{x}$ in M_{N_k} satisfies $|\tilde{x} - x| \le x/2^k$. Then for each term t in the Holant sum $H_{\mathscr{O}}$, its corresponding term $\tilde{t}$ in the sum H_{N_k} satisfies $t(1 - 1/2^k)^n \le \tilde{t} \le t(1 + 1/2^k)^n$, thus $-t(1 - (1 - 1/2^k)^n) \le \tilde{t} - t \le t((1 + 1/2^k)^n - 1)$. Since $1 - (1 - 1/2^k)^n \le (1 + 1/2^k)^n - 1$, which is easy to see by their binomial expansions, we get $|\tilde{t} - t| \le t((1 + 1/2^k)^n - 1)$. Also each term $t \le 1$. Hence if we take $k = 4n + 4$, then

$$|H_{N_k} - H_{\mathscr{O}}| \le 2^{2n}((1 + 1/2^k)^n - 1) < 1/3^{n+1}.$$

Note that $(1 + x)^n - 1 < xn(1 + x)^n < exn$, for $x = 1/2^{4n+4}$. □

We summarize our progress with the following corollary of Lemmas 7.23 and 7.24.

Corollary 7.25. *Let f be an arity 4 signature with complex weights. If M_f is redundant and $\widetilde{M}_f$ is nonsingular, then* Holant(f) *is* #P*-hard. In fact* Pl-Holant(f) *is* #P*-hard.*

Exercise: By Lemma 7.24, Pl-Holant(g) is #P-hard, where g is defined in Equation (7.7). The reduction chain is from the #P-hardness of T(G; 3, 3) of the Tutte polynomial to #EO (by Theorem 6.32), then to Pl-Holant(g) (by Lemma 7.24).

Prove this directly from the #P-hardness of T(G; 3, 3), skipping #EO. (Hint: First apply a rotation (Figure 7.4) to $2g$, then interpolate Pl-Holant($\hat{f}$), where $\hat{f}$ is defined in (6.21).)

In order to make Corollary 7.25 more applicable, we show that for an arity 4 signature f, the redundancy of M_f and the nonsingularity of $\widetilde{M}_f$ are invariant under an invertible holographic transformation.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Lemma 7.26. *Let f be an arity 4 signature with complex weights, $T \in \mathbb{C}^{2\times 2}$ a matrix, and $\hat{f} = T^{\otimes 4} f$. If M_f is redundant, then $M_{\hat{f}}$ is also redundant and*

$$\det(\varphi(M_{\hat{f}})) = \det(\varphi(M_f)) \det(T)^6.$$

Proof. Since $\hat{f} = T^{\otimes 4} f$, we can express $M_{\hat{f}}$ in terms of M_f and T as

$$M_{\hat{f}} = T^{\otimes 2} M_f \left(T^{\mathrm{T}}\right)^{\otimes 2}. \tag{7.12}$$

This can be directly checked. Let $T = (t^i_j)$, where row index i and column index j are from $\{0, 1\}$. The entry $f(a, b, c, d)$ appears in M_f at row $(a, b) \in \{0, 1\}^2$ and column $(d, c) \in \{0, 1\}^2$ in lexicographic order. The matrix $T^{\otimes 2}$ has the form $(t^a_{a'} t^b_{b'})$ at row (a, b) and column (a', b') in lexicographic order. Similarly $(T^{\mathrm{T}})^{\otimes 2}$ has the form $(t^{c'}_c t^{d'}_d)$ at row (c, d) and column (c', d'). Thus row (a, b) and column (d, c) of $T^{\otimes 2} M_f \left(T^{\mathrm{T}}\right)^{\otimes 2}$ has entry $\sum_{a',b',c',d' \in \{0,1\}} t^a_{a'} t^b_{b'} f(a', b', c', d') t^d_{d'} t^c_{c'}$, which is the entry $\hat{f}(a, b, c, d)$ placed at row (a, b) and column (d, c) of $M_{\hat{f}}$.

An alternative, but essentially the same, proof of (7.12) is by linearity, and observing that it is valid for the basis vectors $f = e_a \otimes e_b \otimes e_c \otimes e_d$, for $a, b, c, d \in \{0, 1\}$.

Let

$$\mathcal{E} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then $X \in \mathrm{RM}_4(\mathbb{C})$ iff $\mathcal{E}X = X = X\mathcal{E}$. Then it follows that $M_{\hat{f}} \in \mathrm{RM}_4(\mathbb{C})$ if $M_f \in \mathrm{RM}_4(\mathbb{C})$, since $\mathcal{E} T^{\otimes 2} \mathcal{E} = T^{\otimes 2}$, i.e., a simultaneous row flip $ab \leftrightarrow ba$ and column flip $a'b' \leftrightarrow b'a'$ keep $T^{\otimes 2}$ unchanged. For the two matrices A and B in the definition of φ, we note that $BA = M_g$, where M_g given in Equation (7.7) is the identity element of the semigroup $\mathrm{RM}_4(\mathbb{C})$. Since $M_f \in \mathrm{RM}_4(\mathbb{C})$, we have $BAM_f = M_f = M_f BA$. Then we have

$$\begin{aligned} \varphi(M_{\hat{f}}) = AM_{\hat{f}}B &= A\left(T^{\otimes 2} M_f \left(T^{\mathrm{T}}\right)^{\otimes 2}\right) B \\ &= (AT^{\otimes 2}B)(AM_fB)(A\left(T^{\mathrm{T}}\right)^{\otimes 2} B) \\ &= \varphi(T^{\otimes 2})\varphi(M_f)\varphi(\left(T^{\mathrm{T}}\right)^{\otimes 2}). \end{aligned} \tag{7.13}$$

Another direct calculation shows that

$$\det(\varphi(T^{\otimes 2})) = \det(T)^3 = \det(\varphi(\left(T^{\mathrm{T}}\right)^{\otimes 2})). \tag{7.14}$$

To verify (7.14) one can save some calculation by noticing that for any T_1 and T_2, $\varphi((T_1T_2)^{\otimes 2}) = AT_1^{\otimes 2} T_2^{\otimes 2} B$ by definition, and $AT_1^{\otimes 2}$ has identical two

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

middle columns, thus $AT_1^{\otimes 2} = AT_1^{\otimes 2}BA$, thus φ is multiplicative on $T^{\otimes 2}$, i.e., $\varphi((T_1T_2)^{\otimes 2}) = \varphi(T_1^{\otimes 2})\varphi(T_2^{\otimes 2})$. Note that (7.14) is easy to check for triangular matrices.

Thus, by applying determinant to both sides of Equation (7.13), we have

$$\det(\varphi(M_{\hat{f}})) = \det(\varphi(M_f))\det(T)^6,$$

as claimed. □

In particular, for a nonsingular matrix $T \in \mathbb{C}^{2\times 2}$, M_f is redundant and $\widetilde{M}_f$ is nonsingular iff $M_{\hat{f}}$ is redundant and $\widetilde{M}_{\hat{f}}$ is nonsingular. From Corollary 7.25 and Lemma 7.26 we have the following corollary.

Corollary 7.27. *Let f be an arity 4 signature with complex weights. If there exists a nonsingular matrix $T \in \mathbb{C}^{2\times 2}$ such that $\hat{f} = T^{\otimes 4}f$, where $M_{\hat{f}}$ is redundant and $\widetilde{M}_{\hat{f}}$ is nonsingular, then* Holant(f) *is* #P*-hard. In fact* Pl-Holant(f) *is* #P*-hard.*

The following lemma applies Corollary 7.25.

Lemma 7.28. *Let $f_k = ck\lambda^{k-1} + d\lambda^k$, where $c \neq 0$ and $0 \leq k \leq 4$. Then the problem* Holant$([f_0, f_1, f_2, f_3, f_4])$ *is* #P*-hard unless $\lambda = \pm\mathrm{i}$, in which case $[f_0, f_1, f_2, f_3, f_4]$ is a vanishing signature and the Holant vanishes.*

Proof. If $\lambda = \pm\mathrm{i}$, then $\mathrm{rd}^{\pm}(f) = 1$, $\mathrm{vd}^{\pm}(f) = 3$, and so $f = [f_0, f_1, f_2, f_3, f_4]$ is vanishing by Theorem 7.11. Otherwise, a holographic transformation with the orthogonal matrix $T = \frac{1}{\sqrt{1+\lambda^2}}\begin{bmatrix} 1 & \lambda \\ \lambda & -1 \end{bmatrix}$ transforms f to $\hat{f} = [t, 1, 0, 0, 0]$ for some $t \in \mathbb{C}$ after normalizing the second entry.

This orthogonal transformation can be directly verified; but a more informative derivation is as follows: Consider a symmetric signature $f = [f_0, \ldots, f_n]$ of general arity n, such that $f_k = ck\lambda^{k-1} + d\lambda^k$, where $c \neq 0$, and $\lambda \neq \pm\mathrm{i}$. Let $S = \begin{bmatrix} 1 & \frac{d-1}{n} \\ \lambda & c + \frac{d-1}{n}\lambda \end{bmatrix}$. Note that $\det S = c \neq 0$. Then the signature f can be expressed as

$$f = S^{\otimes n}[1, 1, 0, \ldots, 0],$$

where $[1, 1, 0, \ldots, 0]$ should be understood as a column vector of dimension 2^n, which has a 1 in entries with index weight ≤ 1 and 0 elsewhere. This identity can be verified by observing that

$$[1, 1, 0, \ldots, 0] = [1, 0]^{\otimes n} + \frac{1}{(n-1)!}\,\mathrm{Sym}_n^{n-1}([1, 0]; [0, 1]).$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

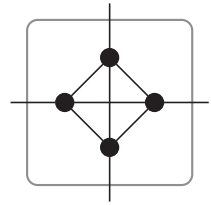


Figure 7.6. The tetrahedron gadget. Each vertex is assigned $\hat{f} = [t, 1, 0, 0, 0]$.

Let $T = \frac{1}{\sqrt{1+\lambda^2}} \begin{bmatrix} 1 & \lambda \\ \lambda & -1 \end{bmatrix}$, then $T = T^{\mathrm{T}} = T^{-1} \in \mathbf{O}_2(\mathbb{C})$ is orthogonal, and $R = TS = \begin{bmatrix} u & w \\ 0 & v \end{bmatrix}$ is upper triangular, where $v, w \in \mathbb{C}$ and $u = \sqrt{1+\lambda^2} \neq 0$. However, $\det R = \det T \det S = (-1)c \neq 0$, so we also have $v \neq 0$. It follows that

$$\begin{aligned} T^{\otimes n} f &= (TS)^{\otimes n}[1, 1, 0, \ldots, 0] \\ &= R^{\otimes n}[1, 1, 0, \ldots, 0] \\ &= R^{\otimes n}\left([1, 0]^{\otimes n} + \frac{1}{(n-1)!}\operatorname{Sym}_n^{n-1}([1, 0]; [0, 1])\right) \\ &= [u, 0]^{\otimes n} + \frac{1}{(n-1)!}\operatorname{Sym}_n^{n-1}([u, 0]; [w, v]) \\ &= [u^n + nu^{n-1}w, u^{n-1}v, 0, \ldots, 0]. \end{aligned}$$

Since $u^{n-1}v \neq 0$, we can normalize the entry of Hamming weight 1 to the value 1 by a scalar multiplication. Thus, we have $[t, 1, 0, \ldots, 0]$ for some $t \in \mathbb{C}$.

Using the tetrahedron gadget in Figure 7.6 with $\hat{f}$ assigned to each vertex, we have a gadget with signature

$$h = [t^4 + 6t^2 + 3, t^3 + 3t, t^2 + 1, t, 1].$$

One way to verify this signature is to treat $[t, 1, 0, 0, 0]$ as a weighted matching signature, and compute the signature of the gadget similarly to a matchgate. Thus, for example, the Hamming weight 0 entry has a contribution t^4 from the empty matching of the six internal edges, a contribution $6t^2$ from the six matchings of a single edge, and a contribution 3 from the three perfect matchings. The other entries can be computed similarly.

Since the determinant of $\widetilde{M}_h$ is 4, the compressed signature matrix of this gadget is nonsingular, so we are done by Corollary 7.25. □

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Notice that Lemma 7.28 does not address the complexity of Pl-Holant($[f_0, f_1, f_2, f_3, f_4]$). In particular Pl-Holant($[0, 1, 0, 0, 0]$) is solvable in polynomial time, since $[0, 1, 0, 0, 0]$ is the PERFECT MATCHING signature. Even though Corollary 7.25 applies to Pl-Holant problems as well, the reduction in Lemma 7.28 is nonplanar, as the tetrahedron gadget in Figure 7.6 is not planar. We will show that Pl-Holant($[t, 1, 0, 0, 0]$) remains #P-hard for all $t \neq 0$ in Chapter 8 (Corollary 8.11).

Now we are ready to prove a dichotomy for a single arity 4 signature. For future convenience we state the following theorem for both the case of an arity 3 and arity 4 signature. The arity 3 case is essentially a restatement of Theorem 6.3. Theorem 7.29 does not address the complexity of Pl-Holant problems; Theorem 8.12 in Chapter 8 will be the version of Theorem 7.29 for Pl-Holant problems for arity 4 signatures.

Theorem 7.29. *If f is a nondegenerate, symmetric, complex-valued signature of arity 3 or arity 4 in Boolean variables, then* Holant(f) *is* #P*-hard unless f is $\mathscr{A}$-transformable, or $\mathscr{P}$-transformable, or vanishing, in which case the problem is in P.*

Proof. We first consider f of arity 3. We show that the tractability conditions in Theorem 6.3 are equivalent to the statement that f is $\mathscr{A}$- or $\mathscr{P}$-transformable, or vanishing, for any symmetric signature of arity 3. Clearly conditions 1 and 2 in Theorem 6.3 imply that f is $\mathscr{A}$- or $\mathscr{P}$-transformable. Condition 3 is $f \in \mathscr{R}_2^{\sigma}$ for $\sigma \in \{+, -\}$, which is the same as $\mathrm{rd}^{\sigma}(f) \leq 1$ by Lemma 7.16. This is equivalent to f is vanishing for arity 3.

Suppose f is $\mathscr{P}$-transformable (this includes the case when f is degenerate). Then for some T, $(=_2)T^{\otimes 2} \in \mathscr{P}$, and $(T^{-1})^{\otimes 3} f \in \mathscr{P}$. Since $(T^{-1})^{\otimes 3} f$ is symmetric and has arity 3, either it is degenerate, which would imply f is degenerate, or

$$(T^{-1})^{\otimes 3} f = \begin{bmatrix} a \\ 0 \end{bmatrix}^{\otimes 3} + \begin{bmatrix} 0 \\ b \end{bmatrix}^{\otimes 3},$$

for some $ab \neq 0$. Let $D = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, then $D \in \mathrm{Stab}(\mathscr{P})$. Then $((TD)^{-1})^{\otimes 3} f = (=_3)$ and $(=_2)(TD)^{\otimes 2} \in \mathscr{P}$.

Suppose f is $\mathscr{A}$-transformable and nondegenerate. Then for some T, $(=_2)T^{\otimes 2} \in \mathscr{A}$, and $(T^{-1})^{\otimes 3} f \in \mathscr{A}$. Since $(T^{-1})^{\otimes 3} f$ is symmetric, it is in $\mathscr{F}_{123} = \mathscr{F}_1 \cup \mathscr{F}_2 \cup \mathscr{F}_3$ defined in Chapter 3. By the form of $\mathscr{F}_{123}$, there exists a matrix

$$M \in \left\{ I, \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \right\}$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

such that

$$((TM)^{-1})^{\otimes 3} f = \lambda \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes 3} + \mathrm{i}^r \begin{bmatrix} 0 \\ 1 \end{bmatrix}^{\otimes 3} \right\},$$

for some $\lambda \neq 0$ and $r \in \{0, 1, 2, 3\}$, and $(=_2)(TM)^{\otimes 2} \in \mathscr{A}$ since $M \in \mathrm{Stab}(\mathscr{A})$. We take $\rho = \mathrm{i}^{3r}$, and $D = \lambda^{\frac{1}{3}} \begin{bmatrix} 1 & 0 \\ 0 & \rho \end{bmatrix} \in \mathrm{Stab}(\mathscr{A})$. Note that $\rho^3 = \mathrm{i}^r$. We conclude that $(=_2)(TMD)^{\otimes 2} \in \mathscr{A}$ and $((TMD)^{-1})^{\otimes 3} f = (=_3)$. Hence f satisfies condition 2 of Theorem 6.3.

