9

Planar Holant for Symmetric Constraints

In this chapter, we consider the complexity of planar Holant problems Pl-Holant($\mathcal{F}$), where $\mathcal{F}$ is any set of complex-valued, symmetric functions on Boolean variables.

9.1 Introduction

In Chapter 8 we presented a complexity dichotomy, Theorem 8.34, for planar #CSP problems Pl-#CSP($\mathcal{F}$), where $\mathcal{F}$ is any set of complex-valued symmetric constraint functions over the Boolean domain. This theorem states that Pl-#CSP($\mathcal{F}$) is either tractable in polynomial time or #P-hard. It is tractable precisely when #CSP($\mathcal{F}$) is tractable without the planarity restriction, and this happens precisely when $\mathcal{F} \subseteq \mathscr{A}$ or $\mathcal{F} \subseteq \mathscr{P}$ (by Theorem 3.7); or when $\mathcal{F}$ is transformable to a set of standard matchgate signatures, i.e., $\mathcal{F} \subseteq \widehat{\mathscr{M}} = H_2\mathscr{M}$, where the Hadamard matrix $H_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. Thus, for those #CSP problems over the Boolean domain that are #P-hard in general but solvable in polynomial time over planar graphs, a *universal* methodology is a suitable holographic transformation followed by Kasteleyn's algorithm given in Chapter 4.

This has been a recurring theme. We have seen this phenomenon already several times in earlier chapters. For example, in Chapter 5, Theorems 5.1 and 5.3 are also examples of this. It can be verified that the complexity classification theorems for the Tutte polynomial $T(G; x, y)$, Theorems 6.28 and 6.29, are also consistent with this theme. Indeed, the only additional tractable cases in Theorem 6.29 over Theorem 6.28 on planar graphs are all captured by $(x-1)(y-1) = 2$. By the correspondence with the Potts model (see Chapter 1, Section 1.2.1), this is precisely the case $q = 2$ (the Boolean domain; thus the Potts model becomes the Ising model, with a symmetric binary signature $[1+\gamma, 1, 1+\gamma]$ on each edge). Note that $[1+\gamma, 1, 1+\gamma] \in \widehat{\mathscr{M}}$.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

In this chapter, we want to achieve a complexity classification of planar Holant problems Pl-Holant($\mathcal{F}$) for symmetric signatures $\mathcal{F}$. This would be a stronger theorem than Theorem 8.34, as Pl-#CSP($\mathcal{F}$) is just Pl-Holant($\mathcal{F} \cup \mathcal{EQ}$). All previously known classification theorems, such as those presented in earlier chapters, suggest the following putative classification of Pl-Holant($\mathcal{F}$) into *exactly* three categories, (1) those that are P-time solvable over general graphs; (2) those that are P-time solvable over planar graphs but #P-hard over general graphs; and (3) those that remain #P-hard over planar graphs. Moreover, category (2) consists of precisely those problems that are holographically reducible to Kasteleyn's algorithm. This theme is so strong that it has become an intuitive and trusty guide for us when we investigate unknown problems and plan proof strategies. In fact, many partial results leading to the main theorem in this chapter are proved in this way.

However, a surprising departure to this theme occurs with Pl-Holant($\mathcal{F}$). It is still true that all Pl-Holant($\mathcal{F}$) problems can be classified into exactly three categories; but there are more problems in category (2) than those that are holographically reducible to Kasteleyn's algorithm. We discover new planar tractable problems that are not expressible by a holographic reduction to matchgates. Thus a holographic reduction to Kasteleyn's algorithm is *not* a universal strategy for Holant problems that are #P-hard in general but solvable in polynomial time over planar graphs. So far, for any class of counting problems expressible in our frameworks for sum-of-product computations, this is the only known exception to the *universality* of holographic reductions to Kasteleyn's algorithm. Moreover, with the discovery of this new class of planar tractable problems the picture is complete for Pl-Holant($\mathcal{F}$): there are no more undiscovered planar tractable problems unless #P collapses to P.

This classification generalizes both the Holant dichotomy, Theorem 7.19, and the planar #CSP classification, Theorem 8.34. At a high level, we can state the main theorem of this chapter as follows.

Theorem 9.1. *Let $\mathcal{F}$ be a set of complex-valued, symmetric functions on Boolean variables. Then there is an effective classification for all possible $\mathcal{F}$, according to which,* Holant($\mathcal{F}$) *is either (1) P-time computable over general graphs, or (2) P-time computable over planar graphs but* #P*-hard over general graphs, or (3)* #P*-hard over planar graphs. Furthermore, category (2) consists precisely of those problems that are holographically reducible to Kasteleyn's algorithm, and those newly discovered problems.*

The complete statement is given in Theorem 9.29. The classification is explicit. The tractability criterion is decidable in polynomial time.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

Let us meet some new tractable problems. A hypergraph $H = (V, E)$ consists of a finite vertex set V and a set E of hyperedges e, each of which is a subset of V. If every e has cardinality 2, then H is an ordinary graph. A perfect matching in H is a subset $M \subseteq E$ such that every $v \in V$ is contained in exactly one $e \in M$. Consider the problem of counting perfect matchings in a hypergraph H. In general, each hyperedge e has a cardinality k_e, and if $k_e = k$ for all e, then the hypergraph is called a k-uniform hypergraph. The number of hyperedges e containing a vertex v is called its degree; if every vertex has the same degree d, then it is called d-regular. Each hypergraph H can be represented by its bipartite edge–vertex incidence graph B_H with vertex bipartition (E, V) and an edge exists in B_H between $v \in V$ and $e \in E$ exactly when $v \in e$. Clearly H is a k-uniform d-regular hypergraph iff B_H is a (k, d)-biregular bipartite graph. We say a hypergraph H is planar when B_H is planar [Zyk74]. Some new tractable problems are counting perfect matchings in planar hypergraphs.

Suppose H is a k-uniform planar hypergraph. If $k = 2$, this is an ordinary planar graph, and Kasteleyn's algorithm can count the number of perfect matchings in polynomial time. We will show that for $k = 3$ and $k = 4$, the problem of counting perfect matchings in k-uniform planar hypergraphs is #P-hard. However, for $k \geq 5$, there is a big surprise that this problem is again computable in polynomial time. The formal statement is given in Theorem 9.28.

These hypergraph perfect matching problems are special cases of a class of graph orientation problems, and their complexity will be derived by looking at this broader class of problems. These can be described as follows. Given a planar (ordinary) graph G, we consider permissible orientations of all edges satisfying certain requirements locally at each incident vertex. We specify two kinds of vertices. The first kind can be either a sink or a source vertex in every permissible orientation. The second kind allows exactly one incoming edge. The goal is to count the number of permissible orientations, i.e., orientations that satisfy these constraints. Clearly this can be expressed as a bipartite Holant problem on the edge–vertex incidence graph of G, where we assign $(\neq_2)$ on each edge, and either the Equality function or the ExactOne function on each vertex. A vertex with an Equality taking the value 0 or 1 corresponds to a source or a sink, respectively. A vertex with an ExactOne is a vertex that allows exactly one incoming edge. Thus this problem can be expressed as Pl-Holant$(\neq_2 |$ Equality, ExactOne$)$. Under a complex-valued holographic transformation $Z = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}$, we have $(\neq_2) = (=_2)Z^{\otimes 2}$. It follows that the problem can also be expressed in a nonbipartite Holant form Holant$(\mathcal{F})$ under a Z-transformation, where $\mathcal{F}$ consists of the Z transformations of Equality and ExactOne. The problem of counting the number of perfect matchings

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

in a hypergraph is the special case of Pl-Holant($\neq_2$| EQUALITY, EXACTONE), where the connections between EQUALITY and EXACTONE via edges labeled ($\neq_2$) form a bipartite graph. In this case, at every vertex with an EQUALITY we absorb all the incident edges with ($\neq_2$) into this EQUALITY, making it an EQUALITY again. Then we obtain the bipartite edge–vertex incidence graph B_H of a hypergraph H: Vertices labeled EXACTONE are vertices of H, and vertices labeled EQUALITY are hyperedges of H, and all connections are by ($=_2$). A valid configuration is to choose a subset S of hyperedges (set to 1), such that every vertex of H is incident to exactly one hyperedge in S.

Suppose vertices assigned EQUALITY functions all have degree k. If $k = 2$, then the bipartite problem Pl-Holant($\neq_2$| EXACTONE, ($=_k$)) can be solved by Kasteleyn's algorithm since every signature is realizable as a matchgate. We show that this problem is #P-hard if $k = 3$ or $k = 4$, but is tractable again if $k \geq 5$. The algorithm involves a recursive procedure that simplifies the instance until it can be solved by known algorithms, including Kasteleyn's algorithm. The algorithm crucially uses global topological properties of a planar graph, in particular Euler's characteristic formula. If the graph is not planar, then this algorithm does not work, and indeed the problem is #P-hard over general graphs.

More generally, we allow vertices of arbitrary degrees to be assigned EQUALITY, with at least one vertex of degree ≥ 3. It turns out that the complexity depends on the greatest common divisor (gcd) of the arities of the EQUALITY functions present. The problem is tractable if gcd ≥ 5 and #P-hard if gcd ≤ 4. Note that it is not the minimum arity k of the EQUALITY functions that is the determining factor, but rather it is the gcd of the arities. (We allow multiple edges in a planar signature grid for Pl-Holant($\neq_2$| EQUALITY, EXACTONE). If $k \geq 6$ and $d \geq 3$, a planar (k, d)-biregular bipartite graph must have multiple edges. This follows from Euler's formula; see Lemma 9.24.)

Most of the reductions in previous Holant dichotomy theorems [HL16, CGW16] do not hold for planar graphs, so new techniques must be developed. An important ingredient in the proof of the Holant dichotomy Theorem 7.19 is Theorem 6.16, the #CSP^d dichotomy by Huang and Lu. Here #CSP^d denotes #CSP where every variable appears a multiple of d times. However, the proof of Theorem 6.16 is not planar, because it uses a pinning technique, and multiple copies of an instance are created and vertices are connected across different copies, which violates planarity. Moreover, as a consequence of the new dichotomy in this chapter, this violation of planarity is unavoidable. Owing to our newly discovered tractable problems, the putative form of a planar #CSP^d dichotomy stating that a holographic reduction to Kasteleyn's algorithm is universal, is *false* when $d \geq 5$. Nevertheless, it turns out that a dichotomy for

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

planar #CSP^2 in the putative form is true. Obtaining a dichotomy for planar #CSP^2 turns out to be an essential step because it captures a significant fraction of planar Holant problems either directly or through reductions. It turns out that one can prove the planar Holant dichotomy without having a planar #CSP^d dichotomy for $d \geq 3$, which is still open. An outline of the proof of the planar #CSP^2 dichotomy will be given.

Owing to the length of the proof of this planar Holant dichotomy, we will omit most of the #P-hardness proof of this dichotomy, as well as the proof of the planar #CSP^2 dichotomy. Our presentation will focus on the polynomial time algorithm for the new planar tractable problems and related issues.

9.2 Some Known Results

We consolidate some known results proved in previous chapters that will be useful for the discussion in this chapter.

Recall that a signature set $\mathcal{F}$ is $\mathscr{C}$-transformable if there exists a $T \in \mathbf{GL}_2(\mathbb{C})$ such that $[1, 0, 1]T^{\otimes 2} \in \mathscr{C}$ and $\mathcal{F} \subseteq T\mathscr{C}$. In this chapter the relevant classes $\mathscr{C}$ for planar Holant tractable problems will be $\mathscr{C} = \mathscr{A}$, $\mathscr{P}$ and $\mathscr{M}$. The putative form of the classification would state that these are the only tractable families for planar Holant problems. This turns out to be inadequate for the final classification; however, we will first need to get a good understanding of $\mathscr{C}$-transformable classes, especially the $\mathscr{M}$-transformable classes.

Here we list several known dichotomies. The first is the dichotomy for Holant problems, Theorem 7.19.

Theorem 9.2. *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. *$\mathcal{F} \subseteq \mathscr{V}^{\sigma} \cup \{f \in \mathscr{R}_2^{\sigma} \mid \text{arity}(f) = 2\}$ for $\sigma \in \{+, -\}$.*
5. *All nondegenerate signatures in $\mathcal{F}$ are in $\mathscr{R}_2^{\sigma}$ for $\sigma \in \{+, -\}$.*

Next we restate a dichotomy for planar Holant problems with additional restrictions. This is a combination of Theorem 6.3 and Theorem 8.12 for arity 3 and 4, respectively. This theorem forms the base case of an inductive proof of our single signature dichotomy, Theorem 9.17.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

Theorem 9.3. *If f is a nondegenerate, symmetric, complex-valued signature of arity* 3 *or* 4 *in Boolean variables, then* Pl-Holant(f) *is* #P*-hard unless f satisfies one of the following conditions, in which case the problem is computable in polynomial time:*

1. Holant(f) *is tractable (i.e., f is $\mathscr{A}$-transformable, $\mathscr{P}$-transformable, or vanishing).*
2. *f is $\mathscr{M}$-transformable.*

Next is a dichotomy theorem about spin systems with prescribed vertex degrees.