Now we prove the dichotomy for a single arity 4 signature. Let $f = [f_0, f_1, f_2, f_3, f_4]$. If the compressed signature matrix $\widetilde{M}_f$ is nonsingular, then $\mathrm{Holant}(f)$ is #P-hard by Corollary 7.25, so assume that the rank of $\widetilde{M}_f$ is ≤ 2. Then we have

$$a \begin{pmatrix} f_0 \\ f_1 \\ f_2 \end{pmatrix} + 2b \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} + c \begin{pmatrix} f_2 \\ f_3 \\ f_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

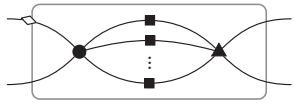
for some $a, b, c \in \mathbb{C}$ that are not all zero. If $a = c = 0$, then $b \neq 0$, so $f_1 = f_2 = f_3 = 0$. In this case, f is a generalized equality Gen-Eq, then $f \in \mathscr{P}$ and in particular f is $\mathscr{P}$-transformable. Now suppose a and c are not both 0. Then f satisfies a second-order recurrence relation. If the roots of the characteristic polynomial of the recurrence relation are distinct, then $f_k = \alpha^{4-k}\beta^k + \gamma^{4-k}\delta^k$ $(0 \leq k \leq 4)$, for some α, β, γ and δ, where $\alpha\delta - \beta\gamma \neq 0$. Let $T_1 = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}$, then $f = (=_4)T_1^{\otimes 4}$, and

$$\mathrm{Holant}(f) \equiv_{\mathrm{T}} \mathrm{Holant}((=_4)T_1^{\otimes 4} \mid (=_2)) \equiv_{\mathrm{T}} \mathrm{Holant}((=_4) \mid T_1^{\otimes 2}(=_2)).$$

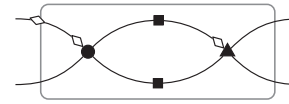
By Theorem 5.61 this problem is #P-hard except when there exists some T_2 such that for $\mathscr{C} \in \{\mathscr{P}, \mathscr{A}\}$, $(=_4)(T_2^{-1})^{\otimes 4} \in \mathscr{C}$ and $(T_2T_1)^{\otimes 2}(=_2) \in \mathscr{C}$. Let $T = (T_2T_1)^{-1}$, then $fT^{\otimes 4} \in \mathscr{C}$ and $(T^{-1})^{\otimes 2}(=_2) \in \mathscr{C}$. Hence f is either $\mathscr{A}$- or $\mathscr{P}$-transformable. Note that Theorem 5.61 is the transformational form of Theorem 5.3, which is equivalent to Theorem 5.1. The case $k = 4$ of Theorem 5.1 is proved in Chapter 5.

Now suppose the characteristic polynomial has a double root λ and there are two cases. In the first case, for any $0 \leq k \leq 4$, $f_k = ck\lambda^{k-1} + d\lambda^k$, where $c \neq 0$. In the second case, for any $0 \leq k \leq 4$, $f_k = c(4-k)\lambda^{3-k} + d\lambda^{4-k}$, where $c \neq 0$. This case is mapped to the first case by a holographic transformation by $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then we are done by Lemma 7.28. □

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press



(a) The circle is assigned f, the triangle is assigned g, and the squares are assigned $(\neq_2)$.



(b) The circle is assigned h', the triangle is assigned h'', and the squares are given $(\neq_2)$.

Figure 7.7. Gadget constructions used in Lemma 7.30. Part (a) constructs h. Its rotated versions h' and h'' are used in (b) to construct r.

7.6 Vanishing Signatures Revisited

7.6.1 Orientation with a Pair from $\mathscr{V}^+$ and $\mathscr{V}^-$

We prove a lemma on an orientation problem with a pair of vanishing signatures of the opposite type. The proof uses techniques similar to those in Section 7.5. This problem is also called #One-In-Or-One-Out-Orientation in Section 7.3, where a set of such signatures is allowed.

Lemma 7.30. *If $f = [0, 1, 0, \ldots, 0]$ and $g = [0, \ldots, 0, 1, 0]$ are both of the same arity n, for some $n \geq 3$, then the problem* Holant$(\neq_2 | \{f, g\})$ *is* #P*-hard.*

Proof. Our goal is to obtain a signature that satisfies the hypothesis of Corollary 7.27.

The gadget in Figure 7.7a, with f assigned to the circle vertex, g assigned to the triangle vertex, and $(\neq_2)$ assigned to the square vertices, has signature h with signature matrix

$$M_h = \begin{bmatrix} 0 & 0 & 0 & m \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

where $m = n - 2$ is positive since $n \geq 3$. Although this signature matrix is redundant, its compressed form is singular. Rotating this gadget 90° clockwise and 90° counterclockwise yield signatures h' and h'' respectively, with signature matrices

$$M_{h'} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & m & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad M_{h''} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & m & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

The gadget in Figure 7.7b, with h' assigned to the circle vertex, h'' assigned to the triangle vertex, and $(\neq_2)$ assigned to the square vertices, has a signature r

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

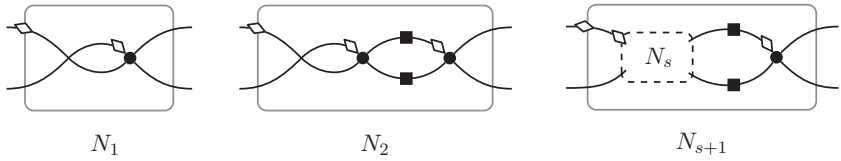


Figure 7.8. Recursive construction to interpolate a signature R that is only a rotation away from having a redundant signature matrix and nonsingular compressed matrix. The circles are assigned r and the squares are assigned $(\neq_2)$.

with signature matrix

$$M_r = M_{h'} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} M_{h''} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & m & m^2+1 & 0 \\ 0 & 1 & m & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

Note that the effect of the $(\neq_2)$ signatures is to reverse all four rows of $M_{h''}$ before multiplying it to the right of $M_{h'}$. Although this signature matrix is not redundant, every entry of Hamming weight 2 is nonzero since m is positive.

Now we claim that we can use r to interpolate the following signature R, for any nonzero value $t \in \mathbb{C}$, via the construction in Figure 7.8. Define $p^{\pm} = (m \pm \sqrt{m^2+4})/2$, $P = \begin{bmatrix} 1 & 1 \\ p^+ & p^- \end{bmatrix}$, and $T = P \begin{bmatrix} t & 0 \\ 0 & t^{-1} \end{bmatrix} P^{-1}$, where $t \in \mathbb{C}$ is any nonzero value. We define the signature R by its signature matrix

$$M_R = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & & & 0 \\ & & T & \\ 0 & & & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \tag{7.15}$$

Consider an instance Ω of Holant$(\neq_2 \mid \mathcal{F} \cup \{R\})$ where $\mathcal{F} = \{f, g\}$. Suppose that R appears n times in Ω. We construct from Ω a sequence of instances Ω_s of Holant$(\neq_2 \mid \mathcal{F})$ indexed by $s \geq 1$. We obtain Ω_s from Ω by replacing each occurrence of R with the gadget N_s in Figure 7.8 with r assigned to the circle vertices and $(\neq_2)$ assigned to the square vertices. In Ω_s, the edge corresponding to the ith significant index bit of N_s connects to the same location as the edge corresponding to the ith significant index bit of R in Ω.

The signature matrix of N_s is the sth power of the matrix obtained from M_r after reversing all rows, and then switching the first and last rows of the final

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

product, namely

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & m & 0 \\ 0 & m & m^2+1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^s$$

$$= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & m & 0 \\ 0 & m & m^2+1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & m & 0 \\ 0 & m & m^2+1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{s-1}.$$

The twist of the two input edges on the left side for the first copy of M_r switches the middle two rows, which is equivalent to a total reversal of all rows, followed by the switching of the first and last rows. The total reversals of rows for all subsequent $s-1$ copies of M_r are due to the presence of $(\neq_2)$ signatures.

After such reversals of rows, it is clear that the matrix is a direct sum of block matrices indexed by $\{00,11\} \times \{00,11\}$ and $\{01,10\} \times \{10,01\}$. Furthermore, in the final product, the block indexed by $\{00,11\} \times \{00,11\}$ is $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Thus in the gadget N_s, the only entries of M_{N_s} that vary with s are the four entries in the middle. These middle four entries of M_{N_s} form the 2×2 matrix $\begin{bmatrix} 1 & m \\ m & m^2+1 \end{bmatrix}^s$. Since $\begin{bmatrix} 1 & m \\ m & m^2+1 \end{bmatrix} = P \begin{bmatrix} \lambda_+ & 0 \\ 0 & \lambda_- \end{bmatrix} P^{-1}$, where $\lambda_\pm = (m^2+2 \pm m\sqrt{m^2+4})/2$ are the eigenvalues, we have

$$\begin{bmatrix} 1 & m \\ m & m^2+1 \end{bmatrix}^s = P \begin{bmatrix} \lambda_+^s & 0 \\ 0 & \lambda_-^s \end{bmatrix} P^{-1}.$$

As $\det \begin{bmatrix} 1 & m \\ m & m^2+1 \end{bmatrix} = 1$, $\lambda_+\lambda_- = 1$, so the eigenvalues are nonzero. Since m is positive, $\lambda_+ > \lambda_-$ and so neither λ_+ nor λ_- is a root of unity.

Now we determine the relationship between Holant_Ω and Holant_{Ω_s}. We can view our construction of Ω_s as first replacing M_R with the RHS of

$$M_R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & 0 \\ & & P & \\ 0 & & & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & t & 0 & 0 \\ 0 & 0 & t^{-1} & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & & & 0 \\ & & P^{-1} & \\ 0 & & & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

which does not change the Holant value, and then replacing the 4×4 matrix in the middle with the matrix

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & \lambda_+^s & 0 & 0 \\ 0 & 0 & \lambda_-^s & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

We stratify the assignments in Ω based on the assignments to the n occurrences of the signature matrix

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & t & 0 & 0 \\ 0 & 0 & t^{-1} & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}. \tag{7.16}$$

The inputs to this matrix are from $\{0,1\}^2 \times \{0,1\}^2$, which correspond to the four input bits. Recall the way rows and columns of a signature matrix are ordered from Definition 1.4. Thus, e.g., the entry t corresponds to the cyclic input bit pattern 0110 in counterclockwise order. We only need to consider the assignments that assign

- i many times the bit pattern 0110,
- j many times the bit pattern 1001, and
- k many times the bit patterns 0011 or 1100,

such that $i + j + k = n$, since any other assignment contributes 0 to the Holant sum. Let c_{ijk} be the sum over all such assignments of the product of evaluations (excluding (7.16), but including the contributions from the block matrices containing P and P^{-1}) on Ω. Then

$$\text{Holant}_\Omega = \sum_{i+j+k=n} t^{i-j} c_{ijk}$$

and the value of the Holant on Ω_s, for $s \geq 1$, is

$$\text{Holant}_{\Omega_s} = \sum_{i+j+k=n} \lambda_+^{si} \lambda_-^{sj} c_{ijk} = \sum_{i+j+k=n} \lambda_+^{s(i-j)} c_{ijk}.$$

This Vandermonde system does not have full rank. However, we can define for $-n \leq \ell \leq n$,

$$c'_\ell = \sum_{\substack{i-j=\ell \\ i+j+k=n}} c_{ijk}.$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Then the Holant of Ω is

$$\text{Holant}_{\Omega} = \sum_{-n \le \ell \le n} t^{\ell} c'_{\ell}$$

and the Holant of Ω_s is

$$\text{Holant}_{\Omega_s} = \sum_{-n \le \ell \le n} \lambda_+^{s\ell} c'_{\ell}.$$

Now this Vandermonde has full rank because λ_+ is neither zero nor a root of unity. Therefore, we can solve for the unknowns $\{c'_{\ell}\}$ in polynomial time and obtain the value of Holant_{Ω}. This proves the claim that we can interpolate the signature R in (7.15), for any nonzero $t \in \mathbb{C}$.

Let $t = (\sqrt{m^2+8} + \sqrt{m^2+4})/2$ so $t^{-1} = (\sqrt{m^2+8} - \sqrt{m^2+4})/2$. Let $a = (\sqrt{m^2+8} - m)/2$ and $b = (\sqrt{m^2+8} + m)/2$; then $ab = 2$ and both a and b are nonzero. One can verify that

$$P \begin{bmatrix} t & 0 \\ 0 & t^{-1} \end{bmatrix} P^{-1} = \begin{bmatrix} a & 1 \\ 1 & b \end{bmatrix}.$$

Thus, the signature matrix for R for this particular choice of t is

$$M_R = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & a & 1 & 0 \\ 0 & 1 & b & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

After a counterclockwise rotation of 90° on the edges of R, we have a signature R' with a redundant signature matrix

$$M_{R'} = \begin{bmatrix} 0 & 0 & 0 & a \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ b & 0 & 0 & 0 \end{bmatrix}.$$

Its compressed signature matrix

$$\widetilde{M_{R'}} = \begin{bmatrix} 0 & 0 & a \\ 0 & 2 & 0 \\ b & 0 & 0 \end{bmatrix}$$

is nonsingular. After a holographic transformation by Z^{-1}, where $Z = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$, the binary DISEQUALITY ($\neq_2$) is transformed to the binary

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

EQUALITY ($=_2$). Thus Holant($\neq_2 | R'$) is transformed to Holant($=_2 | Z^{\otimes 4} R'$), which is the same as Holant($Z^{\otimes 4} R'$). We conclude that this Holant problem Holant($Z^{\otimes 4} R'$) is #P-hard by Corollary 7.27, where we choose $T = Z^{-1}$.

To summarize, Holant($Z^{\otimes 4} R'$) is #P-hard implies that Holant($\neq_2 | R'$) is #P-hard, which implies that Holant($\neq_2 | R$) is #P-hard as R' is a rotated form of R, which implies that Holant($\neq_2 | r$) is #P-hard since it can interpolate Holant($\neq_2 | R$) for all choices of t, which implies that Holant($\neq_2 | \{f, g\}$) is #P-hard since it can directly realize Holant($\neq_2 | r$). □

We note that the planar problem Pl-Holant($\neq_2 | \{f, g\}$) is computable in P by matchgates. The twist introduced in Figure 7.8 is unavoidable; otherwise we would have collapsed #P and P, as all other steps of the reduction chain are planar, all the way back to the #P-hard planar #EULERIAN-ORIENTATIONS problem Pl-Holant($\neq_2 | \mathcal{O}$).

7.6.2 Combining Vanishing Signatures with Others

With Corollary 7.25, Corollary 7.27, and Lemma 7.30 in hand, we consider Holant problems when vanishing signatures are combined with other signatures. We begin with degenerate symmetric signatures, which include unary signatures. Note that in the following lemma proving #P-hardness, to avoid the tractable case 5 in Theorem 7.19, it is necessary to require $\mathrm{rd}^{\sigma}(f) \geq 2$, for otherwise $f \in \mathscr{R}_2^{\sigma}$. We prove that in this case, combining a vanishing signature with any degenerate or unary signature of not exactly the same type produces #P-hardness. For future references we prove the stronger statement that it is #P-hard even restricted to planar graphs.

Lemma 7.31. *Let $f \in \mathscr{V}^{\sigma}$ be a symmetric signature of arity n with $\mathrm{rd}^{\sigma}(f) = d \geq 2$, where $\sigma \in \{+, -\}$. Suppose $v = u^{\otimes m}$ is a symmetric degenerate signature for some unary signature u and some integer $m \geq 1$. If u is not a multiple of $[1, \sigma \mathrm{i}]$, then* Pl-Holant(f, v) *is #P-hard.*

Proof. We consider $\sigma = +$ since the other case is similar. Since $f \in \mathscr{V}^{+}$, we have $n > 2d \geq 4$. Under a holographic transformation by $Z = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$, we have

$$\text{Pl-Holant}(f, v) \equiv \text{Pl-Holant}(\neq_2 | \hat{f}, [a, b]^{\otimes m}),$$

where $\hat{f} = \left(Z^{-1}\right)^{\otimes n} f$ and $[a, b]^{\otimes m} = \left(Z^{-1}\right)^{\otimes m} v = \left(Z^{-1} u\right)^{\otimes m}$. Since u is not a multiple of $[1, \mathrm{i}]$, we have $b \neq 0$. Moreover, $\hat{f} = [\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0]$ with $\hat{f}_d \neq 0$ by Lemma 7.8.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

By $d-2$ self-loops on $\hat{f}$ via $\neq_2$, we get $\widehat{f'} = [\hat{f}_{d-2}, \hat{f}_{d-1}, \hat{f}_d, 0, \ldots, 0]$ of arity $n-2d+4$. This $\widehat{f'}$ is on the RHS in Pl-Holant($\neq_2 | \hat{f}, [a,b]^{\otimes m}$). With two more self-loops, we get $[1,0]^{\otimes(n-2d)}$, also on the RHS.

We claim that we can use $[1,0]^{\otimes(n-2d)}$ and $[a,b]^{\otimes m}$ to create $[a,b]^{\otimes(n-2d)}$ on the RHS. If $n-2d > m$, then we connect $[a,b]^{\otimes m}$ to $[1,0]^{\otimes(n-2d)}$ via $\neq_2$ to get $[1,0]^{\otimes(n-2d-m)}$ up to a nonzero factor $b^m \neq 0$. We repeat this process until we get a tensor power $[1,0]^{\otimes \ell}$ for some $\ell \leq m$. If $m > n-2d$, we can get a tensor power $[a,b]^{\otimes \ell}$ for some $\ell \leq n-2d$ by a similar process. Repeat this process, which is a subtractive Euclidean algorithm, and we end up getting both $[1,0]^{\otimes t}$ and $[a,b]^{\otimes t}$, where $t = \gcd(m, n-2d)$. Then we combine $\frac{n-2d}{t}$ copies of $[a,b]^{\otimes t}$ to get $[a,b]^{\otimes(n-2d)}$.

Now connecting $[a,b]^{\otimes(n-2d)}$ back to $\widehat{f'}$ via $\neq_2$, gives $\widehat{f''} = b^{n-2d}[\widehat{f}_{d-2}, \widehat{f}_{d-1}, \widehat{f}_d, 0, 0]$ of arity 4. Notice that Pl-Holant($\neq_2 | [\widehat{f}_{d-2}, \widehat{f}_{d-1}, \widehat{f}_d, 0, 0]$) $\equiv_{\mathrm{T}}$ Pl-Holant($\neq_2 | [0,0,1,0,0]$). The latter problem is the Eulerian orientation problem over planar 4-regular graphs, which is #P-hard by Theorem 6.32. Thus, Pl-Holant(f, v) is #P-hard. □

Next we consider what binary signatures can be combined with a vanishing signature. Before we prove Lemma 7.33, we prove a simple interpolation lemma.

Lemma 7.32. *Let $x \in \mathbb{C}$. If $x \neq 0$, then for any set $\mathcal{F}$ containing $[x,1,0]$, we have*

$$\text{Holant}(\neq_2 | \mathcal{F} \cup \{[v,1,0]\}) \leq_{\mathrm{T}} \text{Holant}(\neq_2 | \mathcal{F})$$

for any $v \in \mathbb{C}$.

Proof. Consider an instance Ω of Holant($\neq_2 | \mathcal{F} \cup \{[v,1,0]\}$). Suppose that $[v,1,0]$ appears n times in Ω. We stratify the assignments in Ω based on the assignments to $[v,1,0]$. We only need to consider assignments of Hamming weight ≤ 1 since an assignment of Hamming weight 2 contributes a factor of 0. We have

$$\text{Holant}_\Omega = \sum_{i=0}^{n} v^i c_i,$$

where c_i is the sum, over all assignments with exactly i many $[v,1,0]$'s in Ω assigned with a local assignment 00, and $n-i$ many $[v,1,0]$'s in Ω assigned 01 or 10, of the product of evaluations of all other signatures on Ω.

We construct from Ω a sequence of instances Ω_s of Holant($\neq_2 | \mathcal{F}$) indexed by $s \geq 1$, by replacing each occurrence of $[v,1,0]$ with a gadget g_s created from s copies of $[x,1,0]$, connected sequentially but with $(\neq_2) = [0,1,0]$ between

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

each sequential pair. The signature of g_s is $[sx, 1, 0]$, which can be verified by the matrix product

$$\left(\begin{bmatrix} x & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right)^{s-1}\begin{bmatrix} x & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & x \\ 0 & 1 \end{bmatrix}^{s-1}\begin{bmatrix} x & 1 \\ 1 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & (s-1)x \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} sx & 1 \\ 1 & 0 \end{bmatrix}.$$

The Holant on Ω_s is

$$\text{Holant}_{\Omega_s} = \sum_{i=0}^{n} (sx)^i c_i.$$

For $s \geq 1$, this gives a coefficient matrix that is Vandermonde. Since x is nonzero, sx is distinct for each s. Therefore, the Vandermonde system has full rank. We can solve for the unknowns c_i in polynomial time and obtain the value of Holant_Ω. □

For binary signatures we have the following lemma. Owing to the tractable case 4 in Theorem 7.19, the condition $g \notin \mathscr{R}_2^\sigma$ in Lemma 7.33 is needed.

Lemma 7.33. *Let $f \in \mathscr{V}^\sigma$ be a symmetric nondegenerate signature where $\sigma \in \{+, -\}$. If $g \notin \mathscr{R}_2^\sigma$ is a symmetric nondegenerate binary signature, then* $\text{Holant}(\{f, g\})$ *is* #P*-hard.*

Proof. We consider $\sigma = +$ since the other case is similar. A unary signature is degenerate. If a binary symmetric signature f is vanishing, then $\text{vd}^+(f) > 1$, and so $\text{vd}^+(f) \geq 2$; therefore f is also degenerate. Since we assume f is nondegenerate, $\text{arity}(f) \geq 3$.

We prove the lemma by induction on the arity of f. There are two base cases, $\text{arity}(f) = 3$ and $\text{arity}(f) = 4$. However, the arity 3 case is easily reduced to the arity 4 case. We show this first, and then show that the lemma holds for the arity 4 case.

Assume $\text{arity}(f) = 3$. Since $f \in \mathscr{V}^+$, we have $\text{rd}^+(f) < 3/2$, thus $f \in \mathscr{R}_2^+$. From $\text{rd}^+(f) \leq 1$ we have $\text{vd}^+(f) \geq 2$. On the other hand, f is nondegenerate, and so $\text{vd}^+(f) < 3$, thus $\text{vd}^+(f) = 2$ and $\text{rd}^+(f) = 1$.

We connect two copies of f together by one edge to get an arity 4 signature f'. By construction, it may not appear that f' is a symmetric signature; but we show that f' is in fact *symmetric*, nondegenerate and vanishing. It is clearly a vanishing signature, since f is vanishing. Consider the Z transformation, under which f is transformed into $\hat{f} = (Z^{-1})^{\otimes 3} f = [t, 1, 0, 0]$ for some t, up to a nonzero constant. The $(=_2)$ on the connecting edge between the two copies of f

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

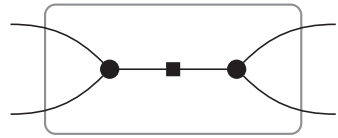


Figure 7.9. The circles are assigned $[t, 1, 0, 0]$ and the square is assigned $(\neq_2)$.

is transformed into $(=_2)Z^{\otimes 2} = (\neq_2)$. In the bipartite setting, our construction is the same as the gadget in Figure 7.9. One can verify that the resulting signature is $\widehat{f'} = [2t, 1, 0, 0, 0]$. The crucial observation is that it takes the same value 0 on inputs 1010 and 1100, where the left two bits are input to one copy of f and the right two bits are for another. The corresponding signature $f' = Z^{\otimes 4}\widehat{f'}$ is symmetric, nondegenerate, with $\mathrm{rd}^+(f') = 1$ and vanishing.

Next we consider the base case of $\mathrm{arity}(f) = 4$. Since $f \in \mathscr{V}^+$, we have $\mathrm{vd}^+(f) > 2$ and $\mathrm{rd}^+(f) < 2$. Since f is nondegenerate we have $\mathrm{rd}^+(f) \neq -1, 0$, hence $\mathrm{rd}^+(f) = 1$ and $\mathrm{vd}^+(f) = 3$. Also by assumption, the given binary $g \notin \mathscr{R}_2^+$, we have $\mathrm{rd}^+(g) = 2$. Once again, consider the holographic transformation by Z. By $\mathrm{rd}^+(g) = 2$, the transformation $(Z^{-1})^{\otimes 2}g$ has the form $[a, b, c]$ with $c \neq 0$ by Lemma 7.8. This gives

$$\begin{aligned}\mathrm{Holant}(=_2 \mid \{f, g\}) &\equiv_{\mathrm{T}} \mathrm{Holant}((=_2)Z^{\otimes 2} \mid \{(Z^{-1})^{\otimes 4}f, (Z^{-1})^{\otimes 2}g\}) \\ &\equiv_{\mathrm{T}} \mathrm{Holant}(\neq_2 \mid \{\hat{f}, \hat{g}\}),\end{aligned}$$

where up to a nonzero constant, $\hat{f} = [t, 1, 0, 0, 0]$ and $\hat{g} = [a, b, 1]$, for some $t, a, b \in \mathbb{C}$. We have $a - b^2 \neq 0$ since g is nondegenerate.

Our next goal is to show that we can realize a signature of the form $[c, 0, 1]$ where $c \neq 0$. If $b = 0$, then $\hat{g}$ is what we want since in this case $a = a - b^2 \neq 0$. Now we assume $b \neq 0$.

Connecting both edges of $\hat{g}$ to $\hat{f}$ via $(\neq_2)$, we get $[t + 2b, 1, 0]$. If $t \neq -2b$, then by Lemma 7.32, we can interpolate any binary signature of the form $[v, 1, 0]$. If $t = -2b$, then we connect two copies of $\hat{g}$ via $(\neq_2)$, and get $\widehat{g'} = [2ab, a + b^2, 2b]$. Connecting this $\widehat{g'}$ to $\hat{f}$ via $(\neq_2)$, we get $[2(a - b^2), 2b, 0]$, using $t = -2b$. Since $a \neq b^2$ and $b \neq 0$, we can interpolate any $[v, 1, 0]$ again by Lemma 7.32.

Hence, we can have any signature $[v, 1, 0]$, where $v \in \mathbb{C}$ is for us to choose. We construct the gadget in Figure 7.10 with the circles assigned $[v, 1, 0]$, the squares assigned $(\neq_2)$, and the triangle assigned $[a, b, 1]$. The resulting gadget has signature $[a + 2bv + v^2, b + v, 1]$, which can be verified by the

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press



Figure 7.10. A sequence of binary gadgets that forms another binary gadget. The circles are assigned $[v, 1, 0]$, the squares are assigned $\neq_2$, and the triangle is assigned $[a, b, 1]$.

matrix product

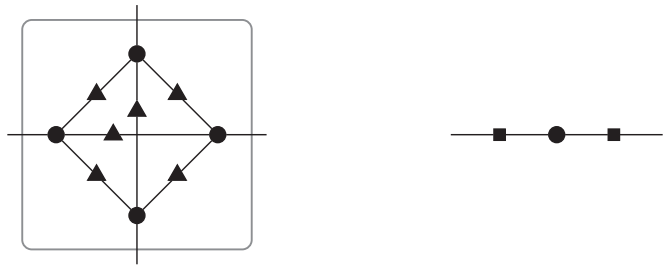
$$\begin{bmatrix} v & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ b & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} v & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} a + 2bv + v^2 & b + v \\ b + v & 1 \end{bmatrix}.$$

By setting $v = -b$, we get $[c, 0, 1]$, where $c = a - b^2 \neq 0$.

Now we construct a signature $\hat{h}$ by the tetrahedron gadget given in Figure 7.11, using the signature $[c, 0, 1]$, in two steps. First in Figure 7.11b we assign $[c, 0, 1]$ to the circle vertex of arity two and $(\neq_2)$ to the two square vertices. The resulting signature $[1, 0, c]$ is then used in Figure 7.11a. We assign $[1, 0, c]$ to the six internal edges labeled with triangles, and assign $\hat{f} = [t, 1, 0, 0, 0]$ to the four circle vertices of arity four in Figure 7.11a. By construction, $\hat{h}$ is a symmetric signature, and in fact

$$\hat{h} = [t^4 + 6ct^2 + 3c^2, t^3 + 3ct, t^2 + c, t, 1].$$

In fact the special case with $c = 1$ has already been used in the proof of Lemma 7.28. We can compute these signature entries similarly to a matchgate. The internal edge function $[1, 0, c]$ is a Gen-Eq signature, and the signature $\hat{f}$



(a) The tetrahedron gadget with edge signatures given in (b).

(b) The gadget representing an edge labeled by a triangle in (a).

Figure 7.11. The tetrahedron gadget with each triangle replaced by the edge in (b), where the circle is assigned $[c, 0, 1]$ and the squares are assigned $\neq_2$. The four circles in (a) are assigned $[t, 1, 0, 0, 0]$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

on the four circle vertices is a weighted version of the matching function AT-MOST-ONE. For example, we may verify that $\hat{h}_0 = t^4 + 6ct^2 + 3c^2$ as follows: For the external input pattern 0000, the term t^4 comes from the empty matching among six internal edges, the term $6ct^2$ comes from the six matchings of size one, each matching edge having weight c, and the term $3c^2$ comes from the three perfect matchings of size two. The other entries can be computed similarly.

The compressed signature matrix of $\hat{h}$ is

$$\widetilde{M}_{\hat{h}} = \begin{bmatrix} 3c^2 + 6ct^2 + t^4 & 2(3ct + t^3) & c + t^2 \\ 3ct + t^3 & 2(c + t^2) & t \\ c + t^2 & 2t & 1 \end{bmatrix}$$

and quite amazingly its determinant is $4c^3 \neq 0$ (check it!). Thus $\widetilde{M}_{\hat{h}}$ is nonsingular. After a holographic transformation by Z^{-1}, where $Z = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$, the binary DISEQUALITY function ($\neq_2$) is transformed to the binary EQUALITY ($=_2$). Thus Holant($\neq_2 | \hat{h}$) is transformed to Holant($=_2 \mid Z^{\otimes 4}\hat{h}$), which is the same as Holant($Z^{\otimes 4}\hat{h}$). Then, by Corollary 7.27, we are done.

Now we do the induction step. The idea is to produce from f and g another symmetric nondegenerate signature $f' \in \mathscr{V}^+$ that has a smaller arity than f. Assume f has arity $n \geq 5$. Since f is nondegenerate, $\mathrm{rd}^+(f) \neq -1, 0$. Under the Z transformation, $f = Z^{\otimes n}\hat{f}$, $g = Z^{\otimes 2}\hat{g}$, where $\hat{f} = [\hat{f}_0, \ldots, \hat{f}_d, 0, \ldots, 0]$, $d = \mathrm{rd}^+(f) < n/2$, and $\hat{g} = [a, b, 1]$ up to a nonzero scalar.

First suppose $\mathrm{rd}^+(f) = 1$. By Lemma 7.12 we can get $f' = \langle f, g \rangle$ by connecting both edges of g to f. This is a symmetric signature of arity $n - 2 \geq 3$ with $\mathrm{rd}^+(f') = \mathrm{rd}^+(f) = 1$. In particular f' is nondegenerate. Since $\mathrm{rd}^+(f') < \mathrm{arity}(f')/2$, $f' \in \mathscr{V}^+$ by definition. So we are done by induction.

We now suppose $\mathrm{rd}^+(f) = d \geq 2$. Since f is nondegenerate it is certainly nonzero. Since it is vanishing, certainly $\mathrm{vd}^+(f) > 0$. Hence we may apply Lemma 7.13. Let f' be obtained from f by a self-loop, then $\mathrm{rd}^+(f') = d - 1 \geq 1$ and $\mathrm{arity}(f') = n - 2$. Clearly f' is still vanishing. We have $f' = Z^{\otimes(n-2)}[\widehat{f'}_0, \ldots, \widehat{f'}_{d-1}, 0, \ldots, 0]$ where $\widehat{f'}_{d-1} \neq 0$. Since $d - 1 \geq 1$, f' is nondegenerate. The proof is complete by induction. □

Finally, we consider a pair of vanishing signatures of the opposite type, both of arity ≥ 3. We show that opposite types of vanishing signatures cannot mix. More formally, vanishing signatures of opposite types, when put together, lead to #P-hardness.

Lemma 7.34. *Let $f \in \mathscr{V}^+$ and $g \in \mathscr{V}^-$ be nondegenerate symmetric signatures both of arity ≥ 3. Then* Holant($\{f, g\}$) *is* #P*-hard.*

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Proof. Let $\mathrm{rd}^+(f) = d, \mathrm{rd}^-(g) = d', \mathrm{arity}(f) = n$ and $\mathrm{arity}(g) = n'$, then $2d < n$ and $2d' < n'$. We perform a holographic transformation by $Z = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$,

$$\begin{aligned} \mathrm{Holant}(=_2 \mid \{f, g\}) &\equiv_{\mathrm{T}} \mathrm{Holant}((=_2)Z^{\otimes 2} \mid \{(Z^{-1})^{\otimes n} f, (Z^{-1})^{\otimes n'} g\}) \\ &\equiv_{\mathrm{T}} \mathrm{Holant}((\neq_2) \mid \{\hat{f}, \hat{g}\}), \end{aligned}$$

where $\hat{f} = (Z^{-1})^{\otimes n} f$ and $\hat{g} = (Z^{-1})^{\otimes n'} g$. By Lemma 7.8 $\hat{f}$ and $\hat{g}$ take the form $\hat{f} = [\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0]$ and $\hat{g} = [0, \ldots, 0, \hat{g}_{n'-d'}, \ldots, \hat{g}_{n'}]$, where $\hat{f}_d \neq 0$ and $\hat{g}_{n'-d'} \neq 0$.