Theorem 9.4. *Let $S \subseteq \mathbb{Z}^+$ containing some $r \geq 3$, let $\mathcal{G} = \{=_k \mid k \in S\}$, and let $d = \gcd(S)$. Further suppose that $f_0, f_1, f_2 \in \mathbb{C}$. Then* Pl-Holant($[f_0, f_1, f_2] \mid \mathcal{G}$) *is* #P*-hard unless one of the following conditions holds:*

1. $f_0 f_2 = f_1^2$.
2. $f_0 = f_2 = 0$.
3. $f_1 = 0$.
4. $f_0 f_2 = -f_1^2$ *and* $f_0^d = -f_2^d \neq 0$.
5. $f_0^d = f_2^d \neq 0$.

In all exceptional cases, the problem is computable in polynomial time.

This theorem follows from Theorem 5.3. The tractability is straightforward after possibly a holographic transformation of the form $\begin{bmatrix} 1 & 0 \\ 0 & \omega \end{bmatrix}$ for some $\omega^{2d} = 1$ (for the matchgate class $\mathscr{M}$) or $\omega^{4d} = 1$ (for the affine class $\mathscr{A}$), where $d = \gcd(S)$. The #P-hardness follows from applying Theorem 5.3 individually to each single arity k for $(=_k) \in \mathcal{G}$.

Theorem 9.4 is explicit and easy to apply. Conceptually, it can be restated as Theorem 9.4′, which supports the putative form of the dichotomy that a holographic reduction to Kasteleyn's algorithm is universal. Note that this is for a single signature, and is consistent with the single signature dichotomy, Theorem 9.17. Denote by $\mathcal{T}_k = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & \omega \end{bmatrix} \mid \omega^k = 1 \right\}$, a set of diagonal matrices of order dividing k.

Theorem 9.4′. *Let $S \subseteq \mathbb{Z}^+$ contain $k \geq 3$, let $\mathcal{G} = \{=_k \mid k \in S\}$, and let $d = \gcd(S)$. Further suppose that f is a nondegenerate, symmetric, complex-valued binary signature in Boolean variables. Then* Pl-Holant($f \mid \mathcal{G}$) *is* #P*-hard unless f satisfies one of the following conditions, in which case, the problem is computable in polynomial time:*

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

1. *There exists* $T \in \mathcal{T}_{4d}$ *such that* $T^{\otimes 2} f \in \mathscr{A}$.
2. $f \in \mathscr{P}$.
3. *There exists* $T \in \mathcal{T}_{2d}$ *such that* $T^{\otimes 2} f \in \widehat{\mathscr{M}}$.

Lastly, we quote the Pl-#CSP dichotomy, Theorem 8.34. It also supports the putative form of a dichotomy, which states that holographic algorithms using matchgates followed by the Kasteleyn's algorithm is a universal strategy.

Theorem 9.5. *Let* $\mathcal{F}$ *be any set of symmetric, complex-valued signatures in Boolean variables. Then* Pl-#CSP($\mathcal{F}$) *is* #P*-hard unless* $\mathcal{F} \subseteq \mathscr{A}$, $\mathcal{F} \subseteq \mathscr{P}$, *or* $\mathcal{F} \subseteq \widehat{\mathscr{M}}$, *in which case the problem is computable in polynomial time.*

9.3 $\mathscr{A}$-, $\mathscr{P}$-, and $\mathscr{M}$-transformable Signatures

In this section, we first restate the properties of $\mathscr{A}$- and $\mathscr{P}$-transformable signatures. Then we investigate the properties of $\mathscr{M}$-transformable signatures. Throughout, we define $\alpha = \frac{1+\mathrm{i}}{\sqrt{2}} = \sqrt{\mathrm{i}} = e^{\frac{\pi \mathrm{i}}{4}}$ and use $\mathbf{O}_2(\mathbb{C})$ to denote the group of 2×2 orthogonal matrices over $\mathbb{C}$.

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

We summarize some results from Section 7.7. The following definition summarizes Definitions 7.39, 7.40, and 7.42.

Definition 9.6. A symmetric signature f of arity n is in, respectively, $\mathscr{A}_1$, or $\mathscr{A}_2$, or $\mathscr{A}_3$ if there exist an $H \in \mathbf{O}_2(\mathbb{C})$ and a nonzero constant $c \in \mathbb{C}$ such that f has the form, respectively,

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

or

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

or

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

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

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

For $k \in \{1, 2, 3\}$, when such an orthogonal H exists, we say that $f \in \mathscr{A}_k$ with transformation H. If $f \in \mathscr{A}_k$ with I_2, then we say f is in the canonical form of $\mathscr{A}_k$.

Lemma 7.41 characterizes the signatures in $\mathscr{A}_2$. We restate it as follows.

Lemma 9.7. *Let f be a symmetric signature of arity n. Then $f \in \mathscr{A}_2$ if and only if $f = c\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}^{\otimes n} + d\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}^{\otimes n}$ for some nonzero $c, d \in \mathbb{C}$. In other words, $f \in \mathscr{A}_2$ iff $f = Z[c, 0, \ldots, 0, d]$.*

Membership in these three sets characterize the $\mathscr{A}$-transformable signatures. We restate Lemma 7.43.

Lemma 9.8. *Let f be a nondegenerate symmetric signature. Then f is $\mathscr{A}$-transformable if and only if $f \in \mathscr{A}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$.*

There is a similar characterization for $\mathscr{P}$-transformable signatures. We restate the definitions for $\mathscr{P}_1$ and $\mathscr{P}_2$ (see Definition 7.45).

Definition 9.9. A symmetric signature f of arity n is in $\mathscr{P}_1$ if there exist an $H \in \mathbf{O}_2(\mathbb{C})$ and a nonzero $c \in \mathbb{C}$ 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 define $\mathscr{P}_2 = \mathscr{A}_2$.

For $k \in \{1, 2\}$, when such an H exists, we say that $f \in \mathscr{P}_k$ with transformation H. If $f \in \mathscr{P}_k$ with I_2, then we say f is in the canonical form of $\mathscr{P}_k$.

Membership in these two sets characterize the $\mathscr{P}$-transformable signatures. We restate Lemma 7.46.

Lemma 9.10. *Let f be a nondegenerate symmetric signature. Then f is $\mathscr{P}$-transformable if and only if $f \in \mathscr{P}_1 \cup \mathscr{P}_2$.*

9.3.2 Characterization of $\mathscr{M}$-transformable Signatures

Now we develop a similar theory for the $\mathscr{M}$-transformable signatures. Recall that a signature set $\mathcal{F}$ is $\mathscr{M}$-transformable, if there exists a $T \in \mathbf{GL}_2(\mathbb{C})$ such that $[1, 0, 1]T^{\otimes 2} \in \mathscr{M}$, and $\mathcal{F} \subseteq T\mathscr{M}$. Since $[1, 0, 1]$ is symmetric, $[1, 0, 1]T^{\otimes 2}$ is also symmetric. However, it is unnecessary to consider all binary signatures in $\mathscr{M}$. We can normalize via *right* multiplication by elements in

$$\mathrm{Stab}(\mathscr{M}) = \{T \in \mathbf{GL}_2(\mathbb{C}) | \mathscr{M}T \subseteq \mathscr{M}\} = \{T \in \mathbf{GL}_2(\mathbb{C}) | T\mathscr{M} \subseteq \mathscr{M}\},$$

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

the stabilizer group of $\mathscr{M}$. The right and left stabilizer groups of $\mathscr{M}$ coincide, since they are both 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 $X = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}$.

After this normalization, it is enough to consider only two cases (case 1 and case 3) in Proposition 7.35, which is restated here.

Proposition 9.11. *Let $T \in \mathbb{C}^{2\times 2}$ be a matrix. Then the following hold:*

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

It follows that

Lemma 9.12. *Let $\mathcal{F}$ be a set of signatures. Then $\mathcal{F}$ is $\mathscr{M}$-transformable if and only if $\mathcal{F} \subseteq Z\mathscr{M}$ for $Z = \begin{bmatrix}1 & 1\\ \mathrm{i} & -\mathrm{i}\end{bmatrix}$, or there exists an $H \in \mathbf{SO}_2(\mathbb{C})$ such that $\mathcal{F} \subseteq H\mathscr{M}$.*

Proof. Sufficiency is easily verified by checking that $(=_2)$ is transformed into $\mathscr{M}$ in both cases. Indeed $(=_2)Z^{\otimes 2} = (\neq_2) \in \mathscr{M}$, and H leaves $(=_2) \in \mathscr{M}$ unchanged.

Suppose $\mathcal{F}$ is $\mathscr{M}$-transformable. The nondegenerate binary signatures in $\mathscr{M}$ 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 $\begin{bmatrix}1 & 0\\ 0 & \nu^{-\frac{1}{2}}\end{bmatrix} \in \mathrm{Stab}(\mathscr{M})$. Thus, we only need to consider $[1, 0, 1]$ and $[0, 1, 0]$. Now we apply Proposition 9.11.

1. If $(=_2)T^{\otimes 2} = [1, 0, 1]$, then by case 1 of Proposition 9.11, we have $T \in \mathbf{O}_2(\mathbb{C})$. If $T \in \mathbf{SO}_2(\mathbb{C})$, then we are done with $H = T$. Otherwise, $T \in \mathbf{O}_2(\mathbb{C}) - \mathbf{SO}_2(\mathbb{C})$. We want to find an $H \in \mathbf{SO}_2(\mathbb{C})$ such that $\mathcal{F} \subseteq H\mathscr{M}$. Let $H = T\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \in \mathbf{SO}_2(\mathbb{C})$. Then
$$\mathcal{F} \subseteq T\mathscr{M} = T\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\mathscr{M} = H\mathscr{M},$$
since $\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \in \mathrm{Stab}(\mathscr{M})$.
2. If $(=_2)T^{\otimes 2} = [0, 1, 0]$, then by case 2 of Proposition 9.11, 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 $\mathcal{F} \subseteq H \begin{bmatrix}1 & 1\\ \mathrm{i} & -\mathrm{i}\end{bmatrix}\mathscr{M}$.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

Furthermore, if $H = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \in \mathbf{SO}_2(\mathbb{C})$, then $a^2 + b^2 = 1$ and

$$\mathcal{F} \subseteq H \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \mathscr{M} = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \begin{bmatrix} a + b\mathrm{i} & 0 \\ 0 & a - b\mathrm{i} \end{bmatrix} \mathscr{M} = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \mathscr{M},$$

since

$$H \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \begin{bmatrix} a + b\mathrm{i} & 0 \\ 0 & a - b\mathrm{i} \end{bmatrix}$$

and

$$\begin{bmatrix} a + b\mathrm{i} & 0 \\ 0 & a - b\mathrm{i} \end{bmatrix} \in \mathrm{Stab}(\mathscr{M}).$$

Otherwise, $H = \begin{bmatrix} a & b \\ b & -a \end{bmatrix} \in \mathbf{O}_2(\mathbb{C}) - \mathbf{SO}_2(\mathbb{C})$, so $a^2 + b^2 = 1$ and

$$\mathcal{F} \subseteq H \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \mathscr{M} = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \begin{bmatrix} 0 & a - b\mathrm{i} \\ a + b\mathrm{i} & 0 \end{bmatrix} \mathscr{M} = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \mathscr{M},$$

since

$$H \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \begin{bmatrix} 0 & a - b\mathrm{i} \\ a + b\mathrm{i} & 0 \end{bmatrix}$$

and

$$\begin{bmatrix} 0 & a - b\mathrm{i} \\ a + b\mathrm{i} & 0 \end{bmatrix} \in \mathrm{Stab}(\mathscr{M}). \qquad \square$$

We use four sets to characterize the $\mathscr{M}$-transformable signatures. The notation Sym is from Definition 7.3.

Definition 9.13. A symmetric signature f of arity n is in, respectively, $\mathscr{M}_1$, or $\mathscr{M}_2$, or $\mathscr{M}_3$, or $\mathscr{M}_4$ if there exist an $H \in \mathbf{O}_2(\mathbb{C})$ and nonzero constants $c, \gamma \in \mathbb{C}$ such that f has the form, respectively,

$$cH^{\otimes n} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}^{\otimes n} \pm i^n \begin{bmatrix} 1 \\ -1 \end{bmatrix}^{\otimes n} \right),$$

or

$$cH^{\otimes n} \left(\begin{bmatrix} 1 \\ \gamma \end{bmatrix}^{\otimes n} \pm \begin{bmatrix} 1 \\ -\gamma \end{bmatrix}^{\otimes n} \right),$$

or

$$cH^{\otimes n}\, \mathrm{Sym}_n^{n-1} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right),$$

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

or

$$cH^{\otimes n}\operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix};\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}\right).$$

For $k \in \{1, 2, 3, 4\}$, when such an H exists, we say that $f \in \mathscr{M}_k$ with transformation H. If $f \in \mathscr{M}_k$ with I_2, then we say f is in the canonical form of $\mathscr{M}_k$.

Notice that $\left\{\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}, \begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}\right\}$ is set-wise invariant under any transformation in $\mathbf{O}_2(\mathbb{C})$ up to nonzero constants. Using this fact, the following lemma gives a characterization of $\mathscr{M}_4$. It says that any signature in $\mathscr{M}_4$ is essentially in canonical form.

Lemma 9.14. *Let f be a symmetric signature of arity n. Then $f \in \mathscr{M}_4$ if and only if*

$$f = c\operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix};\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}\right) \text{ or } f = c\operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix};\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}\right)$$

for some nonzero constant $c \in \mathbb{C}$.

Proof. Suppose $f \in \mathscr{M}_4$, so that $f = cH^{\otimes n}\operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix};\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}\right)$. If $H \in \mathbf{SO}_2(\mathbb{C})$, then $H = \begin{bmatrix}a & b\\ -b & a\end{bmatrix}$ for some $a, b \in \mathbb{C}$ such that $a^2 + b^2 = 1$. Since $H\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix} = (a + b\mathrm{i})\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}$ and $H\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix} = (a - b\mathrm{i})\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}$, it follows that $f = c(a + b\mathrm{i})^{n-1}(a - b\mathrm{i})\operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix};\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}\right)$. Otherwise, $H \in \mathbf{O}_2(\mathbb{C}) - \mathbf{SO}_2(\mathbb{C})$, so $H = \begin{bmatrix}a & b\\ b & -a\end{bmatrix}$ for some $a, b \in \mathbb{C}$ such that $a^2 + b^2 = 1$. Then $f = c(a + b\mathrm{i})(a - b\mathrm{i})^{n-1}\operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix};\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}\right)$.

Now suppose

$$f = c\operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix};\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix}\right)$$

or $f = c\operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}1\\ -\mathrm{i}\end{bmatrix};\begin{bmatrix}1\\ \mathrm{i}\end{bmatrix}\right)$. The first case is already in the standard form of $\mathscr{M}_4$. In the second case, we pick $H = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \in \mathbf{O}_2(\mathbb{C})$. Then $H^{\otimes n}f$ is in the standard form of $\mathscr{M}_4$. □

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

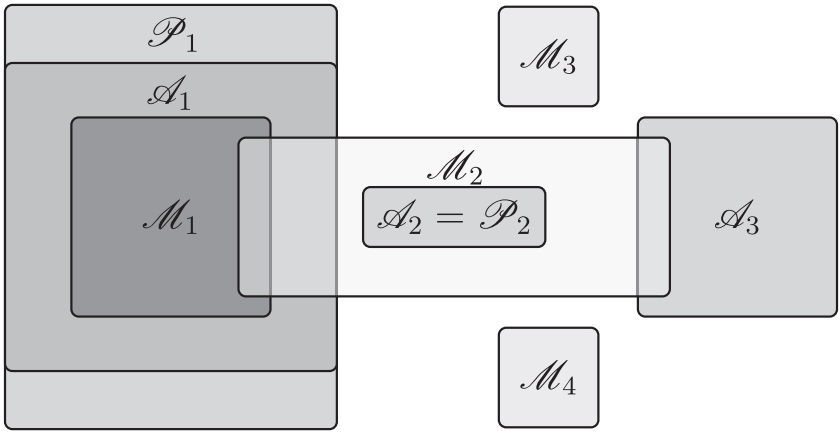


Figure 9.1. Relationships among $\mathscr{A}_1$, $\mathscr{A}_2$, $\mathscr{A}_3$, $\mathscr{P}_1$, $\mathscr{P}_2$, $\mathscr{M}_1$, $\mathscr{M}_2$, $\mathscr{M}_3$, and $\mathscr{M}_4$, for signatures of arity $n \geq 3$. Note that $\mathscr{P}_1 \cap \mathscr{M}_2 \subseteq \mathscr{A}_1$.

We can further split $\mathscr{M}_4$ into two classes $\mathscr{M}_4^{\pm}$. Define

$$\mathscr{M}_4^{\pm} = \left\{ f \;\middle|\; f = c\,\mathrm{Sym}_n^{n-1}\left(\begin{bmatrix}1\\ \pm\mathrm{i}\end{bmatrix}; \begin{bmatrix}1\\ \mp\mathrm{i}\end{bmatrix}\right) \right\}.$$

In other words, $\mathscr{M}_4^{+}$ contains signatures of the form $Z^{\otimes n}[0, 1, 0, \ldots, 0]$ and $\mathscr{M}_4^{-}$ contains signatures of the form $Z^{\otimes n}[0, \ldots, 0, 1, 0]$ up to a scalar, where $Z = \begin{bmatrix}1 & 1\\ \mathrm{i} & -\mathrm{i}\end{bmatrix}$. We will denote $[0, 1, 0, \ldots, 0]$ of arity k by $\textsc{ExactOne}_k$, and $[0, \ldots, 0, 1, 0]$ of arity k by $\textsc{AllButOne}_k$. Note that these are precisely the Perfect Matching signatures and corresponding reversals.

Notice that $\mathscr{M}_1 \subset \mathscr{A}_1 \subset \mathscr{P}_1$ and $\mathscr{A}_2 = \mathscr{P}_2 \subset \mathscr{M}_2$. See Figure 9.1 for a visual description of the relationships among sets.

Next we show that $\mathscr{M}_k$ for $k = 1, 2, 3, 4$ captures all $\mathscr{M}$-transformable signatures.

Lemma 9.15. *Let f be a nondegenerate symmetric signature. Then f is $\mathscr{M}$-transformable if and only if $f \in \mathscr{M}_1 \cup \mathscr{M}_2 \cup \mathscr{M}_3 \cup \mathscr{M}_4$.*

Proof. Assume that f is $\mathscr{M}$-transformable of arity n. By applying Lemma 9.12 to $\{f\}$, we have $f \in \begin{bmatrix}1 & 1\\ \mathrm{i} & -\mathrm{i}\end{bmatrix} \mathscr{M}$ or there exists an $H \in \mathbf{SO}_2(\mathbb{C})$ such that $f \in H\mathscr{M}$. Proposition 8.1 lists all the symmetric signatures in $\mathscr{M}$. Since we are interested only in nondegenerate signatures, we consider only a, b, and λ that are nonzero. Now we consider the possible cases.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

1. Suppose $f \in \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix} \mathscr{M}$.
   - Further, suppose
   $$f = \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}^{\otimes n} \left(\begin{bmatrix} a \\ b \end{bmatrix}^{\otimes n} \pm \begin{bmatrix} a \\ -b \end{bmatrix}^{\otimes n} \right)$$
   for some nonzero $a, b \in \mathbb{C}$. Let $T = \frac{1-\mathrm{i}}{2} \begin{bmatrix} u & v \\ v & -u \end{bmatrix}$, where $u = a + b\mathrm{i}$ and $v = \mathrm{i}(a - b\mathrm{i})$. Then
   $$f = T^{\otimes n} \left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}^{\otimes n} \pm \mathrm{i}^n \begin{bmatrix} 1 \\ -1 \end{bmatrix}^{\otimes n} \right).$$
   Since $T \in \mathbf{O}_2(\mathbb{C})$ up to a nonzero factor of $\sqrt{2ab}$, we have $f \in \mathscr{M}_1$.
   - Further, suppose
   $$f = \lambda \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}^{\otimes n} \mathrm{Sym}_n^{n-1} \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}; \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right)$$
   for some nonzero $\lambda \in \mathbb{C}$. Then we have $f = \lambda \, \mathrm{Sym}_n^{n-1} \left(\begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix}; \begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix} \right)$, so $f \in \mathscr{M}_4$.
   - Further, suppose
   $$f = \lambda \begin{bmatrix} 1 & 1 \\ \mathrm{i} & -\mathrm{i} \end{bmatrix}^{\otimes n} \mathrm{Sym}_n^{n-1} \left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}; \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)$$
   for some nonzero $\lambda \in \mathbb{C}$. Then we have $f = \lambda \, \mathrm{Sym}_n^{n-1} \left(\begin{bmatrix} 1 \\ -\mathrm{i} \end{bmatrix}; \begin{bmatrix} 1 \\ \mathrm{i} \end{bmatrix} \right)$, so $f \in \mathscr{M}_4$ by Lemma 9.14.
2. Suppose $f \in H\mathscr{M}$.
   - Further, suppose
   $$f = H^{\otimes n} \left(\begin{bmatrix} a \\ b \end{bmatrix}^{\otimes n} \pm \begin{bmatrix} a \\ -b \end{bmatrix}^{\otimes n} \right)$$
   for some nonzero $a, b \in \mathbb{C}$. Then we have
   $$f = a^n H^{\otimes n} \left(\begin{bmatrix} 1 \\ \gamma \end{bmatrix}^{\otimes n} \pm \begin{bmatrix} 1 \\ -\gamma \end{bmatrix}^{\otimes n} \right),$$
   where $\gamma = \frac{b}{a}$, so $f \in \mathscr{M}_2$.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

- Further, suppose

$$f = \lambda H^{\otimes n} \operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}1\\0\end{bmatrix}; \begin{bmatrix}0\\1\end{bmatrix}\right)$$

for some nonzero $\lambda \in \mathbb{C}$. Then $f \in \mathscr{M}_3$.

- Further, suppose

$$f = \lambda H^{\otimes n} \operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}0\\1\end{bmatrix}; \begin{bmatrix}1\\0\end{bmatrix}\right)$$

for some nonzero $\lambda \in \mathbb{C}$. Let $H' = H\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix} \in \mathbf{O}_2(\mathbb{C})$. Then we have

$$f = \lambda H'^{\otimes n} \operatorname{Sym}_n^{n-1}\left(\begin{bmatrix}1\\0\end{bmatrix}; \begin{bmatrix}0\\1\end{bmatrix}\right),$$

so $f \in \mathscr{M}_3$.

Conversely, if there exists a matrix $H \in \mathbf{O}_2(\mathbb{C})$ such that $H^{\otimes n} f$ is in one of the canonical forms of $\mathscr{M}_1$, $\mathscr{M}_2$, $\mathscr{M}_3$, or $\mathscr{M}_4$, then one can directly check that f is $\mathscr{M}$-transformable. In fact, the transformations that we applied above are all invertible, except for $\mathscr{M}_1$, if the given orthogonal transformation is of the form $\begin{bmatrix}u & -v\\v & u\end{bmatrix}$, we do $\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$ first followed by $\begin{bmatrix}u & v\\v & -u\end{bmatrix}$. □

Recall from Section 7.9 that a symmetric 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 $0 \le k \le n-2$,

$$af_k + bf_{k+1} + cf_{k+2} = 0.$$

Lemma 7.56 says that any $\mathscr{A}$- or $\mathscr{P}$-transformable signature $f \in \mathscr{P}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3$ of arity $n \ge 3$ satisfies a second-order recurrence relation with nonzero discriminant $b^2 - 4ac$.

This can be extended to $\mathscr{M}$-transformable signatures.

Exercise: Let f be a symmetric $\mathscr{M}$–transformable signature of arity $n \ge 3$. Then f satisfies a second-order recurrence relation. The discriminant is nonzero iff $f \in \mathscr{M}_1 \cup \mathscr{M}_2$. The recurrence relation for $f \in \mathscr{M}_4$ has a double root belonging to $\{\mathrm{i}, -\mathrm{i}\}$. The recurrence relation for $f \in \mathscr{M}_3$ has a double root not belonging to $\{\mathrm{i}, -\mathrm{i}\}$.