If $d \geq 2$, we will use Lemma 7.31. We can do d' many self-loops on $\hat{g}$ via $(\neq_2)$, getting $\hat{g}' = [0, \ldots, 0, \hat{g}_{n'-d'}]$ of arity $n' - 2d' \geq 1$. If we define $g' = Z^{\otimes(n'-2d')}\hat{g}'$, then $g' = [1, -\mathrm{i}]^{\otimes(n'-2d')}$ up to a nonzero constant. This g' is constructible in the RHS of $\mathrm{Holant}(=_2 \mid \{f, g\})$. Then applying Lemma 7.31 on the pair $\{f, g'\}$, we get $\mathrm{Holant}(f, g)$ is #P-hard. (In fact, $\mathrm{Pl\text{-}Holant}(f, g)$ is also #P-hard in this case.) If $d' \geq 2$, we can similarly get $[1, \mathrm{i}]^{\otimes(n-2d)}$ and apply Lemma 7.31. Thus we may assume that $d = d' = 1$.

Therefore, up to a nonzero constant, $\hat{f} = [a, 1, 0, \ldots, 0]$ and $\hat{g} = [0, \ldots, 0, 1, b]$, for some $a, b \in \mathbb{C}$. We show that it is always possible to get two such signatures of the same arity $n^* = \min\{n, n'\}$. Suppose $n > n'$. We form a loop from $\hat{f}$, where the loop is really a path consisting of one vertex and two edges, with the vertex assigned the signature $(\neq_2)$. It is easy to see that this signature is the degenerate signature $2[1, 0]^{\otimes(n-2)}$. Similarly, we can form a loop from $\hat{g}$ and can get $2[0, 1]^{\otimes(n'-2)}$. Thus we have both $[1, 0]^{\otimes(n-2)}$ and $[0, 1]^{\otimes(n'-2)}$. We can connect all $n' - 2$ edges of the second to the first, connected by $(\neq_2)$. This gives $[1, 0]^{\otimes(n-n')}$. We can continue subtracting the smaller arity from the larger one. We continue this process in a subtractive version of the Euclidean algorithm, and end up with both $[1, 0]^{\otimes t}$ and $[0, 1]^{\otimes t}$, where $t = \gcd(n - 2, n' - 2) = \gcd(n - n', n' - 2)$. In particular, $t \mid n - n'$ and by taking $(n - n')/t$ many copies of $[0, 1]^{\otimes t}$, we can get $[0, 1]^{\otimes(n-n')}$. Connecting this back to $\hat{f}$ via $(\neq_2)$, we get a symmetric signature of arity n' consisting of the first $n' + 1$ entries of $\hat{f}$. A similar proof works when $n' > n$.

Thus without loss of generality, we may assume we have two signatures $[a, 1, 0, \ldots, 0]$ and $[0, \ldots, 0, 1, b]$, both of arity $n^* \geq 3$. We also have $[1, 0]^{\otimes t}$ and $[0, 1]^{\otimes t}$ thus both $[1, 0]^{\otimes(n^*-2)}$ and $[0, 1]^{\otimes(n^*-2)}$. Connect $[0, 1]^{\otimes(n^*-2)}$ to $[a, 1, 0, \ldots, 0]$ of arity n^* via $(\neq_2)$ we get $\hat{h} = [a, 1, 0]$. For $a \neq 0$, translating this back by Z, we have a binary signature h together with the given $g \in \mathscr{V}^-$. We claim that $h \notin \mathscr{R}_2^-$. Indeed if it were the case that $h \in \mathscr{R}_2^-$, then $\mathrm{rd}^-(h) \leq 1$ and thus $\hat{h} = [0, *, *]$, a contradiction. By Lemma 7.33, $\mathrm{Holant}(\{g, h\})$ is

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

#P-hard, and thus Holant($\{f, g\}$) is #P-hard. A similar proof works for the case $b \neq 0$.

The only case left is when $\hat{f} = [0, 1, 0, \ldots, 0]$, and $\hat{g} = [0, \ldots, 0, 1, 0]$, both of arity $n^* \geq 3$. This is #P-hard by Lemma 7.30. □

Remark: Notice that in the last case of the proof of Lemma 7.34, $\hat{f} = [0, 1, 0, \ldots, 0]$, and $\hat{g} = [0, \ldots, 0, 1, 0]$. The planar Holant problem Pl-Holant($(\neq_2) \mid \{\hat{f}, \hat{g}\}$) in this case is computable in polynomial time, and therefore so is Pl-Holant(f, g), because $(\neq_2)$, $\hat{f}, \hat{g}$ are all matchgate signatures. One can prove that, excluding this case, the planar Holant problem Pl-Holant(f, g) in Lemma 7.34 is also #P-hard. The additional condition is that either f is not the Z-transform of the matchgate signature $[0, 1, 0, \ldots, 0]$ or g is not the Z-transform of the matchgate signature $[0, \ldots, 0, 1, 0]$, up to a constant. In Chapter 9 these will be called $\mathscr{M}_4$ signatures.

7.7 $\mathscr{A}$-transformable and $\mathscr{P}$-transformable Signatures

In this section, we study $\mathscr{A}$-transformable signatures and $\mathscr{P}$-transformable signatures. We denote by $\alpha = \frac{1+\mathrm{i}}{\sqrt{2}} = \sqrt{\mathrm{i}} = e^{\frac{\pi \mathrm{i}}{4}}$ and use $\mathbf{O}_2(\mathbb{C})$ to denote the group of 2×2 orthogonal matrices over $\mathbb{C}$. Recall that $\mathscr{F}_{123} = \mathscr{F}_1 \cup \mathscr{F}_2 \cup \mathscr{F}_3$, where $\mathscr{F}_1$, $\mathscr{F}_2$, and $\mathscr{F}_3$ are defined in Chapter 3, Section 3.2.

7.7.1 Characterization of $\mathscr{A}$- and $\mathscr{P}$-transformable Signatures

Recall that by definition, if a set of signatures $\mathcal{F}$ is $\mathscr{A}$-transformable (resp. $\mathscr{P}$-transformable), then the binary EQUALITY $(=_2)$ must be simultaneously transformed into $\mathscr{A}$ (resp. $\mathscr{P}$) along with $\mathcal{F}$. We first characterize all possible matrices for such transformations by considering only the transformation of $(=_2)$. Although there are many binary signatures in $\mathscr{A} \cup \mathscr{P}$, it turns out that it is sufficient to consider only three signatures. The following proposition is best understood in terms of a group action, its stabilizer and orbits.

Proposition 7.35. *Let $T \in \mathbb{C}^{2\times 2}$ be a matrix. Then the following hold:*

1. $[1, 0, 1]T^{\otimes 2} = [1, 0, 1]$ *iff* $T \in \mathbf{O}_2(\mathbb{C})$;
2. $[1, 0, 1]T^{\otimes 2} = [1, 0, \mathrm{i}]$ *iff there exists an* $H \in \mathbf{O}_2(\mathbb{C})$ *such that* $T = H \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix}$;
3. $[1, 0, 1]T^{\otimes 2} = [0, 1, 0]$ *iff there exists an* $H \in \mathbf{O}_2(\mathbb{C})$ *such that* $T = \frac{1}{\sqrt{2}} H \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Proof. The first item is clear since

$$[1,0,1]T^{\otimes 2} = [1,0,1] \iff T^{\mathrm{T}} I_2 T = I_2 \iff T^{\mathrm{T}} T = I_2,$$

the definition of a 2×2 orthogonal matrix T.

We let $M_1 = I_2$, $M_2 = \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix}$ and $M_3 = Z = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$, let $T_j = HM_j$ (for $j = 1, 2, 3$), where $H \in \mathbf{O}_2(\mathbb{C})$. Then

$$[1,0,1]T_j^{\otimes 2} = [1,0,1](HM_j)^{\otimes 2} = [1,0,1]M_j^{\otimes 2} = f_j,$$

where $f_1 = [1,0,1]$, $f_2 = [1,0,\mathrm{i}]$ and $f_3 = [0,1,0]$.

Conversely, suppose that $[1,0,1](T_j)^{\otimes 2} = f_j$. Then we have

$$[1,0,1]\big(T_j M_j^{-1}\big)^{\otimes 2} = f_j\big(M_j^{-1}\big)^{\otimes 2} = [1,0,1],$$

so $H = T_j M_j^{-1} \in \mathbf{O}_2(\mathbb{C})$ by case 1. Thus $T_j = HM_j$ as desired. □

We also need the following proposition; the proof is direct.

Proposition 7.36. *If a symmetric signature $f = [f_0, f_1, \ldots, f_n]$ can be expressed in the form $f = a[1,\lambda]^{\otimes n} + b[1,\mu]^{\otimes n}$, for some $a, b, \lambda, \mu \in \mathbb{C}$, then the f_k's satisfy the recurrence relation $f_{k+2} = (\lambda+\mu)f_{k+1} - \lambda\mu f_k$ for $0 \le k \le n-2$.*

Recall the stabilizer group $\mathrm{Stab}(\mathscr{A}) = \mathbb{C}^* \cdot \langle D, H_2 \rangle$ that was determined in Lemma 6.13, where $D = \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{i} \end{bmatrix}$ and $H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Lemma 7.37. *Let $\mathcal{F}$ be a set of signatures. Then $\mathcal{F}$ is $\mathscr{A}$-transformable under T iff $\mathcal{F}$ is $\mathscr{A}$-transformable under any $T' \in T\,\mathrm{Stab}(\mathscr{A})$.*

Proof. Sufficiency is trivial since $I_2 \in \mathrm{Stab}(\mathscr{A})$. If $\mathcal{F}$ is $\mathscr{A}$-transformable under T, then by definition, we have $(=_2)T^{\otimes 2} \in \mathscr{A}$ and $\mathcal{F}' = T^{-1}\mathcal{F} \subseteq \mathscr{A}$. Let $T' = TM \in T\,\mathrm{Stab}(\mathscr{A})$ for any $M \in \mathrm{Stab}(\mathscr{A})$. It then follows that $(=_2)T'^{\otimes 2} = (=_2)T^{\otimes 2}M^{\otimes 2} \in \mathscr{A}M = \mathscr{A}$ and $T'^{-1}\mathcal{F} = M^{-1}\mathcal{F}' \subseteq M^{-1}\mathscr{A} = \mathscr{A}$. Therefore $\mathcal{F}$ is $\mathscr{A}$-transformable under any $T' \in T\,\mathrm{Stab}(\mathscr{A})$. □

After restricting by Proposition 7.35 and normalizing by Lemma 7.37, one only needs to check a small subset of $\mathbf{GL}_2(\mathbb{C})$ to determine if $\mathcal{F}$ is $\mathscr{A}$-transformable.

Lemma 7.38. *Let $\mathcal{F}$ be a set of signatures. Then $\mathcal{F}$ is $\mathscr{A}$-transformable iff there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $\mathcal{F} \subseteq H\mathscr{A}$ or $\mathcal{F} \subseteq H \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \mathscr{A}$.*

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Proof. Sufficiency is easily verified by checking that $(=_2)$ is transformed into $\mathscr{A}$ by H and by $H\begin{bmatrix}1 & 0\\ 0 & \alpha\end{bmatrix}$ in respective cases.

If $\mathcal{F}$ is $\mathscr{A}$-transformable, then by definition, there exists a matrix T such that $(=_2)T^{\otimes 2} \in \mathscr{A}$ and $T^{-1}\mathcal{F} \subseteq \mathscr{A}$. Since $(=_2)$ is nondegenerate and symmetric, $(=_2)T^{\otimes 2} \in \mathscr{A}$ is equivalent to $(=_2)T^{\otimes 2} \in \mathscr{F}_{123}$.

Any signature in $\mathscr{F}_{123}$ is expressible as $c(v_1^{\otimes n} + \mathrm{i}^t v_2^{\otimes n})$, where $t \in \{0, 1, 2, 3\}$ and (v_1, v_2) is a pair of vectors in the set

$$\left\{\left([1, \quad 0], [0, \quad 1]\right), \left([1, \quad 1], [1, \quad -1]\right), \left([1, \quad \mathrm{i}], [1, \quad -\mathrm{i}]\right)\right\}.$$

We use D and $H_2 \in \mathrm{Stab}(\mathscr{A})$ to normalize $\mathscr{F}_{123}$ by Lemma 7.37. In particular, $\mathscr{F}_1 = \mathscr{F}_2 H_2$ and $\mathscr{F}_1 = \mathscr{F}_3(D^{-1}H_2)$. Furthermore, the binary signatures in $\mathscr{F}_1$ are just the four signatures $[1, 0, 1]$, $[1, 0, \mathrm{i}]$, $[1, 0, -1]$, and $[1, 0, -\mathrm{i}]$ up to a scalar. We also normalize these four by D, as $[1, 0, 1] = [1, 0, -1]D^{\otimes 2}$ and $[1, 0, \mathrm{i}] = [1, 0, -\mathrm{i}]D^{\otimes 2}$. Hence $\mathcal{F}$ being $\mathscr{A}$-transformable implies that there exists a matrix T such that $(=_2)T^{\otimes 2} \in \{[1, 0, 1], [1, 0, \mathrm{i}]\}$ and $T^{-1}\mathcal{F} \subseteq \mathscr{A}$. Now we apply Proposition 7.35.

1. If $(=_2)T^{\otimes 2} = [1, 0, 1]$, then by case 1 of Proposition 7.35, we have $T \in \mathbf{O}_2(\mathbb{C})$. Therefore $\mathcal{F} \subseteq H\mathscr{A}$, where $H = T \in \mathbf{O}_2(\mathbb{C})$.
2. If $(=_2)T^{\otimes 2} = [1, 0, \mathrm{i}]$, then by case 2 of Proposition 7.35, there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $T = H\begin{bmatrix}1 & 0\\ 0 & \alpha\end{bmatrix}$. Therefore $\mathcal{F} \subseteq T\mathscr{A} = H\begin{bmatrix}1 & 0\\ 0 & \alpha\end{bmatrix}\mathscr{A}$ where $H \in \mathbf{O}_2(\mathbb{C})$.

This completes the proof. □

Using these two lemmas, we can characterize all $\mathscr{A}$-transformable signatures. We first define the three sets $\mathscr{A}_1$, $\mathscr{A}_2$, and $\mathscr{A}_3$.

Definition 7.39. A symmetric signature f of arity n is in $\mathscr{A}_1$ if there exists an $H \in \mathbf{O}_2(\mathbb{C})$ and a nonzero constant c such that

$$f = cH^{\otimes n}\left(\begin{bmatrix}1\\ 1\end{bmatrix}^{\otimes n} + \beta\begin{bmatrix}1\\ -1\end{bmatrix}^{\otimes n}\right),$$

where $\beta = \alpha^{tn+2r}$ for some $r \in \{0, 1, 2, 3\}$ and $t \in \{0, 1\}$.

When such an H exists, we say that $f \in \mathscr{A}_1$ with transformation H. If $f \in \mathscr{A}_1$ with I_2, then we say f is in the canonical form of $\mathscr{A}_1$. If f is in the canonical

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

form of $\mathscr{A}_1$, then by Proposition 7.36, for any $0 \le k \le n-2$, we have $f_{k+2} = f_k$ and one of the following holds:

- $f_0 = 0$, or
- $f_1 = 0$, or
- $f_1 = \pm \mathrm{i} f_0 \neq 0$, or
- n is odd and $f_1 = \pm(1 \pm \sqrt{2})\mathrm{i} f_0 \neq 0$ (all four sign choices are permissible).

Notice that when n is odd and $t = 1$ in Definition 7.39, it has some complication as described by the factor α^{n+2r} ($r \in \{0, 1, 2, 3\}$), which runs through $\alpha \mathrm{i}^s$ ($s \in \{0, 1, 2, 3\}$).

Definition 7.40. A symmetric signature f of arity n is in $\mathscr{A}_2$ if there exists an $H \in \mathbf{O}_2(\mathbb{C})$ and a nonzero constant c such that

$$f = cH^{\otimes n}\left(\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}^{\otimes n} + \begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}^{\otimes n}\right).$$

Similarly, when such an H exists, we say that $f \in \mathscr{A}_2$ with transformation H. If $f \in \mathscr{A}_2$ with I_2, then we say f is in the canonical form of $\mathscr{A}_2$. If f is in the canonical form of $\mathscr{A}_2$, then by Proposition 7.36, for any $0 \le k \le n-2$, we have $f_{k+2} = -f_k$. Since f is nondegenerate, $f_1 \neq \pm \mathrm{i} f_0$ is implied.

It is worth noting that $H\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}$ is a nonzero multiple of either $\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}$ or $\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}$, for any orthogonal $H \in \mathbf{O}_2(\mathbb{C})$. The same is true for $H\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}$. Using this fact, the following lemma gives a characterization of $\mathscr{A}_2$. It says that any signature in $\mathscr{A}_2$ is essentially in canonical form.