Hence $\mathscr{M}_1 \cup \mathscr{M}_2$, $\mathscr{M}_3$ and $\mathscr{M}_4$ for signatures of arity ≥ 3 are pairwise disjoint.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

9.4 Dichotomy for Pl-#CSP^2

In this section, we state the dichotomy for Pl-#CSP^2. We provide only a sketch of the proof here; the interested reader can find the full proof in Part II of [CFGW15]. Let

$$\mathcal{T}_k = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & \omega \end{bmatrix} \in \mathbb{C}^{2\times 2} \,\middle|\, \omega^k = 1 \right\}.$$

Theorem 9.16. *Let $\mathcal{F}$ be a set of symmetric signatures. Then* Pl-#CSP$^2(\mathcal{F})$ *is* #P*-hard unless $\mathcal{F}$ satisfies one of the following conditions:*

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

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

Proof Sketch. We first define some tractable families of signatures specific to the Pl-#CSP^2 framework. Let

$$\mathcal{T}_k = \mathcal{T}_{2k} \setminus \mathcal{T}_k = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & \omega \end{bmatrix} \,\middle|\, \omega^k = -1 \right\}.$$

Let $\mathscr{A}^\dagger = \mathcal{T}_4\mathscr{A}$ and $\widehat{\mathscr{M}}^\dagger = \mathcal{T}_2\widehat{\mathscr{M}}$ be the sets of signatures transformed by $\mathcal{T}_4$ from the affine family $\mathscr{A}$ and transformed by $\mathcal{T}_2$ from $\widehat{\mathscr{M}}$, respectively. Let

$$\widetilde{\mathscr{A}} = \mathscr{A} \cup \mathscr{A}^\dagger \qquad \text{and} \qquad \widetilde{\mathscr{M}} = \widehat{\mathscr{M}} \cup \widehat{\mathscr{M}}^\dagger.$$

Thus $\widetilde{\mathscr{A}} = \mathscr{A} \cup \begin{bmatrix} 1 & 0 \\ 0 & e^{\pi i/4} \end{bmatrix}\mathscr{A}$ and $\widetilde{\mathscr{M}} = \widehat{\mathscr{M}} \cup \begin{bmatrix} 1 & 0 \\ 0 & \mathrm{i} \end{bmatrix}\widehat{\mathscr{M}}$. These are the $\mathscr{A}$-transformable and $\mathscr{M}$-transformable signatures for Pl-#CSP^2. One can show that $\widetilde{\mathscr{A}}$ covers case 1 above, and $\widetilde{\mathscr{M}}$ covers case 3. The proof will revolve around these tractable classes.

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

The first step is to prove the dichotomy theorem for Pl-#CSP$^2(\mathcal{F})$ when there is at least one nonzero signature of *odd* arity in $\mathcal{F}$. In this case, we can simulate Pl-#CSP$(\mathcal{F})$ by Pl-#CSP$^2(\mathcal{F})$ if $\mathcal{F}$ includes a unary signature $[a, b]$ with $ab \neq 0$. Then we can apply the known dichotomy Theorem 9.5 for Pl-#CSP. However, this strategy (provably) *cannot* work when every signature in $\mathcal{F}$ satisfies the *parity* constraint (in particular, $\mathcal{F}$ does not contain any unary signature $[a, b]$ with $ab \neq 0$). In that case we employ other means. This first step of the proof is relatively uncomplicated.

The second step is to deal with the case when all nonzero signatures in $\mathcal{F}$ have even arity. This is where the real difficulties lie. In this case it is impossible

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

to directly construct *any* unary signature, which makes the proof more difficult. But we prove that there is a way to simulate Pl-#CSP($\mathcal{F}$) by Pl-#CSP^2($\mathcal{F}$) in a *global* fashion, *if* $\mathcal{F}$ includes some tensor power of the form $[a, b]^{\otimes 2}$, where $ab \neq 0$. Moreover, we have a lucky break (for the complexity of the proof) if $\mathcal{F}$ includes a signature that is in $\widehat{\mathscr{M}} \setminus (\mathscr{P} \cup \widetilde{\mathscr{A}})$. In this case, we can construct a special binary signature, and obtain $[1, 1]^{\otimes 2}$ by interpolation. This proof uses some properties of *cyclotomic fields*. This simplifies the proof greatly. For all other cases when $\mathcal{F}$ has only even-arity signatures, the proof gets going in earnest – we will attempt an induction on the arity of signatures.

The lowest arity of this induction will be 2. We will try to reduce the arity to 2 whenever possible; however, for many cases an arity reduction to 2 destroys the #P-hardness at hand. Therefore, the true base case of this induction proof of Pl-#CSP^2 starts with arity 4. Consequently, we will first prove a dichotomy theorem for Pl-#CSP$^2(f)$, where f is a signature of arity 4. Several tools will be used. These include the rank criterion for redundant signatures Corollary 7.27, a dichotomy for planar spin systems Theorem 9.4, and a trick called the *Three Stooges* by domain pairing.

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

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

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

the proof of this Pl-#CSP^2 dichotomy for a single signature of even arity, the no-mixing lemma becomes *unconditional*.

Finally the dichotomy for a single signature of even arity is logically extended to a dichotomy theorem for Pl-#CSP$^2(\mathcal{F})$ where all signatures in $\mathcal{F}$ have even arity. Together with the first main step when $\mathcal{F}$ contains some nonzero signature of odd arity, this completes the proof of Theorem 9.16. □

9.5 Single-Signature Dichotomy

By Lemmas 9.8, 9.10 and 9.15, f is $\mathscr{A}$-, or $\mathscr{P}$-, or $\mathscr{M}$-transformable, or vanishing if and only if

$$f \in \mathscr{A}_1 \cup \mathscr{A}_2 \cup \mathscr{A}_3 \cup \mathscr{P}_1 \cup \mathscr{P}_2 \cup \mathscr{M}_1 \cup \mathscr{M}_2 \cup \mathscr{M}_3 \cup \mathscr{M}_4 \cup \mathscr{V}.$$

However, since $\mathscr{M}_1 \subset \mathscr{A}_1 \subset \mathscr{P}_1$, $\mathscr{A}_2 = \mathscr{P}_2 \subset \mathscr{M}_2$, we have the equivalent expression

$$f \in \mathscr{P}_1 \cup \mathscr{M}_2 \cup \mathscr{A}_3 \cup \mathscr{M}_3 \cup \mathscr{M}_4 \cup \mathscr{V}.$$

In fact, for signatures f of arity $n \geq 3$, $f \in \mathscr{M}_4$ implies $f \in \mathscr{V}$. So for arity $n \geq 3$, we could further replace it by $f \in \mathscr{P}_1 \cup \mathscr{M}_2 \cup \mathscr{A}_3 \cup \mathscr{M}_3 \cup \mathscr{V}$. However, for the analysis of cases we will actually consider the equivalent expression

$$f \in \mathscr{P}_1 \cup \mathscr{P}_2 \cup (\mathscr{M}_2 \setminus \mathscr{P}_2) \cup \mathscr{A}_3 \cup \mathscr{M}_3 \cup \mathscr{M}_4 \cup (\mathscr{V} \setminus \mathscr{M}_4).$$

We state Theorem 9.17, the single-signature dichotomy for Pl-Holant problems. This theorem says that for a single signature the putative form of the Pl-Holant dichotomy holds. That is, for a nondegenerate symmetric signature f of arity ≥ 3, Pl-Holant(f) is either #P-hard, or f is $\mathscr{A}$-, or $\mathscr{P}$-, or $\mathscr{M}$-transformable, or vanishing.

Theorem 9.17. *If f is a nondegenerate symmetric signature of arity $n \geq 3$ with complex weights in Boolean variables, then* Pl-Holant(f) *is* #P*-hard unless $f \in \mathscr{P}_1 \cup \mathscr{M}_2 \cup \mathscr{A}_3 \cup \mathscr{M}_3 \cup \mathscr{M}_4 \cup \mathscr{V}$, in which case the problem is computable in polynomial time.*

The proof of Theorem 9.17 is by induction on the arity; but we will not give the details. In fact, the proof of Theorem 9.17 already has to investigate the mixing properties of various tractable signatures. Interested readers can find the proof in the paper [CFGW15].

Having established Theorem 9.17, the main question is for a *set* of tractable signatures what kinds of mixture are possible. A great deal of effort is spent trying to prove various tractable classes cannot mix, much in the same way as in the proof of the no-mixing lemma for Pl-#CSP^2 problems. Previously we had always found that two tractable signatures of different types cannot be mixed

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

together; i.e., any signature set containing both signatures leads to #P-hardness. For Pl-Holant problems, it turns out that this is no longer true for all the tractable classes.

But the non-mixing property is still true for $f \in \mathscr{P}_1 \cup (\mathscr{M}_2 \setminus \mathscr{P}_2) \cup \mathscr{A}_3 \cup \mathscr{M}_3$. The technique is similar to that of Section 7.7.2, and we can reduce the case to Theorem 9.16, the dichotomy for Pl-#CSP^2 problems. For example, suppose $\mathcal{F}$ contains some $f \in \mathscr{P}_1$ of arity $n \geq 3$, we can write it as $f = H[a, 0, \ldots, 0, b]$, where $H \in \mathbf{O}_2(\mathbb{C})$ and $ab \neq 0$. After a holographic transformation by H in Pl-Holant$(=_2 \mid \mathcal{F})$, we have Pl-Holant$(=_2 \mid H^{-1}\mathcal{F})$, where the generalized equality $[a, 0, \ldots, 0, b] \in H^{-1}\mathcal{F}$. Then we can follow the proof of Lemma 7.48, which remains valid for planar Holant problems, and obtain all equalities of even arity $\mathcal{EQ}_2$. Basically, by repeated self-loops we can get a generalized equality of arity 3 or 4. If it is 3, connecting two copies with one edge gives a generalized equality of arity 4. From arity 4, a self-loop gives a generalized equality of arity 2. Then we can get generalized equalities of all even arities. Finally we get $\mathcal{EQ}_2$ by interpolation, or in case the ratio is a root of unity, by a direct construction. Once we get $\mathcal{EQ}_2$ we can apply Theorem 9.16, the dichotomy for Pl-#CSP^2 problems.

Some more effort is needed to prove non-mixing for $\mathscr{V} \setminus \mathscr{M}_4$. So the real difficulty is the mixing question of $\mathscr{P}_2$ and $\mathscr{M}_4$.

The following lemma deals with a case of mixing property for some signatures from $\mathscr{P}_2$ and $\mathscr{M}_4$. There are three signatures, the first two signatures $Z^{\otimes 4}[0, 1, 0, 0, 0]$ and $Z^{\otimes 4}[0, 0, 0, 1, 0]$ both belong to $\mathscr{M}_4$. The third signature $Z^{\otimes 4}g$ is the Z-transformation of the conjunction $g(x_1, x_2, x_3, x_4) = (x_1 = x_3) \wedge (x_2 = x_4) \wedge (x_1 \neq x_2)$. This is the conjunction of $Z^{\otimes 2}[1, 0, 1]_{x_1, x_3}$, $Z^{\otimes 2}[1, 0, 1]_{x_2, x_4}$ and $Z^{\otimes 2}[0, 1, 0]_{x_1, x_2}$; each belongs to $\mathscr{P}_2$. This lemma is already needed in the proof of Theorem 9.17, and is part of the proof of the no-mixing for Pl-Holant problems. We will also use it to investigate the complexity of hypergraph perfect matching problems in Section 9.6. The lemma is stated in the form after the Z-transformation.

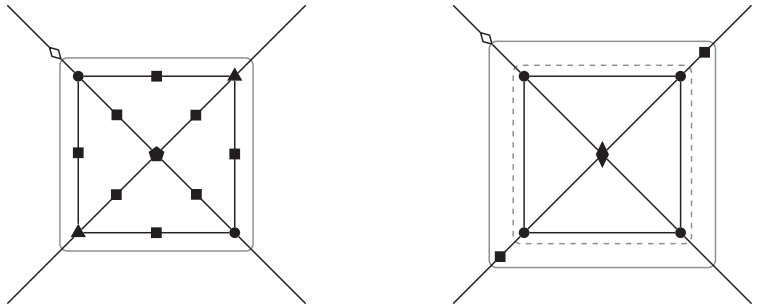
Lemma 9.18. *Let g be the arity* 4 *signature whose matrix is*

$$M_g = \begin{bmatrix} g_{0000} & g_{0010} & g_{0001} & g_{0011} \\ g_{0100} & g_{0110} & g_{0101} & g_{0111} \\ g_{1000} & g_{1010} & g_{1001} & g_{1011} \\ g_{1100} & g_{1110} & g_{1101} & g_{1111} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{9.1}$$

Then Pl-Holant$(\neq_2 \mid [0, 1, 0, 0, 0], [0, 0, 0, 1, 0], g)$ *is* #P*-hard.*

Proof. Recall that we list the columns of a signature matrix in reverse lexicographic order so that matrix product corresponds to linking the signature in a

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press



(a) $(\neq_2 \mid [0,1,0,0,0],[0,0,0,1,0],g)$-gate on right side (b) Simpler construction with the same signature

Figure 9.2. Two gadgets with the same signature used in Lemma 9.18.

chain. The only nonzero entries of M_g are $g_{0101} = g_{1010} = 1$. Thus g is a conditional crossover signature where in a planar cyclic order the two wires crossing each other must carry unequal values 0 and 1.

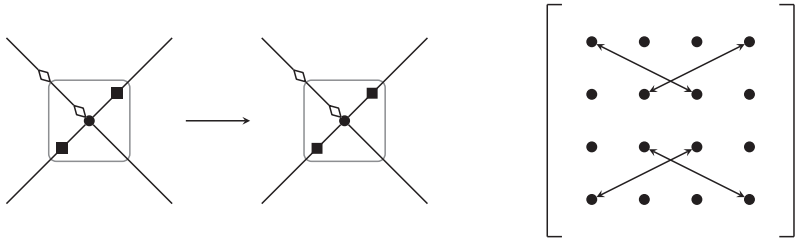
Consider the gadget in Figure 9.2a. We assign $[0, 0, 0, 1, 0]$ to the triangle vertices, $[0, 1, 0, 0, 0]$ to the circle vertices, g to the pentagon vertex, and $[0, 1, 0]$ to the square vertices. Let h be the signature of this gadget. If we append two binary DISEQUALITY signatures consecutively at the lower left and also at the upper right dangling edges, we do not change the signature h. Then we absorb four adjacent DISEQUALITY signatures surrounding each $[0, 0, 0, 1, 0]$ signature at the two triangle vertices, and we convert the ALLBUTONE_4 signature $[0, 0, 0, 1, 0]$ to the EXACTONE_4 signature $[0, 1, 0, 0, 0]$. If we also absorb the two DISEQUALITY signatures at the upper left and lower right positions of the pentagon vertex, the signature g there becomes $(=_4)$. This is now depicted as the diamond vertex in Figure 9.2b, which has the same signature h as in Figure 9.2a, where the circle vertices are still assigned $[0, 1, 0, 0, 0]$, the square vertices are still assigned $[0, 1, 0]$, and the diamond vertex is assigned $(=_4)$. To compute the signature h, first compute the signature h' of the inner gadget enclosed by the dashed line, which has signature matrix

$$M_{h'} = \begin{bmatrix} 3 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}.$$

Then by Figure 9.3, the signature matrix of h is

$$M_h = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

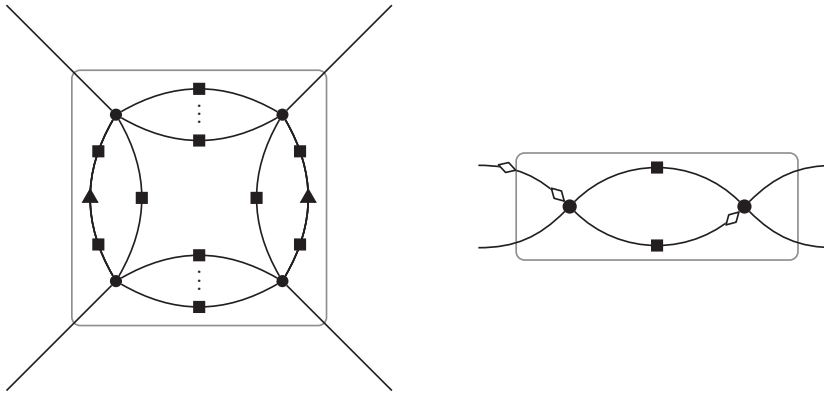


(a) Negating the second and fourth inputs (b) Movement of even Hamming weight entries

Figure 9.3. The movement of the even Hamming weight entries in the signature matrix of a quaternary signature under the negation of the second and fourth inputs (i.e., the square vertices are assigned [0, 1, 0]).

We need one more symmetrizing construction. Consider the gadget in Figure 9.4b. We assign h to the circle vertices and [0, 1, 0] to the square vertices. The signature of the resulting gadget is r with signature matrix M_r (see Figure 7.4 for the signature of a rotated copy of h that appears as the second circle vertex in Figure 9.4b), where

$$M_r = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 3 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right) \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 3 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 6 & 4 & 0 \\ 0 & 4 & 2 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$



(a) Gadget that realizes a partial crossover (b) Gadget with a useful signature matrix

Figure 9.4. Two quaternary gadgets used in the proof of Lemma 9.18.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

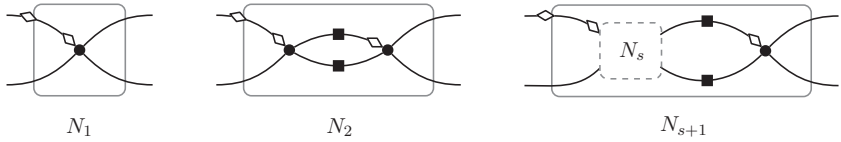


Figure 9.5. Linear recursive construction used for interpolation in a nonstandard basis.

Now we consider the signature

$$M_R = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 3 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

By a rotation (see Figure 7.4) R becomes R', with

$$M_{R'} = \begin{bmatrix} 0 & 0 & 0 & 3 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}.$$

The signature matrix $M_{R'}$ is redundant and the compressed signature matrix is nonsingular. Hence by Corollary 7.27 using the Z transformation, Pl-Holant($\neq_2 | R'$) is #P-hard. Therefore so is Pl-Holant($\neq_2 | R$).

Now we prove a reduction:

$$\text{Pl-Holant}(\neq_2 | R) \leq_{\text{T}} \text{Pl-Holant}(\neq_2 | r).$$

Consider an instance Ω of Pl-Holant($\neq_2 | R$). Suppose that R appears n times in Ω. We construct from Ω a sequence of instances Ω_s of Pl-Holant($\neq_2 | r$) indexed by $s \geq 1$. We obtain Ω_s from Ω by replacing each occurrence of R with the gadget N_s in Figure 9.5 with r assigned to the circle vertices and $[0, 1, 0]$ 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.

We can express the signature matrix of N_s as

$$M_{N_s} = X(XM_r)^s = XP\,\text{diag}(1, 4 + 2\sqrt{3}, 4 - 2\sqrt{3}, 1)^s P^{-1},$$

where

$$X = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & \sqrt{3} & -\sqrt{3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

Since $M_R = XP\,\mathrm{diag}(1, 1+\sqrt{3}, 1-\sqrt{3}, 1)P^{-1}$, we can view our construction of Ω_s as first replacing M_R with $XP\,\mathrm{diag}(1, 1+\sqrt{3}, 1-\sqrt{3}, 1)P^{-1}$, which does not change the Holant value, and then replacing the diagonal matrix $\mathrm{diag}(1, 1+\sqrt{3}, 1-\sqrt{3}, 1)$ with the diagonal matrix $\mathrm{diag}(1, 4+2\sqrt{3}, 4-2\sqrt{3}, 1)^s$.

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

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1+\sqrt{3} & 0 & 0 \\ 0 & 0 & 1-\sqrt{3} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{9.2}$$

We only need to consider the assignments that assign

- i many times the bit patterns 0000 or 1111,
- j many times the bit pattern 0110, and
- k many times the bit pattern 1001,

since any other assignment contributes a factor 0. Let c_{ijk} be the sum over all such assignments of the products of evaluations of all signatures (including the signatures corresponding to the signature matrices X, P, and P^{-1}) in Ω except for the signature corresponding to the signature matrix in (9.2). Then

$$\mathrm{Holant}_{\Omega} = \sum_{i+j+k=n} (1+\sqrt{3})^j (1-\sqrt{3})^k c_{ijk}$$

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

$$\begin{aligned} \mathrm{Holant}_{\Omega_s} &= \sum_{i+j+k=n} ((4+2\sqrt{3})^j (4-2\sqrt{3})^k)^s c_{ijk} \\ &= \sum_{i+j+k=n} ((4+2\sqrt{3})^{j-k} 4^k)^s c_{ijk}. \end{aligned}$$

We argue that this Vandermonde system has full rank, which is to say that $(4+2\sqrt{3})^{j-k}4^k \neq (4+2\sqrt{3})^{j'-k'}4^{k'}$ unless $(j,k) = (j',k')$. If $(4+2\sqrt{3})^{j-k}4^k = (4+2\sqrt{3})^{j'-k'}4^{k'}$, then we have $(4+2\sqrt{3})^{j-k-(j'-k')}4^{k-k'} = 1$. Since any nonzero integer power of $4+2\sqrt{3}$ is not rational, we must have $j-k = j'-k'$. And in this case, $4^{k-k'} = 1$, and hence $k = k'$ and $j = j'$.

Therefore, we can solve for the unknown c_{ijk}'s and obtain the value of Holant_{Ω}. This completes the proof that Pl-Holant($\neq_2 | r$) is #P-hard by interpolation. Thus the problem Pl-Holant($\neq_2 \mid [0,1,0,0,0], [0,0,0,1,0], g$) is also #P-hard. □

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

9.6 Mixing $\mathscr{P}_2$ and $\mathscr{M}_4$ – Equalities and Matchgates in the Z Basis

Given a set $\mathcal{F}$ of symmetric signatures, by Theorem 9.17, Pl-Holant($\mathcal{F}$) is #P-hard unless every nondegenerate signature f of arity ≥ 3 in $\mathcal{F}$ is in $\mathscr{P}_1 \cup \mathscr{M}_2 \cup \mathscr{A}_3 \cup \mathscr{M}_3 \cup \mathscr{M}_4 \cup \mathscr{V}$. In this section we consider mixing signatures from $\mathscr{P}_2$ and $\mathscr{M}_4$. It is in this setting that the new planar tractable problems are discovered.

We first do a holographic transformation by Z. Then the problem becomes

$$\text{Pl-Holant}(\neq_2 \mid =_k, \text{EXACTONE}_d), \tag{9.3}$$

with various arities k and d. Recall that EXACTONE_d denotes the exact one function $[0, 1, 0, \ldots, 0]$ of arity d. These are the signatures for PERFECT MATCHING and they are the basic components of matchgates.

Surprisingly, contrary to the putative form of a complexity classification for planar counting problems, the complexity of Pl-Holant($\neq_2 \mid =_k, \text{EXACTONE}_d$) depends nontrivially on the values of d and k, and the problem is tractable for all large k. Furthermore, the tractable cases of this problem *cannot* be captured by a holographic reduction to Kasteleyn's algorithm, or any other known algorithm. Thus for planar problems the paradigm of holographic algorithms using matchgates (i.e., being $\mathscr{M}$-transformable) is *not* universal.

The problem (9.3) is tractable if $d \leq 2$, since in this case all signatures are in $\mathscr{P}$. Hence we assume $d \geq 3$. Let $\mathcal{EO} = \{\text{EXACTONE}_d \mid d \geq 3\}$.

9.6.1 Hardness When $k = 3$ or 4

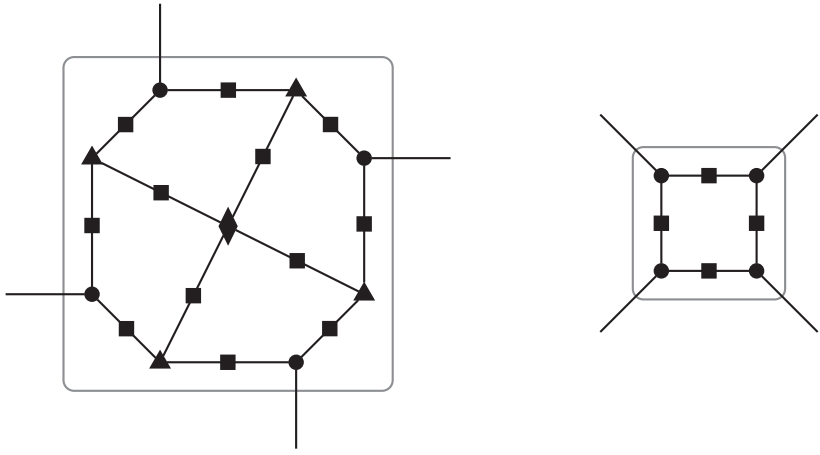
We begin with some hardness results.

Lemma 9.19. Pl-Holant($\neq_2 \mid =_3, \text{EXACTONE}_3$) *is* #P*-hard.*

Proof. By connecting two copies of $\text{EXACTONE}_3 = [0, 1, 0, 0]$ together via $(\neq_2)$, we have $\text{EXACTONE}_4 = [0, 1, 0, 0, 0]$ on the right. Consider the gadget in Figure 9.6a. We assign $(=_3)$ to the triangle vertices, $[0, 1, 0, 0]$ to the circle vertices, $(\neq_2)$ to the square vertices, and $[0, 1, 0, 0, 0]$ on the diamond vertex in the middle. Let f be the signature of this gadget.