Lemma 7.41. *Let f be a nondegenerate symmetric signature. Then $f \in \mathscr{A}_2$ iff f is of the form*

$$c\left(\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}^{\otimes n} + \beta\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}^{\otimes n}\right)$$

for some $c, \beta \neq 0$.

Proof. Assume that

$$f = c\left(\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}^{\otimes n} + \beta\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}^{\otimes n}\right)$$

for some $c, \beta \neq 0$. Consider the orthogonal transformation $H = \begin{bmatrix}a & b\\ b & -a\end{bmatrix}$, where $a = \frac{1}{2}\left(\beta^{\frac{1}{2n}} + \beta^{-\frac{1}{2n}}\right)$ and $b = \frac{1}{2\mathrm{i}}\left(\beta^{\frac{1}{2n}} - \beta^{-\frac{1}{2n}}\right)$. We pick a and b in this

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

way so that $a + b\mathrm{i} = \beta^{\frac{1}{2n}}$, $a - b\mathrm{i} = \beta^{-\frac{1}{2n}}$, and $(a + b\mathrm{i})(a - b\mathrm{i}) = a^2 + b^2 = 1$. Also $\left(\frac{a+b\mathrm{i}}{a-b\mathrm{i}}\right)^n = \beta$. Then

$$\begin{aligned} H^{\otimes n} f &= c\left(\begin{bmatrix} a+b\mathrm{i} \\ -a\mathrm{i}+b \end{bmatrix}^{\otimes n} + \beta \begin{bmatrix} a-b\mathrm{i} \\ a\mathrm{i}+b \end{bmatrix}^{\otimes n}\right) \\ &= c\left((a+b\mathrm{i})^n \begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}^{\otimes n} + (a-b\mathrm{i})^n \beta \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}^{\otimes n}\right) \\ &= c\sqrt{\beta}\left(\begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}^{\otimes n} + \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}^{\otimes n}\right), \end{aligned}$$

so f can be written as

$$f = c\sqrt{\beta}(H^{-1})^{\otimes n}\left(\begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}^{\otimes n} + \begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}^{\otimes n}\right).$$

Therefore $f \in \mathscr{A}_2$.

On the other hand, the desired form

$$f = c\left(\begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}^{\otimes n} + \beta \begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}^{\otimes n}\right)$$

follows from the fact that

$$\left\{\begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}, \begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}\right\}$$

is fixed setwise under any orthogonal transformation up to nonzero constants. □

Definition 7.42. A symmetric signature f of arity n is in $\mathscr{A}_3$ if there exists an $H \in \mathbf{O}_2(\mathbb{C})$ and a nonzero constant c such that

$$f = cH^{\otimes n}\left(\begin{bmatrix} 1 \\ \alpha \end{bmatrix}^{\otimes n} + \mathrm{i}^r \begin{bmatrix} 1 \\ -\alpha \end{bmatrix}^{\otimes n}\right)$$

for some $r \in \{0, 1, 2, 3\}$.

Again, when such an H exists, we say that $f \in \mathscr{A}_3$ with transformation H. If $f \in \mathscr{A}_3$ with I_2, then we say f is in the canonical form of $\mathscr{A}_3$. If f is in the canonical form of $\mathscr{A}_3$, then by Proposition 7.36, for any $0 \le k \le n-2$, we have $f_{k+2} = \mathrm{i} f_k$ and one of the following holds:

- $f_0 = 0$, or
- $f_1 = 0$, or
- $f_1 = \pm\alpha\mathrm{i} f_0 \neq 0$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Now we characterize the $\mathscr{A}$-transformable signatures.

Lemma 7.43. *Let f be a nondegenerate symmetric signature. Then f is $\mathscr{A}$-transformable iff $f \in \mathscr{A}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$.*

Proof. Assume that f is $\mathscr{A}$-transformable of arity n. By applying Lemma 7.38 to $\{f\}$, there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $f \in H\mathscr{A}$ or $f \in H \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \mathscr{A}$. This is equivalent to $(H^{-1})^{\otimes n} f \in \mathscr{A}$ or $(H^{-1})^{\otimes n} f \in \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \mathscr{A}$. Since f is nondegenerate and symmetric, we can replace $\mathscr{A}$ in the previous expressions with $\mathscr{F}_{123}$. Now we consider the possible cases.

1. If $(H^{-1})^{\otimes n} f \in \mathscr{F}_1$, then a further transformation by $H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbf{O}_2(\mathbb{C})$ puts it into the canonical form of $\mathscr{A}_1$, i.e., $f \in \mathscr{A}_1$ with transformation HH_2.
2. If $(H^{-1})^{\otimes n} f \in \mathscr{F}_2$, then it is already in the canonical form of $\mathscr{A}_1$.
3. If $(H^{-1})^{\otimes n} f \in \mathscr{F}_3$, then it is already of the equivalent form of $\mathscr{A}_2$ given by Lemma 7.41.
4. If $(H^{-1})^{\otimes n} f \in \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \mathscr{F}_1$, then a further transformation by $H_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbf{O}_2(\mathbb{C})$ puts it into the canonical form of $\mathscr{A}_1$, i.e., $f \in \mathscr{A}_1$ with transformation HH_2.
5. If $(H^{-1})^{\otimes n} f \in \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \mathscr{F}_2$, then it is already in the canonical form of $\mathscr{A}_3$.
6. If $(H^{-1})^{\otimes n} f \in \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix} \mathscr{F}_3$, then it is of the form $\begin{bmatrix} 1 \\ \alpha^3 \end{bmatrix}^{\otimes n} + \mathrm{i}^r \begin{bmatrix} 1 \\ -\alpha^3 \end{bmatrix}^{\otimes n}$ and a further transformation by $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \in \mathbf{O}_2(\mathbb{C})$ puts it into the canonical form of $\mathscr{A}_3$. To see this,

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^{\otimes n} \left(\begin{bmatrix} 1 \\ \alpha^3 \end{bmatrix}^{\otimes n} + \mathrm{i}^r \begin{bmatrix} 1 \\ -\alpha^3 \end{bmatrix}^{\otimes n} \right)$$
$$= \begin{bmatrix} -\alpha^3 \\ 1 \end{bmatrix}^{\otimes n} + \mathrm{i}^r \begin{bmatrix} \alpha^3 \\ 1 \end{bmatrix}^{\otimes n} = \left(-\alpha^3\right)^n \left(\begin{bmatrix} 1 \\ -\frac{1}{\alpha^3} \end{bmatrix}^{\otimes n} + (-1)^n \mathrm{i}^r \begin{bmatrix} 1 \\ \frac{1}{\alpha^3} \end{bmatrix}^{\otimes n} \right)$$
$$= \left(-\alpha^3\right)^n \left(\begin{bmatrix} 1 \\ \alpha \end{bmatrix}^{\otimes n} + \mathrm{i}^{2n+r} \begin{bmatrix} 1 \\ -\alpha \end{bmatrix}^{\otimes n} \right).$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

So $f \in \mathscr{A}_3$ with transformation $H\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. We have shown that if f is $\mathscr{A}$-transformable then $f \in \mathscr{A}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$.

Conversely, if there exists a matrix $H \in \mathbf{O}_2(\mathbb{C})$ such that $(H^{-1})^{\otimes n} f$ is in one of the canonical forms of $\mathscr{A}_1$, $\mathscr{A}_2$, or $\mathscr{A}_3$, then f is $\mathscr{A}$-transformable by applying Lemma 7.38. □

We also have a similar characterization for $\mathscr{P}$-transformable signatures. We define the stabilizer group of $\mathscr{P}$ similar to $\mathrm{Stab}(\mathscr{A})$. It is easy to see the left and right stabilizers of $\mathscr{P}$ coincide, which we denote by $\mathrm{Stab}(\mathscr{P})$. Furthermore, $\mathrm{Stab}(\mathscr{P})$ is generated by nonzero scalar multiples of matrices of the form $\begin{bmatrix} 1 & 0 \\ 0 & \nu \end{bmatrix}$ for any nonzero $\nu \in \mathbb{C}$ and $R = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Lemma 7.44. *Let $\mathcal{F}$ be a set of signatures. Then $\mathcal{F}$ is $\mathscr{P}$-transformable iff there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $\mathcal{F} \subseteq H\mathscr{P}$ or $\mathcal{F} \subseteq H\begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}\mathscr{P}$.*

Proof. Sufficiency is easily verified by checking that $=_2$ is transformed into $\mathscr{P}$ in both cases. In particular, H leaves $=_2$ unchanged.

If $\mathcal{F}$ is $\mathscr{P}$-transformable, then by definition, there exists a matrix T such that $(=_2)T^{\otimes 2} \in \mathscr{P}$ and $T^{-1}\mathcal{F} \subseteq \mathscr{P}$. The nondegenerate binary signatures in $\mathscr{P}$ are either $[0, 1, 0]$ or of the form $[1, 0, \nu]$, up to a scalar. However, notice that $[1, 0, 1] = [1, 0, \nu]\begin{bmatrix} 1 & 0 \\ 0 & \nu^{-\frac{1}{2}} \end{bmatrix}^{\otimes 2}$ and $\begin{bmatrix} 1 & 0 \\ 0 & \nu^{-\frac{1}{2}} \end{bmatrix} \in \mathrm{Stab}(\mathscr{P})$. Thus, we only need to consider $[1, 0, 1]$ and $[0, 1, 0]$. Now we apply Proposition 7.35.

1. If $(=_2)T^{\otimes 2} = [1, 0, 1]$, then by case 1 of Proposition 7.35, we have $T \in \mathbf{O}_2(\mathbb{C})$. Therefore $\mathcal{F} \subseteq H\mathscr{P}$ where $H = T \in \mathbf{O}_2(\mathbb{C})$.
2. If $(=_2)T^{\otimes 2} = [0, 1, 0]$, then by case 3 of Proposition 7.35, there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $T = \frac{1}{\sqrt{2}} H\begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$. Therefore $\mathcal{F} \subseteq H\begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}\mathscr{P}$ where $H \in \mathbf{O}_2(\mathbb{C})$.

□

We also have similar definitions of the sets $\mathscr{P}_1$ and $\mathscr{P}_2$.

Definition 7.45. A symmetric signature f of arity n is in $\mathscr{P}_1$ if there exists $H \in \mathbf{O}_2(\mathbb{C})$ and a nonzero constant c such that $f = cH^{\otimes n}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}^{\otimes n} + \beta \begin{bmatrix} 1 \\ -1 \end{bmatrix}^{\otimes n}\right)$, where $\beta \neq 0$.

We also define $\mathscr{P}_2 = \mathscr{A}_2$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

When such an H exists, we say that $f \in \mathscr{P}_1$ with transformation H. If $f \in \mathscr{P}_1$ with I_2, then we say f is in the canonical form of $\mathscr{P}_1$. If f is in the canonical form of $\mathscr{P}_1$, then by Proposition 7.36, for any $0 \le k \le n-2$, we have $f_{k+2} = f_k$. Since f is nondegenerate, $f_1 \neq \pm f_0$ is implied.

It is easy to check that $\mathscr{A}_1 \subset \mathscr{P}_1$. Now we characterize the $\mathscr{P}$-transformable signatures as we did for the $\mathscr{A}$-transformable signatures in Lemma 7.43.

Lemma 7.46. *Let f be a nondegenerate symmetric signature. Then f is $\mathscr{P}$-transformable iff $f \in \mathscr{P}_1 \cup \mathscr{P}_2$.*

Proof. Assume that f is $\mathscr{P}$-transformable of arity n. By applying Lemma 7.44 to $\{f\}$, there exists an $H \in \mathbf{O}_2(\mathbb{C})$ such that $f \in H\mathscr{P}$ or $f \in H\begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}\mathscr{P}$. This is equivalent to $(H^{-1})^{\otimes n} f \in \mathscr{P}$ or $(H^{-1})^{\otimes n} f \in \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}\mathscr{P}$.

The nondegenerate symmetric signatures in $\mathscr{P}$ take the form $[0, 1, 0]$, or $[a, 0, \ldots, 0, b] = a[1, 0]^{\otimes n} + b[0, 1]^{\otimes n}$, where $ab \neq 0$. Now we consider the possible cases.

1. If $(H^{-1})^{\otimes n} f = [0, 1, 0]^{\mathrm{T}} = \frac{1}{2\mathrm{i}}\left(\begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}^{\otimes 2} - \begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}^{\otimes 2}\right)$, then it is already of the equivalent form of $\mathscr{P}_2 = \mathscr{A}_2$ given by Lemma 7.41.
2. If $(H^{-1})^{\otimes n} f = a\begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes n} + b\begin{bmatrix} 0 \\ 1 \end{bmatrix}^{\otimes n}$, then a further transformation by $H_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \in \mathbf{O}_2(\mathbb{C})$ puts it into the canonical form of $\mathscr{P}_1$, i.e., $f \in \mathscr{P}_1$ with transformation HH_2.
3. If $(H^{-1})^{\otimes n} f = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}^{\otimes 2} [0, 1, 0]^{\mathrm{T}} = 2[1, 0, 1]^{\mathrm{T}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}^{\otimes 2} + \begin{bmatrix} 1 \\ -1 \end{bmatrix}^{\otimes 2}$, then it is already in the canonical form of $\mathscr{P}_1$.
4. If $(H^{-1})^{\otimes n} f = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}^{\otimes n}\left(a\begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes n} + b\begin{bmatrix} 0 \\ 1 \end{bmatrix}^{\otimes n}\right)$, then it is already of the equivalent form of $\mathscr{P}_2 = \mathscr{A}_2$ given by Lemma 7.41.

Conversely, if there exists a matrix $H \in \mathbf{O}_2(\mathbb{C})$ such that $(H^{-1})^{\otimes n} f$ is in one of the canonical forms of $\mathscr{P}_1$ or $\mathscr{P}_2$, then one can directly check that f is $\mathscr{P}$-transformable by applying Lemma 7.44. □

Combining Lemmas 7.43 and 7.46, we have a necessary and sufficient condition for a single nondegenerate symmetric signature to be $\mathscr{A}$- or $\mathscr{P}$-transformable.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Corollary 7.47. *Let f be a nondegenerate symmetric signature. Then f is $\mathscr{A}$- or $\mathscr{P}$-transformable iff $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$.*

Notice that our definitions of $\mathscr{P}_1$, $\mathscr{P}_2$, and $\mathscr{A}_3$ each involve an orthogonal transformation. For any single signature $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$, Holant($f$) is tractable. However, this does not imply that Holant($\mathscr{P}_1$), Holant($\mathscr{P}_2$), or Holant($\mathscr{A}_3$) is tractable. One can check, using Theorem 7.19, that Holant($\mathscr{P}_2$) is tractable while Holant($\mathscr{P}_1$) and Holant($\mathscr{A}_3$) are #P-hard.

Given a symmetric signature f, whether f is $\mathscr{A}$- or $\mathscr{P}$-transformable is decidable in polynomial time in the size of the symmetric signature f given as $n + 1$ numbers. We will discuss some necessary conditions in Section 7.9.

7.7.2 Dichotomies When $\mathscr{A}$- or $\mathscr{P}$-transformable Signatures Appear

Our characterizations of $\mathscr{A}$-transformable signatures in Lemma 7.43 and $\mathscr{P}$-transformable signatures in Lemma 7.46 are up to transformations in $\mathbf{O}_2(\mathbb{C})$. Since an orthogonal transformation never changes the complexity of the problem, in the proofs of the following lemmas, we assume any signature in $\mathscr{A}_i$ for $i = 1, 2, 3$, or $\mathscr{P}_j$ for $j = 1, 2$, is already in the canonical form. These lemmas state the general principle that if a set $\mathcal{F}$ of symmetric signatures contains a nondegenerate $\mathscr{A}$-transformable or a $\mathscr{P}$-transformable signature of arity $n \geq 3$, then the only way Holant($\mathcal{F}$) is tractable is when the entire set is $\mathscr{P}$-transformable or $\mathscr{A}$-transformable.

We first give a simple lemma to interpolate $(=_4)$.

Lemma 7.48. *Let $a, b \in \mathbb{C}$. If $ab \neq 0$, then for any set $\mathcal{F}$ of complex-weighted signatures containing $[a, 0, \ldots, 0, b]$ of arity $m \geq 3$,*

$$\text{Holant}(\mathcal{F} \cup \mathcal{EQ}_2) \leq_{\text{T}} \text{Holant}(\mathcal{F}),$$

where $\mathcal{EQ}_2 = \{=_2, =_4, \ldots, =_{2k}, \ldots\}$ is the set of all EQUALITIES *with an even arity.*

The same reduction also works for planar graphs, thus Pl-Holant($\mathcal{F} \cup \mathcal{EQ}_2$) $\leq_{\text{T}}$ Pl-Holant($\mathcal{F}$).