We claim that this signature f is 0-1 valued, and its support is $\{0011, 0110, 1100, 1001\}$. To see this, for each $(=_3)$ at the four triangle vertices, we imagine all three adjacent DISEQUALITY signatures $(\neq_2)$ are absorbed to $(=_3)$, making it another $(=_3)$ signature. Thus the gadget in Figure 9.6a is functionally equivalent to another virtual gadget Figure 9.6a*, where all square

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press



(a) Cycle-like gadget used more than once (b) Gadget to realize $\hat{g}$

Figure 9.6. Two gadgets used in the proof of Lemma 9.19.

vertices are given simply the EQUALITY signature ($=_2$). Now in Figure 9.6a* the signature [0, 1, 0, 0, 0] in the middle forces exactly one of the four equality signatures (triangles) to take value 1 and the other three equality signatures to take value 0. The two [0, 1, 0, 0]'s (circle vertices) adjacent to two equality signatures both assigned 0 must have 1 going out, and the other two [0, 1, 0, 0]'s having one adjacent equality assigned 0 and another adjacent equality assigned 1 must have 0 going out.

Now we consider the gadget in Figure 9.6a again. This time we place [0, 1, 0, 0] on each triangle, ($=_3$) on each circle, f on the middle diamond, and again ($\neq_2$) on each square. In this gadget notice that each support vector of f makes two [0, 1, 0, 0]'s (triangles) that are cyclically adjacent on the outer cycle to become [0, 1, 0] and the other two [1, 0, 0]. It is easy to see that the resulting signature is 0-1 valued, and has support {0111, 1011, 1101, 1110}. Therefore it is the reversed EXACTONE$_4$ signature [0, 0, 0, 1, 0] (namely ALLBUTONE$_4$). The composite gadget as a construction for a signature on the RHS in Pl-Holant($\neq_2 \mid =_3$, EXACTONE$_3$) is illustrated in Figure 9.7, where each circle is assigned [0, 1, 0, 0], triangle ($=_3$), and square ($\neq_2$).

Finally, we build the gadget in Figure 9.6b. We place ($=_3$) on each circle and ($\neq_2$) on each square. It is easy to see that there are only two support vectors of the resulting signature, which are 0101 and 1010. Recall the definition (9.1) of the conditional crossover g. This gadget realizes exactly g.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

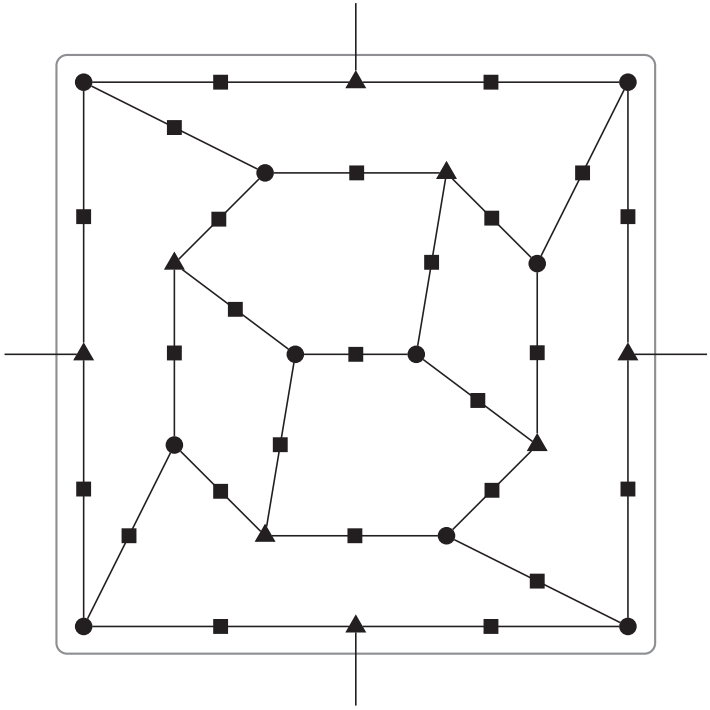


Figure 9.7. The whole gadget to realize [0, 0, 0, 1, 0].

By Lemma 9.18, Pl-Holant($\neq_2$ | [0, 1, 0, 0, 0], [0, 0, 0, 1, 0], g) is #P-hard. We have constructed [0, 1, 0, 0, 0], [0, 0, 0, 1, 0], and g on the right-hand side (RHS). Therefore Pl-Holant($\neq_2$ | $=_3$, [0, 1, 0, 0]) is #P-hard. □

For $k = 4$, we need the following lemma.

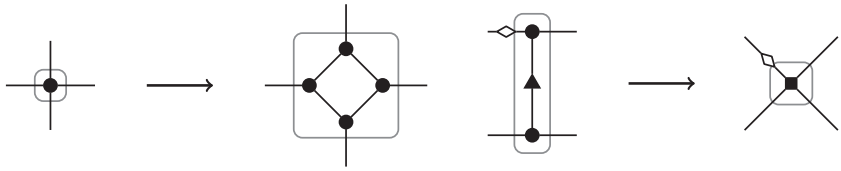
Lemma 9.20. *Let f be the arity* 4 *signature whose matrix is*

$$M_f = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Then Pl-Holant(f) *is* #P*-hard.*

Proof. Let $h = [2, 1, 1]$. We show that Pl-#CSP(h) $\leq_{\mathrm{T}}$ Pl-Holant(f) in two steps. In each step, we begin with a signature grid and end with a new signature grid such that the Holant values of both signature grids are the same.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press



(a) Step one: Degree 4 vertex example (b) Step two: Contract edges

Figure 9.8. A reduction from Pl-Holant($\mathcal{EQ} \mid h$) to Pl-Holant(g) for any binary signature h and a quaternary signature g that depends on h. The circle vertices are assigned $(=_4)$ or $(=_3)$ respectively, the triangle vertex is assigned h, and the square vertex is assigned the signature of the gadget to its left.

Then by Theorem 9.5, we are done. Or more explicitly, since Pl-#CSP(h) $\equiv$ Pl-Holant($\mathcal{EQ} \mid h$), by Theorem 9.4, we are done.

For step 1, let $G = (U, V, E)$ be an instance of Pl-Holant($\mathcal{EQ} \mid h$). Fix an embedding of G in the plane. This defines a cyclic ordering of the edges incident to each vertex. Consider a vertex $u \in U$ of degree k. It is assigned the signature $(=_k)$. We split u into k vertices, and connect the k edges originally incident to u to these k new vertices so that each vertex is incident to exactly one edge. We also connect these k new vertices in a cycle according to the cyclic ordering induced on them by their incident edges. Each of these vertices has degree 3, and we assign them $(=_3)$. Clearly the Holant value is unchanged. This completes step 1. An example of this step applied to a vertex of degree 4 is given in Figure 9.8a. The resulting graph has the following properties: (1) it is planar; (2) every vertex has either degree 2 (in V and assigned h) or degree 3 (newly created and assigned $(=_3)$); (3) each degree 2 vertex is connected to two degree 3 vertices; and (4) each degree 3 vertex is connected to one degree 2 vertex and two other degree 3 vertices.

Now step 2. For every $v \in V$, v has degree 2. We contract the two edges incident to v, or equivalently, we replace the two circle vertices and one triangle vertex boxed in Figure 9.8b with a single (square) vertex of degree 4. The resulting graph $G' = (V', E')$ is planar and 4-regular.

Next we determine the signature on each $v' \in V'$ after this contraction. Note that the two inputs to each original circle have to be the same. Therefore the support of this signature is $\{0000, 0110, 1001, 1111\}$, listed starting from the diamond and going counterclockwise. Moreover, due to the triangle assigned h in the middle, the weight on 0000 is 2, and every other weight is 1. Hence it is exactly the signature f, with the diamond in Figure 9.8b marking the first input bit. This finishes the proof. □

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

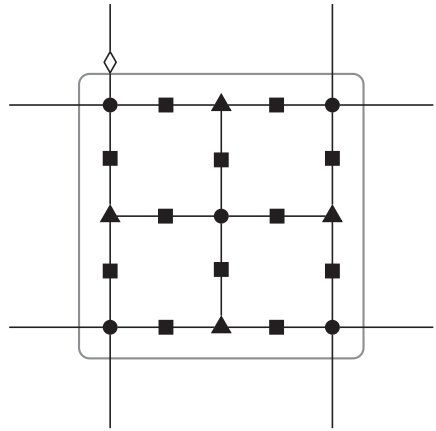


Figure 9.9. Grid-like gadget used in the proof of Lemma 9.21, whose support vectors are 00110011, 11001100, and 11111111. Each square is assigned a binary disequality ($\neq_2$), circle ($=_4$), and triangle [0, 1, 0, 0].

Remark: The two-step construction starting from a plane graph G produces what is known as the medial graph of G. See Definition 6.30 in Section 6.5, and Figure 6.9.

Our next goal is to prove the same result as Lemma 9.19, where $(=_3)$ is replaced by $(=_4)$. This proof is more challenging, because any gadget construction for the RHS of Pl-Holant($\neq_2 \mid =_4$, ExactOne$_3$) is very "fragile." It is not easy to produce gadgets with signatures that have nontrivial support with interesting nonzero values.

Lemma 9.21. Pl-Holant($\neq_2 \mid =_4$, ExactOne$_3$) *is* #P-*hard.*

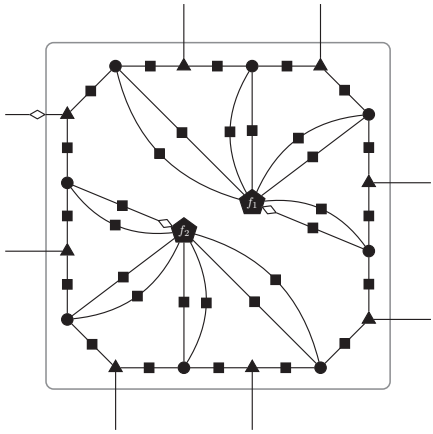
Proof. Consider the gadget in Figure 9.9. We assign binary disequality $(\neq_2)$ to the square vertices, $(=_4)$ to the circle vertices, and ExactOne$_3$ = [0, 1, 0, 0] to the triangle vertices. We show that the resulting signature is 0-1 valued, and its support set is

$$\{00110011, 11001100, 11111111\},$$

where each vector is the assignment ordered counterclockwise starting from the diamond point.

We call the equality signature $(=_4)$ in the middle the center vertex. There are two possible assignments at the center vertex. If it is assigned 0, then every adjacent perfect matching signature [0, 1, 0, 0] is matched to the half edge toward

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press



(a) Gadget with signature g. Each square is assigned a binary disequality ($\neq_2$), circle ($=_4$), triangle $[0, 1, 0, 0]$, and pentagon f.

f_1	f_2	g
00000000	00000000	11111111
00110011	00000000	01111000
11001100	00000000	11110000
00000000	00110011	10000111
00110011	00110011	00000000
11001100	00110011	-
00000000	11001100	00001111
00110011	11001100	-
11001100	11001100	00000000

(b) Support of g. Each vector is an assignment ordered counterclockwise from the diamond.

Figure 9.10. Another gadget used in the proof of Lemma 9.21 and a table listing the support of its signature.

the center vertex, and every equality ($=_4$) is forced to be 1. This gives the support vector 11111111.

The other possibility is that the center vertex is 1. In this case, we can remove the center vertex and its four adjacent ($\neq_2$), changing the four [0, 1, 0, 0] (triangles) into ($\neq_2$) = [0, 1, 0]. So the outer cycle effectively becomes a cycle of four equalities ($=_4$) (circle vertices) connected by ($\neq_2$), since a chain of three ($\neq_2$)'s is just a single ($\neq_2$). Hence in this case there are exactly two support vectors, which are 00110011 and 11001100.

Every pair of dangling edges at each corner vertex always takes the same value. We further connect each pair of these edges to a distinct copy of ($=_4$) via two copies of ($\neq_2$). This results in a gadget with a 0-1 valued signature f whose support is formed by taking complement bit patterns of the original support, namely, {11001100, 00110011, 00000000}.

Now consider the gadget in Figure 9.10a. We assign ($\neq_2$) to the square vertices, ($=_4$) to the circle vertices, $[0, 1, 0, 0]$ to the triangle vertices, and f to the pentagon vertex. In Figure 9.10a, two edges coming out of the pentagon vertex, which connect to the same circle vertex ($=_4$) on the outer cycle via ($\neq_2$), are from the same corner of the gadget in Figure 9.9 used to realize f, and thus must carry equal values 00 or 11. We now study the signature g of arity 8 defined by the gadget Figure 9.10a. The signature g is in the RHS of Pl-Holant($\neq_2 \mid =_4, \text{EXACTONE}_3$).

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

Notice that if a $(=_4)$ on the outer cycle is assigned 0, then the two adjacent perfect matchings must match half edges toward that $(=_4)$, and their outgoing edges must be 0. Furthermore, the two $(=_4)$ one more step away must be 1. A further observation is that any pair of consecutive $(=_4)$'s on the outer cycle cannot be both 0, and if a pair of consecutive $(=_4)$'s are both 1, then the [0, 1, 0, 0] in the middle must have a 1 going out. In Figure 9.10a, we call f_1 the pentagon connecting to four equalities $(=_4)$ on the upper right part, and call the other one f_2. We further order the external wires of f_1, f_2, and g counterclockwise, each starting from the edge marked with a diamond. With this notation and these observations, we get Table 9.10b listing the support of g. The support of g is

$$\{11111111, 01111000, 11110000, 10000111,$$
$$00000000, 00001111, 00000000\},$$

and 00000000 has multiplicity 2.

Next we use domain pairing. In Pl-Holant$(\neq_2 \mid =_4, \text{EXACTONE}_3, g)$, each $(=_4)$ is connected to four $(\neq_2)$'s on the left-hand side (LHS). We can move $(=_4)$ to the LHS by contracting four $(\neq_2)$'s into it. So we get a reduction Pl-Holant$(=_4 \mid g) \leq_{\text{T}}$ Pl-Holant$(\neq_2 \mid =_4, \text{EXACTONE}_3, g)$. We apply domain pairing on the problem Pl-Holant$(=_4 \mid g)$. Specifically, we can use $(=_4)$ as $(=_2)$ by pairing a pair of edges together, and also pair adjacent two outputs of g counterclockwise, starting from the diamond point. Each pair of output wires of g is connected to a pair of wires from $(=_4)$ on the left hand side. Note that $(=_4)$ enforces that each pair of edges always takes the same value. We reinterpret 00 or 11 as 0 or 1 in the Boolean domain. In this way, we can treat g as an arity 4 signature g' in the Boolean domain. So the reduction is

$$\text{Pl-Holant}(=_2 \mid g') \leq_{\text{T}} \text{Pl-Holant}(=_4 \mid g).$$

We get the expression of g' next. The two support bit strings 01111000 and 10000111 of g are eliminated as they do not agree on adjacent paired outputs. So in the paired (Boolean) domain, the support of g' becomes $\{1111, 1100, 0011, 0000\}$, where 0000 has multiplicity 2. We further rotate g' as a Boolean domain signature such that the support is $\{1111, 0110, 1001, 0000\}$. Now it is easy to see that the matrix of g', an arity 4 signature in the Boolean domain, is

$$M_{g'} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

By Lemma 9.20 Pl-Holant(g') is #P-hard. Hence Pl-Holant($\neq_2 \mid =_4$, EXACTONE$_3$) is also #P-hard. □

To extend Lemma 9.19 and Lemma 9.21 to general EXACTONE$_d$ functions, we show that we can always realize the pinning functions $\Delta_0 = [1, 0]$ and $\Delta_1 = [0, 1]$ in this setting.

Lemma 9.22. *For any integer $k \geq 3$ and $d \geq 3$,*

$$\text{Pl-Holant}(\neq_2 \mid =_k, \text{EXACTONE}_d, \Delta_0, \Delta_1) \leq_{\text{T}} \text{Pl-Holant}(\neq_2 \mid =_k, \text{EXACTONE}_d).$$

Proof. Given an instance Ω of Pl-Holant($\neq_2 \mid =_k$, EXACTONE$_d$, Δ_0, Δ_1) with underlying planar bipartite graph G, if any Δ_0 from the RHS is used in Ω, then it is applied to a ($\neq_2$) of the LHS, and can be combined with ($\neq_2$) to become a Δ_1 on the LHS. Then it is applied to a signature of the RHS: If it's Δ_1 it becomes a constant 1 and they can be removed; if it's Δ_0 the Holant value is 0; if it's EXACTONE$_d$ it becomes $\Delta_0^{\otimes(d-1)}$; if it's ($=_k$), it becomes $\Delta_1^{\otimes(k-1)}$. These newly created pinning signatures on the RHS can be applied recursively.

By a similar analysis, the only nontrivial occurrences of Δ_1's in Ω are those attached to EXACTONE$_d$ via ($\neq_2$). We may therefore assume that the only appearances among Δ_0's and Δ_1's are possibly some Δ_1's applied to EXACTONE$_d$ via ($\neq_2$).

We can construct ($=_{\ell k}$) for any integer $\ell \geq 1$, by ($\neq_2$) on the left and ($=_k$) on the right. In fact if we connect two copies of ($=_k$) via ($\neq_2$) on one wire as a planar gadget we get a signature of arity $2k-2$ with $k-1$ consecutive external wires labeled "+" and the others labeled "−," where wires of the same label take equal values, but oppositely labeled wires ("+−") take opposite values. As $k \geq 3$, we can take two wires of the $k-1$ wires labeled "−" and attach to two copies of ($=_k$) via two ($\neq_2$). This creates a signature of arity $4k-6$ with $3(k-1)$ consecutive wires labeled "+" and the other $k-3$ wires labeled "−". Finally connect $k-3$ pairs of adjacent ("+−") labeled wires by ($\neq_2$) recursively. This creates a planar gadget with an equality signature of arity $2k$. This can be extended to any ($=_{\ell k}$) by applying the same process on any consecutive k wires.

Next we construct $\Delta_1^{\otimes r}$ in the RHS, for some integer $r \geq 1$. We get $\Delta_0^{\otimes(d-2)}$ by a self-loop of EXACTONE$_d$ via ($\neq_2$), ignoring the factor 2. We pick an integer ℓ large enough so that $d-2 < \ell k$. Then we connect $\Delta_0^{\otimes(d-2)}$ to ($=_{\ell k}$) via ($\neq_2$) to get $\Delta_1^{\otimes(\ell k-d+2)}$. Then we take $r = \ell k - d + 2$.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

One more construction we will use is $\text{EXACTONE}_{2+\ell(d-2)}$ for any integer $\ell \geq 1$. This is realizable by connecting ℓ many copies of EXACTONE_d sequentially via $(\neq_2)$.

Consider the dual graph G^* of G. Take a spanning tree T of G^*, with the external face as the root. In each face F, let c_F be the number of Δ_1's in the face. We start from the leaves to recursively move all the pinnings of Δ_1 to the external face. Suppose the face F is a current leaf of T. If $c_F = 0$ then we just remove the leaf from T and recurse on another leaf. Otherwise we temporarily remove all Δ_1's in F. Let s be the smallest integer such that $sr \geq c_F$. We replace the $(\neq_2)$ edge bordering between F and its parent F' by a path with three signatures: $(\neq_2)$, $\text{EXACTONE}_{2+\ell(d-2)}$ and $(\neq_2)$, where ℓ is large such that $\ell(d-2) \geq sr - c_F$. The bipartiteness is preserved with the two $(\neq_2)$'s from the LHS and the $\text{EXACTONE}_{2+\ell(d-2)}$ from the RHS. Two edges of $\text{EXACTONE}_{2+\ell(d-2)}$ are connected to the two adjacent copies of $(\neq_2)$ on the path. Of the other $\ell(d-2)$ edges we will put $sr - c_F$ many dangling edges in F, and the remaining $\ell(d-2) - (sr - c_F)$ dangling edges in F'. Hence there are sr dangling edges in F, including those c_F many dangling edges that were connected to Δ_1's via $(\neq_2)$ before we removed them. We put s copies of $\Delta_1^{\otimes r}$ inside the face F to pin all of them in a planar way. Since this $\Delta_1^{\otimes r}$ is from the RHS, it is connected via $(\neq_2)$, and so the value being pinned are all 0, including its effect on $\text{EXACTONE}_{2+\ell(d-2)}$. In the end all $\ell(d-2)$ dangling edges of $\text{EXACTONE}_{2+\ell(d-2)}$ in both F and F' will be pinned to 0, thus its effect on the path $(\neq_2)$, $\text{EXACTONE}_{2+\ell(d-2)}$ and $(\neq_2)$, is equivalent to a path of three $(\neq_2)$'s, thus equivalent to the original single $(\neq_2)$. To complete the description of the construction at face F, we add $\ell(d-2) - (sr - c_F)$ to $c_{F'}$, and they are all connected to Δ_1 via $(\neq_2)$. Remove the leaf F from T, and recurse.

After the process, all Δ_1's are in the external face of G. Suppose the number is p. We may use $\Delta_1^{\otimes r}$ to reduce this number to $0 \leq p < r$. We put r disjoint copies of G together to form a planar signature grid. Use p copies of $\Delta_1^{\otimes r}$ in the outer face in a planar way, making available a total of pr many Δ_1's. This is now an instance of Pl-Holant$(\neq_2 \mid =_k, \text{EXACTONE}_d)$ and the Holant value is the rth power of that of Ω. Since the Holant value of Ω is a nonnegative integer, we can take the rth root and finish the reduction. □

Once we have the pinning functions, it is easy to construct EXACTONE_3 from EXACTONE_d. Therefore combining Lemmas 9.22 with 9.19 and 9.21 we get the following.

Corollary 9.23. *If* $d \geq 3$ *and* $k \in \{3, 4\}$, *then* Pl-Holant$(\neq_2 \mid =_k, \text{EXACTONE}_d)$ *is* #P-*hard.*

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

The implication of Corollary 9.23 for counting planar hypergraph perfect matchings will be discussed in Theorem 9.28.

9.6.2 Tractability When $k \geq 5$

On the other hand, if the arity k of the equality signature is ≥ 5, then the problem Pl-Holant($\neq_2 \mid =_k, \mathcal{EO}$) is tractable. In this subsection we will first prove that the problem is tractable for $k \geq 6$. After that we will return to $(=_5)$.

In fact we consider the problem Pl-Holant($\neq_2 \mid =_k, \mathcal{EO}, \neq_2, \Delta_0, \Delta_1$). Note that $(\neq_2) = [0, 1, 0] = \text{EXACTONE}_2$ and $\Delta_1 = [0, 1] = \text{EXACTONE}_1$. Given an instance of the problem with underlying graph G, we can do some preprocessing as follows. Any occurrence of $(\neq_2)$ on the RHS is connected to two copies of $(\neq_2)$ on the LHS. We can replace the three $(\neq_2)$'s by a single $(\neq_2)$ on the LHS, thus eliminating the $(\neq_2)$ on the RHS. Any pinning signature Δ_0 or Δ_1 can be applied and possibly propagated. Any Δ_0 on the RHS is first transformed to Δ_1 by $(\neq_2)$ on the LHS and then applied to a signature on the RHS. If it is applied to Δ_1 then it produces a constant 1, and can be eliminated. If it is applied to Δ_0 then the Holant value is 0 and we are done. If it is applied to $(=_k)$ then this $(=_k)$ is changed to $\Delta_1^{\otimes(k-1)}$, which can be recursively applied. If it is applied to EXACTONE_d (for some $d \geq 3$) then this EXACTONE_d becomes $\Delta_0^{\otimes(d-1)}$, which can be recursively applied. The case with any occurrence of Δ_1 on the RHS is similar. It is first transformed to Δ_0 on the LHS. Then if it is applied to Δ_0 or Δ_1 on the RHS, the case is simple. Applying to $(=_k)$ produces $\Delta_0^{\otimes(k-1)}$, which can be recursively applied. Applying Δ_0 to EXACTONE_d (for some $d \geq 3$) produces EXACTONE_{d-1}, which is either in $\mathcal{EO}$ (if $d > 3$) or $(\neq_2)$ which can be eliminated.

Hence we may consider the instance is from Pl-Holant($\neq_2 \mid =_k, \mathcal{EO}$).