Proof. We only need to show how to get $(=_4)$. Note that a self-loop on $(=_4)$ gives $(=_2)$, and a chain of m copies of $(=_4)$ gives $(=_{2m+2})$, where we link one edge of the ith copy with one edge of the $(i + 1)$st copy $(1 \leq i \leq m - 1)$.

Since $a \neq 0$, we can normalize the first entry to get $[1, 0, \ldots, 0, x]$, where $x \neq 0$. First, we show how to obtain an arity 4 GEN-EQ signature. If $m = 3$, then we connect two copies together by a single edge to get an arity 4 signature. For

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

larger arities, we form self-loops until realizing a signature of arity 3 or 4. By this process, we have a signature $g = [1, 0, 0, 0, y]$, where $y \neq 0$. If y is a pth root of unity, then we can directly realize $(=_4)$ by connecting p copies of g together, two edges at a time as in Figure 7.2. Otherwise, y is not a root of unity and we can interpolate $(=_4)$ as follows.

Consider an instance Ω of Holant$(\mathcal{F} \cup \{=_4\})$. Suppose that $(=_4)$ appears n times in Ω. We stratify the assignments in Ω based on the assignments to $(=_4)$. We only need to consider the all-zero and all-one assignments since any other assignment contributes a factor of 0. We have

$$\text{Holant}_\Omega = \sum_{i=0}^{n} c_i,$$

where c_i is the sum, over all assignments with exactly i many $(=_4)$'s in Ω assigned 1111, and $n - i$ many $(=_4)$'s in Ω assigned 0000, of the product of evaluations of all other signatures on Ω.

We construct from Ω a sequence of instances Ω_s of Holant$(\mathcal{F})$ indexed by $s \geq 1$. We obtain Ω_s from Ω by replacing each occurrence of $(=_4)$ with a gadget g_s created from s copies of $[1, 0, 0, 0, y]$, connecting two edges together at a time as in Figure 7.2. Then

$$\text{Holant}_{\Omega_s} = \sum_{i=0}^{n} (y^s)^i c_i.$$

For $s \geq 1$, this gives a coefficient matrix that is Vandermonde. Since y is neither 0 nor a root of unity, y^s is distinct for each s. Therefore, the Vandermonde system $(1 \leq s \leq n + 1)$ has full rank. We can solve for the unknowns c_i in polynomial time and obtain the value of Holant$_\Omega$. □

Lemma 7.49. *Let $\mathcal{F}$ be a set of symmetric signatures. Suppose $\mathcal{F}$ contains a nondegenerate signature $f \in \mathscr{P}_1$ of arity $n \geq 3$. Then* Holant$(\mathcal{F})$ *is* #P*-hard unless $\mathcal{F}$ is $\mathscr{P}$-transformable or $\mathscr{A}$-transformable.*

Proof. Without loss of generality we assume f is in canonical form of $\mathscr{P}_1$ after an orthogonal transformation. The conclusion that $\mathcal{F}$ is $\mathscr{P}$-transformable or $\mathscr{A}$-transformable is invariant under this transformation. So by assumption, for any $0 \leq k \leq n - 2$, $f_{k+2} = f_k$ and $f_1 \neq \pm f_0$ since f is not degenerate. We can express f (after an orthogonal transformation) as

$$f = a_0 \begin{bmatrix} 1 \\ 1 \end{bmatrix}^{\otimes n} + a_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix}^{\otimes n},$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

where $a_0 = (f_0 + f_1)/2$, $a_1 = (f_0 - f_1)/2$, and $a_0 a_1 \neq 0$. For this f, we can further perform an orthogonal transformation by $H_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ so that f is transformed into the Gen-Eq signature $[a_0, 0, \ldots, 0, a_1]$ of arity n, up to a nonzero constant. By Lemma 7.48, we can obtain $(=_4)$. With $(=_4)$, we can realize any equality signature of even arity. Thus, $\#\mathrm{CSP}^2(H_2\mathcal{F}) \leq_{\mathrm{T}} \mathrm{Holant}(\mathcal{F})$.

Now we apply Theorem 6.16, the $\#\mathrm{CSP}^d$ dichotomy for $d = 2$, to the set $H_2\mathcal{F}$. If $\#\mathrm{CSP}^2(H_2\mathcal{F})$ is #P-hard, then $\mathrm{Holant}(\mathcal{F})$ is #P-hard as well. Otherwise, this problem is tractable. Therefore, there exists some T of the form $\begin{bmatrix} 1 & 0 \\ 0 & \alpha^k \end{bmatrix}$, where the integer $k \in \{0, 1, \ldots, 7\}$, such that $TH_2\mathcal{F}$ is a subset of $\mathscr{A}$ or $\mathscr{P}$.

If $TH_2\mathcal{F} \subseteq \mathscr{P}$, then we have $H_2\mathcal{F} \subseteq T^{-1}\mathscr{P}$. Notice that $T \in \mathrm{Stab}(\mathscr{P})$, so $T^{-1}\mathscr{P} = \mathscr{P}$. Thus, $\mathcal{F}$ is $\mathscr{P}$-transformable under this H_2 transformation. Otherwise, $TH_2\mathcal{F} \subseteq \mathscr{A}$. It is easy to verify that $(=_2)((TH_2)^{-1})^{\otimes 2}$ is $[1, 0, \mathrm{i}^{-k}] \in \mathscr{A}$. Thus, $\mathcal{F}$ is $\mathscr{A}$-transformable under this TH_2 transformation. □

Note that the hypothesis $f \in \mathscr{P}_1$ in Lemma 7.49 may seem to demand that all $\mathcal{F}$ is $\mathscr{P}$-transformable in order that the problem $\mathrm{Holant}(\mathcal{F})$ remain not #P-hard. However, this is not so. Since $\mathscr{A}_1$ is a proper subset of $\mathscr{P}_1$, the condition $f \in \mathscr{P}_1$ is consistent with possibly $f \in \mathscr{A}_1$. In that case $\mathcal{F}$ might be $\mathscr{A}$-transformable and not $\mathscr{P}$-transformable, and then $\mathrm{Holant}(\mathcal{F})$ is tractable (thus not #P-hard if #P does not collapse to P).

Lemma 7.50. *Let $\mathcal{F}$ be a set of symmetric signatures. Suppose $\mathcal{F}$ contains a nondegenerate signature $f \in \mathscr{P}_2$ of arity $n \geq 3$. Then* $\mathrm{Holant}(\mathcal{F})$ *is* #P*-hard unless $\mathcal{F}$ is $\mathscr{P}$-transformable or $\mathscr{A}$-transformable.*

Proof. By assumption, for any $0 \leq k \leq n-2$, $f_{k+2} = -f_k$ and $f_1 \neq \pm \mathrm{i} f_0$ since f is not degenerate. By Lemma 7.41, we have

$$f = a_0 \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}^{\otimes n} + a_1 \begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}^{\otimes n},$$

where $a_0 = (f_0 - \mathrm{i} f_1)/2$ and $a_1 = (f_0 + \mathrm{i} f_1)/2$, and $a_0 a_1 \neq 0$. Then under the holographic transformation

$$Z' = \begin{bmatrix} \sqrt[n]{a_0} & \sqrt[n]{a_1} \\ \mathrm{i}\sqrt[n]{a_0} & -\mathrm{i}\sqrt[n]{a_1} \end{bmatrix}^{-1},$$

we have

$$Z'^{\otimes n} f = (=_n) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes n} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}^{\otimes n}$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

and

$$\begin{aligned}\text{Holant}(=_2 \mid \mathcal{F} \cup \{f\}) &\equiv_{\text{T}} \text{Holant}([1,0,1](Z'^{-1})^{\otimes 2} \mid Z'\mathcal{F} \cup \{Z'^{\otimes n} f\}) \\ &\equiv_{\text{T}} \text{Holant}([0,1,0] \mid Z'\mathcal{F} \cup \{=_n\}).\end{aligned}$$

Thus, we have a bipartite graph having $(=_n)$ together with $Z'\mathcal{F}$ on the right and $(\neq_2) = [0,1,0]$ on the left up to a nonzero scalar. Inductively we can realize $(=_{kn})$ on the right for every integer $k > 0$. Suppose we have already constructed $(=_{(k-1)n})$. Place one copy of $(=_{(k-1)n})$ and three copies of $(=_n)$ on the right, and call them A, B, and C. Connect $(n-2)$ edges of A to $(=_{(k-1)n})$, one edge of A to one edge of B, and one edge of A to one edge of C, with all connections via $(\neq_2)$ on the left. This creates a $(=_{kn})$ on the right. Moreover, we can move these EQUALITY signatures to the left side by connecting each edge of $(=_{kn})$ by a copy of $(\neq_2)$. Thus, $\#\text{CSP}^n(Z'\mathcal{F}) \leq_{\text{T}} \text{Holant}(\mathcal{F})$.

Now we apply Theorem 6.16, the $\#\text{CSP}^d$ dichotomy with $d = n$, to the set $Z'\mathcal{F}$. Note that this n is a constant independent of the input size of any signature grid Ω. If $\#\text{CSP}^n(Z'\mathcal{F})$ is #P-hard, then $\text{Holant}(\mathcal{F})$ is #P-hard as well. Otherwise, $\#\text{CSP}^n(Z'\mathcal{F})$ is tractable. Let ω be a primitive $4n$th root of unity. Then under the holographic transformation $T = \begin{bmatrix} 1 & 0 \\ 0 & \omega^k \end{bmatrix}$ for some integer k, $TZ'\mathcal{F}$ is a subset of $\mathscr{A}$ or $\mathscr{P}$. If $TZ'\mathcal{F} \subseteq \mathscr{P}$, then we have $Z'\mathcal{F} \subseteq T^{-1}\mathscr{P}$. Notice that $T \in \text{Stab}(\mathscr{P})$, so $T^{-1}\mathscr{P} = \mathscr{P}$. Also recall that $(=_2)(Z'^{-1})^{\otimes 2}$ is a multiple of $(\neq_2)$. Thus, $\mathcal{F}$ is $\mathscr{P}$-transformable under this Z' transformation.

Otherwise, $TZ'\mathcal{F} \subseteq \mathscr{A}$. It is easy to verify that $(=_2)((TZ')^{-1})^{\otimes 2}$ is $[0,1,0] \in \mathscr{A}$ up to a scalar. Thus, $\mathcal{F}$ is $\mathscr{A}$-transformable under this TZ' transformation. □

Lemma 7.51. *Let $\mathcal{F}$ be a set of symmetric signatures. Suppose $\mathcal{F}$ contains a nondegenerate signature $f \in \mathscr{A}_3$ of arity $n \geq 3$. Then* $\text{Holant}(\mathcal{F})$ *is* #P*-hard unless $\mathcal{F}$ is $\mathscr{A}$-transformable.*

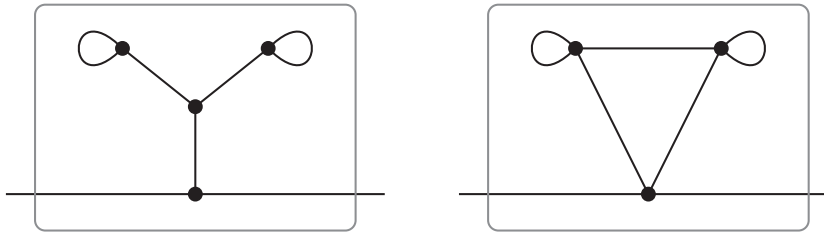
Proof. After an orthogonal transformation we can express f as

$$f = \lambda \left(\begin{bmatrix} 1 \\ \alpha \end{bmatrix}^{\otimes n} + \mathrm{i}^r \begin{bmatrix} 1 \\ -\alpha \end{bmatrix}^{\otimes n} \right),$$

for some $\lambda \neq 0$ and integer r. In this canonical form, we have $f_{k+2} = \mathrm{i} f_k$ for any $0 \leq k \leq n-2$.

A self-loop on f yields f', where $f'_k = f_k + f_{k+2} = (1+\mathrm{i}) f_k$. Thus up to the constant $(1+\mathrm{i})$, f' is just the first $n-2$ entries of f. By doing zero or more self-loops, we eventually obtain an arity 4 signature when n is even or a ternary one when n is odd. There are eight cases depending on the first two entries of f and the parity of n. However, in each case, we can realize the signature $[1, 0, \mathrm{i}]$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press



(a) Vertices assigned $[1, \alpha\mathrm{i}, \mathrm{i}, -\alpha]$. (b) Vertices assigned $[1, \alpha\mathrm{i}, \mathrm{i}, -\alpha, -1]$.

Figure 7.12. Constructions to realize [1, 0, i].

We list them here. (In the calculations below, we omit certain nonzero constant factors without explanation.)

- [0, 1, 0, i]: Another self-loop gives [0, 1]. Connect it back to the ternary, we get [1, 0, i].
- [1, 0, i, 0]: Another self-loop gives [1, 0]. Connect it back to the ternary, we get [1, 0, i].
- $[1, \alpha\mathrm{i}, \mathrm{i}, -\alpha]$: Another self-loop gives $[1, \alpha\mathrm{i}]$. Connect two copies of it to the ternary, we get $[1, -\alpha]$. Then connect this back to the ternary, we finally get [1, 0, i]. See Figure 7.12a.
- $[1, -\alpha\mathrm{i}, \mathrm{i}, \alpha]$: Same construction as in the previous case.
- [0, 1, 0, i, 0]: Another self-loop gives [0, 1, 0]. Connect it back to the arity 4 signature, we get [1, 0, i].
- [1, 0, i, 0, −1]: Another self-loop gives [1, 0, i] directly.
- $[1, \alpha\mathrm{i}, \mathrm{i}, -\alpha, -1]$: Another self-loop gives $[1, \alpha\mathrm{i}, \mathrm{i}]$. Connect two copies of it together, we get $[1, -\alpha, -\mathrm{i}]$. Connect this back to the arity 4 signature, we get [1, 0, i]. See Figure 7.12b.
- $[1, -\alpha\mathrm{i}, \mathrm{i}, \alpha, -1]$: Same construction as in the previous case.

With [1, 0, i] in hand, we can connect three copies to get [1, 0, −i]. Now we construct a bipartite graph, with $\mathcal{F} \cup \{=_2\}$ on the right side and [1, 0, −i] on the left. This problem is reducible to Holant($\mathcal{F}$). We perform a holographic transformation by $M^{-1} = \frac{1}{2\alpha}\begin{bmatrix} 1 & 1 \\ \alpha & -\alpha \end{bmatrix}$, with $M = \begin{bmatrix} \alpha & 1 \\ \alpha & -1 \end{bmatrix}$, and get

$$\begin{aligned}
&\text{Holant}([1, 0, -\mathrm{i}] \mid \mathcal{F} \cup \{f, =_2\}) \\
&\quad \equiv_{\mathrm{T}} \text{Holant}([1, 0, -\mathrm{i}](M^{-1})^{\otimes 2} \mid M\mathcal{F} \cup \{M^{\otimes n} f, M^{\otimes 2}(=_2)\}) \\
&\quad \equiv_{\mathrm{T}} \text{Holant}([1, 0, 1] \mid M\mathcal{F} \cup \{[1, 0, \ldots, 0, \mathrm{i}^r], [1, \mathrm{i}, 1]\}) \\
&\quad \equiv_{\mathrm{T}} \text{Holant}\left(M\mathcal{F} \cup \{[1, 0, \ldots, 0, \mathrm{i}^r], [1, \mathrm{i}, 1]\}\right).
\end{aligned}$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Note that $M^{\otimes n} f$ is $[1, 0, \ldots, 0, \mathrm{i}^r]$ (after normalizing the first entry); similarly $M^{\otimes 2}(=_2)$ is $[1, \mathrm{i}, 1]$. On the left side, $[1, 0, -\mathrm{i}](M^{-1})^{\otimes 2}$ is a nonzero multiple of $[1, 0, 1] = (=_2)$. Therefore, we can construct all EQUALITY signatures of even arity as follows: Use four copies of $[1, 0, \ldots, 0, \mathrm{i}^r]$ of arity n linked as a chain by three edges we get $(=_{4n-6})$ with even arity $4n - 6 \geq 6$. Then we can get $(=_2)$ and $(=_4)$ by self-loops. Finally, connecting k copies of $(=_4)$ as a chain by $k - 1$ edges gives $(=_{4k-2(k-1)}) = (=_{2k+2})$ for all $k \geq 2$. Thus,

$$\#\mathrm{CSP}^2(M\mathcal{F} \cup \{[1, \mathrm{i}, 1]\}) \leq_{\mathrm{T}} \mathrm{Holant}(\mathcal{F}).$$

Now we apply Theorem 6.16, the $\#\mathrm{CSP}^d$ dichotomy for $d = 2$, to the set $M\mathcal{F} \cup \{[1, \mathrm{i}, 1]\}$. If $\#\mathrm{CSP}^2(M\mathcal{F} \cup \{[1, \mathrm{i}, 1]\})$ is #P-hard, then $\mathrm{Holant}(\mathcal{F})$ is #P-hard as well. Otherwise, $\#\mathrm{CSP}^2(M\mathcal{F} \cup \{[1, \mathrm{i}, 1]\})$ is tractable, and there exists some T of the form $\begin{bmatrix} 1 & 0 \\ 0 & \alpha^d \end{bmatrix}$, where the integer $d \in \{0, 1, \ldots, 7\}$, such that $TM\mathcal{F} \cup \{T^{\otimes 2}[1, \mathrm{i}, 1]\}$ is a subset of $\mathscr{A}$ or $\mathscr{P}$.

However, $T^{\otimes 2}[1, \mathrm{i}, 1]$ cannot be in $\mathscr{P}$ for a diagonal T. Thus $TM\mathcal{F} \cup \{T^{\otimes 2}[1, \mathrm{i}, 1]\} \subseteq \mathscr{A}$. Further notice that if $d \in \{1, 3, 5, 7\}$ in the expression of T, then $T^{\otimes 2}[1, \mathrm{i}, 1]$ is not in $\mathscr{A}$. Hence, T must be of the form $\begin{bmatrix} 1 & 0 \\ 0 & \mathrm{i}^d \end{bmatrix}$, where the integer $d \in \{0, 1, 2, 3\}$. For such T, $T^{\otimes 2}[1, \mathrm{i}, 1] \in \mathscr{A}$, and $T^{-1}\mathscr{A} = \mathscr{A}$ as $T \in \mathrm{Stab}(\mathscr{A})$. Thus, $TM\mathcal{F} \cup \{T^{\otimes 2}[1, \mathrm{i}, 1]\} \subseteq \mathscr{A}$ simply becomes $M\mathcal{F} \subseteq \mathscr{A}$. Moreover, $(=_2)(M^{-1})^{\otimes 2}$ is a nonzero multiple of $[1, -\mathrm{i}, 1] \in \mathscr{A}$. Therefore, $\mathcal{F}$ is $\mathscr{A}$-transformable under the transformation by $M = \begin{bmatrix} \alpha & 1 \\ \alpha & -1 \end{bmatrix}$. □

7.8 Proof of Theorem 7.19

In this section, we finish the proof of the main theorem of this chapter, Theorem 7.19. We begin with a dichotomy for a single signature, which we prove by induction on its arity.

Theorem 7.52. *If f is a nondegenerate symmetric signature of arity ≥ 3 with complex weights in Boolean variables, then* $\mathrm{Holant}(f)$ *is* #P*-hard unless $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$ or f is vanishing, in which case the problem is in* P.

Recall that $\mathscr{A}_1 \subset \mathscr{P}_1$ and $\mathscr{A}_2 = \mathscr{P}_2$. Thus $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$ iff f is $\mathscr{A}$-transformable or $\mathscr{P}$-transformable by Corollary 7.47.

Proof. Let the arity of f be n. The base cases of $n = 3$ and $n = 4$ are proved in Theorem 7.29. Now assume $n \geq 5$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Let f' be the signature of arity ≥ 3 obtained from f by forming a self-loop. We consider the cases separately whether f' is degenerate or not.

- Suppose $f' = [a, b]^{\otimes(n-2)}$ is degenerate. There are three cases to consider.

  1. If $a = b = 0$, then f' is the all-zero signature. For f, this means $f_{k+2} = -f_k$ for $0 \leq k \leq n-2$, so $f \in \mathscr{A}_2 = \mathscr{P}_2$ by Lemma 7.41, and therefore Holant(f) is tractable.
  2. If $a^2 + b^2 \neq 0$, then f' is nonzero and $[a, b]$ is not a constant multiple of either $[1, \mathrm{i}]$ or $[1, -\mathrm{i}]$. We may normalize so that $a^2 + b^2 = 1$. Then the orthogonal transformation $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$ transforms the column vector $[a, b]$ to $[1, 0]$. Let $\hat{f}$ be the transformed signature from f, and $\widehat{f'} = [1, 0]^{\otimes(n-2)}$ the transformed signature from f'.

     Since an orthogonal transformation keeps $(=_2)$ invariant, this transformation commutes with the operation of taking a self-loop, i.e., $\widehat{f'} = (\hat{f})'$. Here $(\hat{f})'$ is the function obtained from $\hat{f}$ by taking a self-loop. So $\hat{f}_0 + \hat{f}_2 = 1$ and for every integer $1 \leq k \leq n-2$, we have $\hat{f}_k = -\hat{f}_{k+2}$. With one or more self-loops, we eventually obtain either $[1, 0]$ when n is odd or $[1, 0, 0]$ when n is even. In either case, we connect an appropriate number of copies of this signature to $\hat{f}$ to get an arity 4 signature $\hat{g} = [\hat{f}_0, \hat{f}_1, \hat{f}_2, -\hat{f}_1, -\hat{f}_2]$. We show that Holant$(\hat{g})$ is #P-hard. To see this, we first compute $\det(\widetilde{M_{\hat{g}}}) = -2(\hat{f}_0 + \hat{f}_2)(\hat{f}_1^2 + \hat{f}_2^2) = -2(\hat{f}_1^2 + \hat{f}_2^2)$, since $\hat{f}_0 + \hat{f}_2 = 1$. Therefore if $\hat{f}_1^2 + \hat{f}_2^2 \neq 0$, Holant$(\hat{g})$ is #P-hard by Corollary 7.25. Otherwise $\hat{f}_1^2 + \hat{f}_2^2 = 0$, and we consider $\hat{f}_2 = \mathrm{i}\hat{f}_1$ since the other case $\hat{f}_2 = -\mathrm{i}\hat{f}_1$ is similar. Since f is nondegenerate, $\hat{f}$ is nondegenerate, which implies $\hat{f}_2 \neq 0$. We can express $\hat{g}$ as $[1, 0]^{\otimes 4} - \hat{f}_2[1, \mathrm{i}]^{\otimes 4}$. Under the holographic transformation by $T = \begin{bmatrix} 1 & (-\hat{f}_2)^{1/4} \\ 0 & \mathrm{i}(-\hat{f}_2)^{1/4} \end{bmatrix}$, we have

$$\begin{aligned} \text{Holant}(=_2 \mid \hat{g}) &\equiv_{\mathrm{T}} \text{Holant}([1, 0, 1]T^{\otimes 2} \mid (T^{-1})^{\otimes 4}\hat{g}) \\ &\equiv_{\mathrm{T}} \text{Holant}(\hat{h} \mid =_4), \end{aligned}$$

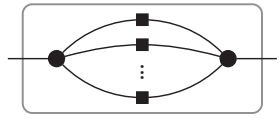
where

$$\hat{h} = [1, 0, 1]T^{\otimes 2} = [1, (-\hat{f}_2)^{1/4}, 0]$$

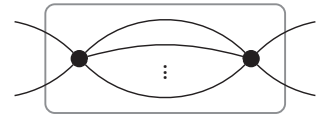
and $\hat{g}$ is transformed by T^{-1} into the arity 4 equality $(=_4)$, which can be verified by (also the reason for the choice of T)

$$T^{\otimes 4}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes 4} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}^{\otimes 4}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes 4} - \hat{f}_2 \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}^{\otimes 4} = \hat{g}.$$

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press



(a) The circles are assigned $\hat{f}$ and the squares are assigned $(\neq_2)$.



(b) The circles are assigned f.

Figure 7.13. Two gadgets used when $f' = [1, \pm\mathrm{i}]^{\otimes(n-2)}$.

By Theorem 5.3 for the case $k = 4$ (which is proved in Chapter 5), Holant$(\hat{h} \mid =_4)$ is #P-hard as $\hat{f}_2 \neq 0$.

3. If $a^2 + b^2 = 0$ but $(a, b) \neq (0, 0)$, then $[a, b]$ is a nonzero multiple of $[1, \pm\mathrm{i}]$. We have $f' = [1, \mathrm{i}]^{\otimes(n-2)}$ or $[1, -\mathrm{i}]^{\otimes(n-2)}$ up to a nonzero constant. We consider the first case since the other case is similar.

   Under the holographic transformation $Z = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$, $f' = Z^{\otimes(n-2)}[1, 0]^{\otimes(n-2)}$, up to a nonzero constant, and we have

$$\begin{aligned} \text{Holant}(=_2 \mid f) &\equiv_{\mathrm{T}} \text{Holant}(=_2 \mid f, f') \\ &\equiv_{\mathrm{T}} \text{Holant}([1, 0, 1]Z^{\otimes 2} \mid (Z^{-1})^{\otimes n} f, (Z^{-1})^{\otimes(n-2)} f') \\ &\equiv_{\mathrm{T}} \text{Holant}([0, 1, 0] \mid \hat{f}, [1, 0]^{\otimes(n-2)}), \end{aligned}$$

   where $\hat{f} = (Z^{-1})^{\otimes n} f$ takes the form $[\hat{f}_0, \hat{f}_1, 0, \ldots, 0, \hat{f}_n]$ with $\hat{f}_1 \neq 0$. We can normalize $\hat{f}_1 = 1$. If $\hat{f}_n = 0$ then $f \in \mathscr{V}^+$ and we are done. Suppose $\hat{f}_n \neq 0$.

   Now consider the gadget in Figure 7.13a with $\hat{f}$ assigned to both vertices, which clearly has a symmetric signature. The Hamming weight 0 entry of this binary signature is 0. This is because the arity of $\hat{f}$ is $n \geq 5$ and so either the left or the right side copy of $\hat{f}$ has at least two incident edges assigned 1 as they are linked by $(\neq_2)$ signatures. But $\hat{f}$ takes value 0 for input of Hamming weight w if $2 \leq w < n$. Similarly one can see that for input 01, the only nonzero configuration is all incident edges to the left side copy of $\hat{f}$ take 0, and those of the right take 1, which has the value $\hat{f}_0\hat{f}_n$. Finally, for input 11, we have two nonzero configurations and get $2\hat{f}_n$ as $\hat{f}_1 = 1$. Thus the binary signature of this gadget is $[0, \hat{f}_0\hat{f}_n, 2\hat{f}_n]$, which is equivalent to $[0, \hat{f}_0, 2]$ since $\hat{f}_n \neq 0$.

   By Lemma 7.8, $g = Z^{\otimes 2}[0, \hat{f}_0, 2]$ is a binary symmetric signature with $\mathrm{vd}^+(g) = 2$, and so $g \notin \mathscr{R}_2^+$.

   If $\hat{f}_0 \neq 0$, then $g = Z^{\otimes 2}[0, \hat{f}_0, 2]$ is nondegenerate. In this case we construct some function in $\mathscr{V}^+$. First we connect f' back to f, getting a binary

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

signature $Z^{\otimes 2}[0, 0, \hat{f}_n]$. This is best seen in the Z basis, where we connect all $n-2$ edges of $[1, 0]^{\otimes(n-2)}$ via $(\neq_2)$ to $\hat{f}$. So up to a nonzero factor we have $Z^{\otimes 2}[0, 0, 1]$. Then we connect both edges of $Z^{\otimes 2}[0, 0, 1]$ to f, the resulting signature is $h = Z^{\otimes(n-2)}[\hat{f}_0, 1, 0, \ldots, 0]$ of arity $n-2 \geq 3$, up to a nonzero scalar. This is again best seen in the Z basis. Notice that h is nondegenerate and $h \in \mathscr{V}^+$. By Lemma 7.33, Holant$(\{h, g\})$ is #P-hard, and hence Holant(f) is also #P-hard.

Otherwise suppose $\hat{f}_0 = 0$. Then we have $g = Z^{\otimes 2}[0, 0, 2] = 2Z^{\otimes 2}[0, 1]^{\otimes 2} = [1, -\mathrm{i}]^{\otimes 2}$. Connecting this degenerate signature to f, we get a signature $h = \langle f, g \rangle$. Under the Z basis, $(Z^{-1})^{\otimes(n-2)}h = [\hat{f}_0, \hat{f}_1, 0, \ldots, 0] = [0, 1, 0, \ldots, 0]$ of arity $n-2 \geq 3$, up to a nonzero scalar. Hence h is nondegenerate and $\mathrm{rd}^+(h) = 1$. Therefore $h \in \mathscr{V}^+$.

Moreover, assigning f to both vertices in the gadget of Figure 7.13b, we get a nondegenerate signature $h' \in \mathscr{V}^-$ of arity 4. To see this, consider this gadget after a holographic transformation by Z. In this bipartite setting, it is the same as assigning $\hat{f} = [0, 1, 0, \ldots, 0, \hat{f}_n]$ to both the circle and triangle vertices in the gadget of Figure 7.7a. The square vertices there are still assigned $(\neq_2) = [0, 1, 0]$. Although it is not apparent from the gadget's geometry, this signature is in fact symmetric. In particular, its values on inputs 1010 and 1100 are both zero. Note that as $n \geq 5$, the number of parallel paths with $(\neq_2)$ in the middle between the two copies of $\hat{f}$ is $n-2 \geq 3$, and thus at least one copy of $\hat{f}$ has two internal incident edges set to 1 due to the $(\neq_2)$ in the middle. For Hamming weight ≥ 2, the only nonzero value of $\hat{f}$ is at weight n. From this, one can see that the resulting signature is $\widehat{h'} = (Z^{-1})^{\otimes 4}h' = [0, 0, 0, \hat{f}_n, 0]$. Hence $\mathrm{rd}^-(h') = 1$ by Lemma 7.8. Therefore h' is nondegenerate and $h' \in \mathscr{V}^-$.

By Lemma 7.34, Holant$(\{h, h'\})$ is #P-hard, hence Holant(f) is also #P-hard.

- Suppose f' is nondegenerate. If Holant(f') is #P-hard then so is Holant(f). We now assume Holant(f') is not #P-hard. Then, by the inductive hypothesis, $f' \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$ or f' is vanishing. If $f' \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$, then applying Lemma 7.49, Lemma 7.50, or Lemma 7.51 to f' and the set $\{f, f'\}$, we either have that Holant$(\{f, f'\})$ is #P-hard, so Holant(f) is #P-hard as well, or that f is $\mathscr{A}$- or $\mathscr{P}$-transformable, so by Corollary 7.47, $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$.

  Otherwise, f' is vanishing, so $f' \in \mathscr{V}^\sigma$ for $\sigma \in \{+, -\}$ by Theorem 7.11. We may assume that $f' \in \mathscr{V}^+$; the other case is similar. Let $\mathrm{rd}^+(f') = d - 1$. Then $f' \in \mathscr{V}^+$ implies that $2d < n$ and f' is nondegenerate implies that $d \geq 2$.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Under the transformation $Z = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$, we have

$$\begin{aligned} \text{Holant}(=_2 \mid f) &\equiv_{\mathrm{T}} \text{Holant}((=_2)Z^{\otimes 2} \mid (Z^{-1})^{\otimes n} f) \\ &\equiv_{\mathrm{T}} \text{Holant}(\neq_2 \mid \hat{f}), \end{aligned}$$

where $\hat{f} = (Z^{-1})^{\otimes n} f$. We have $\hat{f} = [\hat{f}_0, \hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0, \hat{f}_n]$, with $\hat{f}_d \neq 0$, because $\mathrm{rd}^+(f') = d - 1$. If $\hat{f}_n = 0$ then f is vanishing, and we are done. We may assume $\hat{f}_n \neq 0$.

Note that $d < n/2$ and $\hat{f}_d$ occurs strictly in the first half, and there are $n - d - 1$ many 0's between $\hat{f}_d$ and $\hat{f}_n$. Taking a self-loop in the original setting is equivalent to connecting both edges of $(\neq_2)$ to two inputs of a signature after this transformation. Thus, doing this once on $\hat{f}$, we can get $\widehat{f'} = 2[\hat{f}_1, \ldots, \hat{f}_d, 0, \ldots, 0]$ corresponding to f', namely $(Z^{-1})^{\otimes(n-2)} f'$. Doing this $d - 2$ times on $\hat{f}$, up to a nonzero factor, we get a signature $\hat{h} = [\hat{f}_{d-2}, \hat{f}_{d-1}, \hat{f}_d, 0, \ldots, 0, 0/\hat{f}_n]$ of arity $n - 2(d-2) = n - 2d + 4$. The last entry is $\hat{f}_n$ when $d = 2$ and is 0 when $d > 2$.

As $n > 2d$, we may do this two more times on $\hat{h}$ and get $[\hat{f}_d, 0, \ldots, 0]$ of arity $k = n - 2d$. Now connect this signature back to $\hat{f}$ via $(\neq_2)$. We get a signature of arity $n - k = 2d$. This signature consists of the last $2d + 1$ signature entries of $\hat{f}$, up to a nonzero factor. We may repeat this operation zero or more times, which reduces the arity of $\hat{f}$ by k each time, until the arity k' of the resulting signature is less than or equal to k. We claim that this signature has the form $\hat{g} = [0, \ldots, 0, \hat{f}_n]$. In other words, the $k' + 1$ entries of $\hat{g}$ consist of the last $\hat{f}_n$ and k' many 0's in the signature $\hat{f}$, all appearing after $\hat{f}_d$. This is because there are $n - d - 1$ many 0 entries in the signature $\hat{f}$ after $\hat{f}_d$, and $n - d - 1 \geq k \geq k'$.

Hence, in the RHS of $\text{Holant}(\neq_2 \mid \hat{f})$ we can get both $[\hat{f}_d, 0, \ldots, 0]$ of arity k and $[0, \ldots, 0, \hat{f}_n]$ of arity $k' \leq k$. If $k > k'$ we may repeat this process in the way of a subtractive Euclidean algorithm. The result is that we can get both $[1, 0]^{\otimes t}$ and $[0, 1]^{\otimes t}$, where $t = \gcd(k, k')$. Now putting k/t many copies of $[0, 1]^{\otimes t}$ together, we get $[0, 1]^{\otimes k}$.

Then we connect $[0, 1]^{\otimes k}$ back to $\hat{h} = [\hat{f}_{d-2}, \hat{f}_{d-1}, \hat{f}_d, 0, \ldots, 0, 0/\hat{f}_n]$ of arity $n - 2d + 4 = k + 4$ via $(\neq_2)$. Thus we get a signature of arity 4, which is $[\hat{f}_{d-2}, \hat{f}_{d-1}, \hat{f}_d, 0, 0]$. Note that the last entry is 0 (and not $\hat{f}_n$), because $k \geq 1$.

However, $\text{Holant}(\neq_2 \mid [\hat{f}_{d-2}, \hat{f}_{d-1}, \hat{f}_d, 0, 0])$ is equivalent to $\text{Holant}(\neq_2 \mid [0, 0, 1, 0, 0])$ when $\hat{f}_d \neq 0$. This is the Eulerian orientation problem #EO on 4-regular graphs and is #P-hard by Theorem 6.32. □

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Finally we are ready to give the hardness proof in Theorem 7.19. We prove the contrapositive: If Holant($\mathcal{F}$) is not #P-hard then $\mathcal{F}$ must be in one of the five cases listed in Theorem 7.19. Combined with the tractability proof given in Section 7.2, this completes the proof of Theorem 7.19.

Proof of hardness in Theorem 7.19. Assume that Holant($\mathcal{F}$) is not #P-hard. If all of the nondegenerate signatures in $\mathcal{F}$ are of arity ≤ 2, then it is tractable case 1. Otherwise we have some nondegenerate signatures of arity ≥ 3. For any such f, by Theorem 7.52, $f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$ or f is vanishing. If any of them is in $\mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{A}_3$, then by Lemma 7.49, Lemma 7.50, or Lemma 7.51, we have that $\mathcal{F}$ is $\mathscr{A}$- or $\mathscr{P}$-transformable, which are tractable cases 2 and 3.

Now we assume all nondegenerate signatures of arity ≥ 3 in $\mathcal{F}$ are vanishing, and there is a nonempty set of such signatures in $\mathcal{F}$. By Lemma 7.34, they must all be in $\mathscr{V}^{\sigma}$ for the same $\sigma \in \{+, -\}$. By Lemma 7.33, we know that any nondegenerate binary signature in $\mathcal{F}$ must belong to $\mathscr{R}_2^{\sigma}$. Furthermore, if there is an $f \in \mathscr{V}^{\sigma}$ from $\mathcal{F}$ such that $\mathrm{rd}^{\sigma}(f) \geq 2$, then by Lemma 7.31, the only unary signature that is allowed in $\mathcal{F}$ is $[1, \sigma \mathrm{i}]$, and all degenerate signatures in $\mathcal{F}$ are a tensor power of $[1, \sigma \mathrm{i}]$, up to a scalar. Thus, all nondegenerate signatures of arity ≥ 3 as well as all degenerate signatures belong to $\mathscr{V}^{\sigma}$, and all nondegenerate binary signatures belong to $\mathscr{R}_2^{\sigma}$. This is tractable case 4.

Finally, we have the following: (i) all nondegenerate signatures of arity ≥ 3 in $\mathcal{F}$ belong to $\mathscr{V}^{\sigma}$; (ii) all signatures $f \in \mathcal{F} \cap \mathscr{V}^{\sigma}$ have $\mathrm{rd}^{\sigma}(f) \leq 1$, which implies that $f \in \mathscr{R}_2^{\sigma}$; and (iii) all nondegenerate binary signatures in $\mathcal{F}$ belong to $\mathscr{R}_2^{\sigma}$. Hence all nondegenerate signatures in $\mathcal{F}$ belong to $\mathscr{R}_2^{\sigma}$. All unary signatures also belong to $\mathscr{R}_2^{\sigma}$ by definition. This is indeed tractable case 5. The proof is complete. □

7.9 Decidability of the Dichotomy

Theorem 7.19 is a full dichotomy for Holant($\mathcal{F}$) for a set of symmetric signatures on Boolean variables. It says that if $\mathcal{F}$ satisfies one of the tractability conditions then Holant($\mathcal{F}$) is computable in P; otherwise it is #P-hard. A natural question is given a finite set $\mathcal{F}$, how can one decide which case it is. This is the decidability question of the dichotomy theorem.

Clearly conditions 1, 4, and 5 are all easily decidable in polynomial time. Hence we may assume $\mathcal{F}$ contains at least one nondegenerate signature of arity ≥ 3, and the real question is how can one decide whether $\mathcal{F}$ is $\mathscr{A}$- or $\mathscr{P}$-transformable. This problem is nontrivial because it involves a potentially infinite set of transformational matrices.

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

Theorem 7.53. *Given any finite set $\mathcal{F}$ of symmetric signatures with complex weights in Boolean variables containing a nondegenerate signature f of arity ≥ 3, it is decidable in polynomial time in the input size of $\mathcal{F}$ whether $\mathcal{F}$ is $\mathscr{A}$- or $\mathscr{P}$-transformable. Thus the tractability criterion of Theorem 7.19 is decidable in polynomial time.*

In this book we will not give the full proof of this theorem; interested readers can find it in [CGW14]. In this section we discuss some necessary conditions of the tractability criterion based on some structural properties, that both shed new light on the dichotomy and also help in many practical cases to test the criterion on particular functions. The decision algorithm of Theorem 7.53 is built on top of these ideas with some additional technical extensions. We will omit some technical proofs in this section.

By Corollary 7.47, f is $\mathscr{A}$- or $\mathscr{P}$-transformable iff $f \in \mathscr{A}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$, or $f \in \mathscr{P}_1 \cup \mathscr{P}_2$, where $\mathscr{A}_1 \subset \mathscr{P}_1$ and $\mathscr{A}_2 = \mathscr{P}_2$. We say a signature $f = [f_0, f_1, \ldots, f_n]$ satisfies a second order recurrence relation, if there exist $a, b, c \in \mathbb{C}$ not all zero, such that for all $0 \leq k \leq n-2$, $af_k - bf_{k+1} + cf_{k+2} = 0$. We say the recurrence has type $\langle a, b, c\rangle$. The characteristic equation of the recurrence relation is $a - bX + cX^2 = 0$. The discriminant is $b^2 - 4ac$.

Lemma 7.54. *Let f be a nondegenerate symmetric signature of arity $n \geq 3$. If f satisfies a second-order recurrence relation then the recurrence relation is unique up to a nonzero multiple.*

Under any holographic transformation, the property that a signature satisfies a second order recurrence relation, and if so, the condition $b^2 - 4ac \neq 0$ are both invariant.

Lemma 7.55. *Let f be a symmetric signature of arity n and $\hat{f} = T^{\otimes n} f$ for some $T \in \mathbf{GL}_2(\mathbb{C})$. Then f satisfies a second-order recurrence relation iff $\hat{f}$ does. Furthermore, in this case the discriminant for f is nonzero iff the one for $\hat{f}$ is nonzero.*

Proof. If we write the $2^2 \times 2^{n-2}$ signature matrix of f as M_f, with rows indexed by the first two input bits and the rest are the other $n-2$ input bits, as in Definition 1.4, then the corresponding signature matrix of $T^{\otimes n} f$ is $T^{\otimes 2} M_f (T^{\mathrm{T}})^{\otimes (n-2)}$. f satisfies a second-order recurrence relation of type $\langle a, b, c\rangle$ iff $(a, -b/2, -b/2, c)$ is a solution of $(x, y, y, z)M_f = 0$ iff $(a, -b/2, -b/2, c)(T^{-1})^{\otimes 2}$ is a solution of $(x', y', y', z')M_{T^{\otimes n} f} = 0$. In this case, the discriminant of f is $\det \begin{bmatrix} a & -b/2 \\ -b/2 & c \end{bmatrix}$, and that of $T^{\otimes n} f$ is multiplied by $(\det T)^2$. □

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

A simple but important observation is that if f is $\mathscr{A}$- or $\mathscr{P}$-transformable, then f satisfies a second-order recurrence relation with a nonzero discriminant. In fact, in their canonical forms, $f \in \mathscr{A}_1$, $\mathscr{A}_2$, and $\mathscr{A}_3$ have types $\langle -1, 0, 1\rangle$, $\langle 1, 0, 1\rangle$, and $\langle -\mathrm{i}, 0, 1\rangle$, with characteristic polynomials $X^2 - 1$, $X^2 + 1$, and $X^2 - \mathrm{i}$, respectively. For $f \in \mathscr{P}_1$ and $\mathscr{P}_2$, the types are $\langle -1, 0, 1\rangle$ and $\langle 1, 0, 1\rangle$. By Lemma 7.55, any $f \in \mathscr{P}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$ satisfies a second-order recurrence relation with a nonzero discriminant. Hence,

Lemma 7.56. *Let $f \in \mathscr{P}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$ be a nondegenerate symmetric signature of arity $n \geq 3$. Then f satisfies a second-order recurrence relation with nonzero discriminant. It has the form $f = \begin{bmatrix}\alpha\\ \gamma\end{bmatrix}^{\otimes n} + \begin{bmatrix}\beta\\ \delta\end{bmatrix}^{\otimes n}$, where* $\det \begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix} \neq 0$.

The following definition of the θ function is crucial.

Definition 7.57. For a pair of linearly independent vectors $u = \begin{bmatrix}\alpha\\ \gamma\end{bmatrix}$ and $v = \begin{bmatrix}\beta\\ \delta\end{bmatrix}$, we define

$$\theta(u, v) = \left(\frac{\langle u, v\rangle}{\det[u, v]}\right)^2 = \left(\frac{\alpha\beta + \gamma\delta}{\alpha\delta - \beta\gamma}\right)^2,$$

where $\langle u, v\rangle$ is the dot product, and $\det[u, v]$ is the determinant with column vectors u and v. Furthermore, if $f = u^{\otimes n} + v^{\otimes n}$ for linearly independent u and v and $n \geq 3$, then we define $\theta(f) = \theta(u, v)$.

By linear independence, $\det[u, v] \neq 0$, so $\theta(u, v)$ is well defined.

Lemma 7.58. *For two linearly independent vectors $u, v \in \mathbb{C}^2$, and nonzero λ and $\mu \in \mathbb{C}$, $\theta(\lambda u, \mu v) = \theta(u, v)$. If $H \in \mathbf{O}_2(\mathbb{C})$, let $\widehat{u} = Hu$ and $\widehat{v} = Hv$. Then $\theta(\widehat{u}, \widehat{v}) = \theta(u, v)$. In particular $\theta(v, u) = \theta(u, v)$.*

Proof. The first claim is obvious. Because H is orthogonal, the dot product is preserved, $\langle \widehat{u}, \widehat{v}\rangle = \langle u, v\rangle$. Also $\det[\widehat{u}, \widehat{v}] = \pm \det[u, v]$. Hence $\theta(\widehat{u}, \widehat{v}) = \theta(u, v)$. □

Intuitively, $\theta(u, v)$ is the square of the cotangent of the angle from u to v. (The cotangent function is extended to the complex domain.) Let $f = u^{\otimes n} + v^{\otimes n}$ be a nondegenerate signature of arity $n \geq 3$. Since f is nondegenerate, u and v are linearly independent. The next proposition implies that this expression

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

for f via u and v is unique up to a root of unity. Therefore, $\theta(f)$ from Definition 7.57 is well defined from f; it is independent of the particular expression of f.

Proposition 7.59. *Let $u, v, u', v' \in \mathbb{C}^2$. Suppose u and v are linearly independent. If for some $n \geq 3$, we have*

$$u^{\otimes n} + v^{\otimes n} = u'^{\otimes n} + v'^{\otimes n}, \tag{7.17}$$

then with possibly exchanging u and v, there exist ρ and $\omega \in \mathbb{C}$, such that $u' = \rho u$, $v' = \omega v$ where $\rho^n = \omega^n = 1$.

We omit the proof. Now we give some necessary conditions for membership in $\mathscr{A}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$. Recall that $\mathscr{A}_1 \subseteq \mathscr{P}_1$.

Lemma 7.60. *Let f be a nondegenerate symmetric signature of arity ≥ 3. Then*

1. $f \in \mathscr{P}_1 \implies \theta(f) = 0$,
2. $f \in \mathscr{A}_2 \implies \theta(f) = -1$, *and*
3. $f \in \mathscr{A}_3 \implies \theta(f) = -\frac{1}{2}$.

Proof. The result clearly holds when f is in the canonical form of each set. This extends to the rest of each set by Lemma 7.58. □

These results imply the following lemma.

Lemma 7.61. *Let f be a nondegenerate symmetric signature f of arity $n \geq 3$. If f is $\mathscr{A}$-transformable, then f satisfies a second-order recurrence relation with nonzero discriminant, thus has the form $f = u^{\otimes n} + v^{\otimes n}$, where u and v are linearly independent, and $\theta(f) \in \{0, -1, -\frac{1}{2}\}$.*

Note that the condition in Lemma 7.61 on $\theta(f)$ is only a necessary condition for f being $\mathscr{A}$-transformable. This condition is not sufficient to determine if $f \in \mathscr{A}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$. In fact, it can be directly verified that $\theta(u, v) = -1$ for linearly independent u and v if and only if at least one of u, v is $\begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}$ or $\begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}$, up to a nonzero scalar. Thus, if $f = u^{\otimes n} + v^{\otimes n}$ with $u = \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}$ and v is *not* a multiple of $\begin{bmatrix} 1 \\ \pm\mathrm{i} \end{bmatrix}$, then $\theta(f) = -1$. But f is not in $\mathscr{A}_2$ by Proposition 7.59, and Lemma 7.41. A full characterization can be achieved with additional work; but we will not pursue it further here [CGW14].

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press

To obtain the decision algorithm for Theorem 7.19 as stated in Theorem 7.53 will require more careful work concerning a common transformation for the whole set $\mathcal{F}$. We will not present all the details, but it starts with the following.

Lemma 7.38 showed that every $\mathscr{A}$-transformable set $\mathcal{F}$ is contained in either $H\mathscr{A}$ or $H\begin{bmatrix}1 & 0\\ 0 & \alpha\end{bmatrix}\mathscr{A}$ for some $H \in \mathbf{O}_2(\mathbb{C})$. For nondegenerate symmetric signatures, we can replace $\mathscr{A}$ by $\mathscr{F}_1 \cup \mathscr{F}_2 \cup \mathscr{F}_3$. We have the following classification.

$$\theta(f) = \begin{cases} 0 & \text{if } f \in \mathscr{F}_1 \cup \mathscr{F}_2 \cup \begin{bmatrix}1 & 0\\ 0 & \alpha\end{bmatrix}\mathscr{F}_1, \\ -1 & \text{if } f \in \mathscr{F}_3, \\ -\frac{1}{2} & \text{if } f \in \begin{bmatrix}1 & 0\\ 0 & \alpha\end{bmatrix}(\mathscr{F}_2 \cup \mathscr{F}_3). \end{cases} \tag{7.18}$$

Regarding $\mathscr{P}$-transformability, Lemma 7.60 also gives a useful necessary condition for f being $\mathscr{P}$-transformable. As $\mathscr{P}_2 = \mathscr{A}_2$, the same discussion for $\mathscr{A}_2$ applies to $\mathscr{P}_2$.

The next lemma tells us how to decide membership in $\mathscr{P}_1$ for signatures of arity ≥ 3.

Lemma 7.62. *Let $f = v_0^{\otimes n} + v_1^{\otimes n}$ be a symmetric signature of arity $n \geq 3$, where v_0 and v_1 are linearly independent. Then $f \in \mathscr{P}_1$ iff $\theta(f) = 0$.*

With some additional work one can obtain the decision algorithm stated in Theorem 7.53. Details can be found in [CGW14].

https://doi.org/10.1017/9781107477063.008 Published online by Cambridge University Press