Any self-loop on $(=_k)$ by $(\neq_2)$ gives the Holant value 0. Any self-loop on an EXACTONE_d by $(\neq_2)$ changes it to a $\Delta_0^{\otimes(d-2)}$ with factor 2. These pinning signatures can be applied recursively. Whenever an EXACTONE_d and another EXACTONE_ℓ are connected by a $(\neq_2)$, we replace it by a single $\text{EXACTONE}_{d+\ell-2}$, shrinking the edge between (and remove the connecting $(\neq_2)$). (If there are $m > 1$ edges connecting them by $(\neq_2)$, then $m - 1$ self-loops result, which either lead to pinning signatures or Holant value 0 if $m > 2$.) On the other hand, consider a connected component made of $(=_k)$ and $(\neq_2)$. We call such a component an E_k-block. An E_k-block has either exactly two or zero support vectors. This depends on whether or not there exists a contradiction, which is formed by an odd cycle of $(=_k)$ connected by $(\neq_2)$. We say an E_k-block is trivial if it has no support. An E_k-block is nontrivial iff the graph with $(=_k)$ as vertices and $(\neq_2)$ as edges is bipartite. This is easy to check in

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

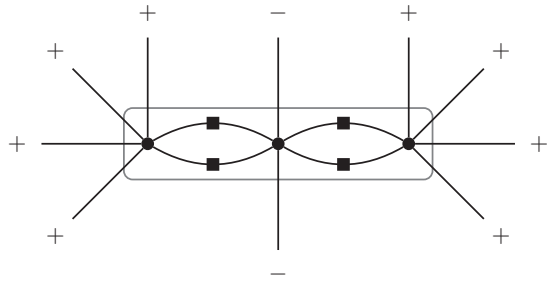


Figure 9.11. Example E_6-block. Circle vertices are assigned $(=_6)$ and square vertices are assigned $(\neq_2)$.

polynomial time, and if so we can find the two support assignments of a nontrivial E_k-block that are complements of each other. We mark dangling edges of a nontrivial E_k-block by "+" or "−" signs. Dangling edges marked with the same sign take the same value on both support assignments while dangling edges marked with different signs take opposite values on both support assignments. Let n_+ and n_- denote the number of dangling edges marked by $+$ and $-$ respectively. Then it is easy to show by induction that

$$n_+ \equiv n_- \pmod{k} \tag{9.4}$$

An example of an E_6-block is illustrated in Figure 9.11, with 8 "+" signs and 2 "−" signs.

Exercise: Prove (9.4).

Exercise: Prove that for any n_+ and $n_- \geq 0$ such that $n_+ + n_- = n$ and $n_+ \equiv n_- \mod k$, if $b_1 b_2 \dots b_n \in \{+, -\}^n$ is a string with n_+ many $+$'s and n_- many $-$'s, then there is a linear sized E_k-block such that all its dangling edges are on the outer face, and has labels $b_1 b_2 \dots b_n$ in cyclic order. (Since the string is arbitrary, the statements for clockwise cyclic order and counterclockwise cyclic order are equivalent.)

After contracting all edges between ExactOne$_d$'s and forming E_k-block's we obtain a bipartite graph G' connected between ExactOne$_d$'s and E_k-block's by edges labeled by $(=_2)$. We claim that G' is planar. Note that even though the given graph G is planar, and contracting edges is a planar operation, the planarity of the (abstract) graph G' where E_k-block's and (contracted) ExactOne$_d$'s are vertices requires a justification. In particular, an E_k-block B in G is a connected planar graph that may have dangling edges in several distinct faces, and the topological structure between the E_k-block's and ExactOne$_d$'s in G can be arbitrarily nested and complicated. Each E_k-block has one outer

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

face and possibly several internal faces, each of which may have some dangling edges connected to EXACTONE$_d$'s. To show planarity of G', we order the dangling edges on the outer face F of B counterclockwise (if any), and order the dangling edges on each internal face (if any) $F', F'', \ldots$, clockwise, and then concatenate their labels into a single string σ of $+$'s and $-$'s. If an internal face F' with dangling edges exists, we will simply move the part within F' connected to these dangling edges to the outer face. Then we replace this E_k-block by a construction (see Exercise above) that has all dangling edges ordered cyclically according to σ and placed on the outer face. After the replacement we reconnect all dangling edges, now all on the outer face. This way we maintain planarity, and the replacement E_k-block can be contracted to a single point. After replacing every E_k-block by an equivalent E_k-block the planarity of the resulting graph shows that G' is also planar.

Now a key observation is that if a planar (bipartite) graph has moderately high degrees, then it cannot be a simple graph, i.e., it must have parallel edges.

Lemma 9.24. *Let $G = (L \cup R, E)$ be a planar bipartite graph with parts L and R. If every vertex in L has degree ≥ 6 and every vertex in R has degree ≥ 3, then G is not simple.*

Proof. Suppose G is simple. Let V, E, and F be the total number of vertices, edges, and faces, respectively. Let v_i are the number of vertices of degree i in L, where $i \geq 6$, and let u_j be the number of vertices of degree j in R, where $j \geq 3$. Since G is simple and bipartite, each face has at least four edges. Thus,

$$2\mathsf{E} \geq 4\mathsf{F}. \tag{9.5}$$

Furthermore, it is easy to see that

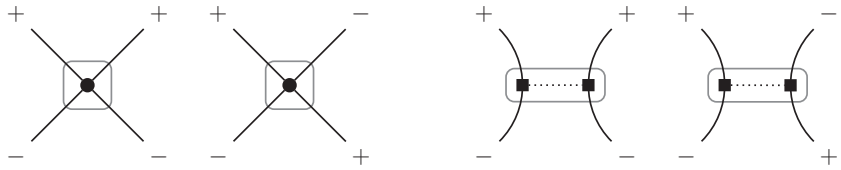
$$\mathsf{V} = \sum_{i\geq 6} v_i + \sum_{j\geq 3} u_j \qquad \text{and} \qquad \mathsf{E} = \sum_{i\geq 6} i v_i = \sum_{j\geq 3} j u_j. \tag{9.6}$$

Then starting from Euler's characteristic formula for planar graphs, we have

$$\begin{aligned} 2 &= \mathsf{V} - \mathsf{E} + \mathsf{F} \\ &\leq \mathsf{V} - \frac{\mathsf{E}}{2} && \text{(By (9.5))} \\ &= \sum_{i\geq 6} v_i + \sum_{j\geq 3} u_j - \frac{1}{6}\sum_{i\geq 6} i v_i - \frac{1}{3}\sum_{j\geq 3} j u_j && \text{(By (9.6))} \\ &= \sum_{i\geq 6} \frac{6-i}{6} v_i + \sum_{j\geq 3} \frac{3-j}{3} u_j \leq 0, \end{aligned}$$

a contradiction. □

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press



(a) Two different arity 4 E_k-blocks. (b) Replace them by parallel $(\neq_2)$'s.

Figure 9.12. Arity 4 E_k-blocks.

Lemma 9.24 does not give us tractability for the case of $k \geq 6$ yet. The reason is that, after preprocessing, given an instance of Pl-Holant$(\neq_2 \mid =_k, \mathcal{EO})$, we form E_k-blocks to make the graph bipartite. It is possible to have E_k-blocks of arity < 6. If so Lemma 9.24 does not apply. However, for $k \geq 6$ and a nontrivial E_k-block of arity n where $n < 6$, by (9.4) and the fact that $0 \leq n_+, n_- \leq n < k$, we see that $n_+ = n_-$, and $n = n_+ + n_-$ must be even. Moreover, if $n = 2$, then this means that the E_k-block is just $(\neq_2)$, in which case we can replace it by a single $(\neq_2)$ connecting signatures from $\mathcal{EO}$. If this $(\neq_2)$ connects two different signatures from $\mathcal{EO}$ then it produces a new ExactOne signature in $\mathcal{EO}$. If it connects the same signature from $\mathcal{EO}$, this produces pinning Δ_0's and can be handled as before. So the only problematic case is when $n = 4$. We identify two possibilities of such E_k-blocks up to a rotation in Figure 9.12a.

We may replace any E_k-block of arity 4 by an equivalent E_k-block of arity 4 where all dangling edges are on the outer face, as depicted in Figure 9.12a. Functionally an E_k-block of arity 4 can be viewed as a pair of $(\neq_2)$ in parallel, but there is a correlation between them, namely their support vectors are paired up in a unique way. If we replace the contracted E_k-block of arity 4 by two parallel edges with $(\neq_2)$ as indicated in Fig. 9.12b, we maintain planarity (but we cannot claim it maintains functional equivalence due to the nontrivial correlation.)

We will show in the following lemma how to replace E_k-block of arity 4 by two parallel edges with $(\neq_2)$ while keeping track of the Holant value.

Lemma 9.25. *For any integer $k \geq 6$,* Pl-Holant$(\neq_2 \mid =_k, \mathcal{EO}, \neq_2, \Delta_0, \Delta_1)$ *is tractable.*

Proof. Let Ω be an instance of Pl-Holant$(\neq_2 \mid =_k, \mathcal{EO}, \neq_2, \Delta_0, \Delta_1)$. Without loss of generality, we assume that Ω is connected. We already discussed the preprocessing so that we may assume Ω is an instance of Pl-Holant$(\neq_2 \mid =_k, \mathcal{EO})$.

We call an edge pinned if it has the same value in all satisfying assignments, *if there is any*. When an edge is pinned to a known value, we can get a smaller

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

instance of the problem Pl-Holant($\neq_2 \mid =_k, \mathcal{EO}$) without changing the number of satisfying assignments. In our algorithm we may also find a contradiction and simply return 0. If e is a pinned edge, then it is adjacent to another edge e' via ($\neq_2$) on the LHS, and both e and e' are pinned. We remove e, e', and ($\neq_2$), and perform the following on e (and on e' as well). If the other endpoint of e in the RHS is ($=_k$), then all edges incident to this ($=_k$) are pinned to the same value which we can recursively apply. If the other endpoint of e is $\text{EXACTONE}_d \in \mathcal{EO}$, then we replace this signature by EXACTONE_{d-1} if the pinned value is 0; or if the pinned value is 1 then the remaining $d-1$ edges of this EXACTONE_d are pinned to 0 which we recursively apply. Notice that we may create an EXACTONE_2 (i.e., ($\neq_2$)) on the RHS when we pin 0 on EXACTONE_3. Such ($\neq_2$)'s are replaced as described before. It is easy to see that all these procedures do not change the number of satisfying assignments, and work in polynomial time.

Claim: There always exists an edge in Ω that is pinned, unless Ω does not contain ($=_k$), or does not contain EXACTONE_d functions (for some $d \geq 3$), or there is a contradiction. Furthermore, if there are ($=_k$) and EXACTONE_d functions (for some $d \geq 3$), in polynomial time we can find a pinned edge with a known value, or return that there is a contradiction. (If there is a contradiction in Ω, we may still return a purported pinned edge with a known value, which we can apply and simplify Ω. The contradiction will eventually be found.)

Note that if Ω does not contain ($=_k$), or does not contain EXACTONE_d functions (for some $d \geq 3$), then the problem is tractable, since Ω is an instance of $\mathscr{M}$, or an instance of $\mathscr{P}$, respectively. The lemma follows from the claim, for we either recurse on a smaller instance or have a tractable instance.

Suppose Ω is an instance where at least one ($=_k$) and at least one $\text{EXACTONE}_d \in \mathcal{EO}$ appear. Let's examine how the signatures can be connected by ($\neq_2$). Clearly we may assume there is no self-loop on any ($=_k$); otherwise we return the Holant value 0. If a signature $\text{EXACTONE}_d \in \mathcal{EO}$ is connected to itself by a self-loop through a ($\neq_2$), then there are two choices for the assignment on this pair of edges through the ($\neq_2$), but the remaining $d-2 \geq 1$ edges are pinned to 0. We can keep track of the factor 2 and have found a pinned edge with a known value. Thus we may assume there are no self-loops via ($\neq_2$) on EXACTONE signatures.

Next we consider the case that two separate signatures EXACTONE_d and EXACTONE_ℓ from $\mathcal{EO}$ are connected by some number of ($\neq_2$)'s. Depending on the number of connecting edges, there are three cases:

1. The connection is by a single ($\neq_2$). We contract the connecting edge, maintaining planarity, and replace these three signatures by an $\text{EXACTONE}_{d+\ell-2}$

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

to get a new instance Ω'. If an edge is pinned in Ω' then it is also pinned in Ω to the same value. We continue with Ω'.
2. The connection is by two $(\neq_2)$'s. There are two choices for the assignment on these two pairs of edges through $(\neq_2)$, but the remaining $d + \ell - 4 \geq 2$ edges are pinned to 0.
3. The connection is by at least three $(\neq_2)$'s. The three $(\neq_2)$'s cannot be all satisfied, so there is no satisfying assignment, a contradiction. We return the value 0.

Hence, we may assume there is no connection via any number of $(\neq_2)$'s among ExactOne signatures in $\mathcal{EO}$.

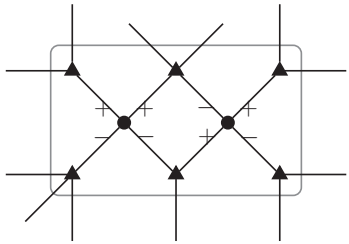
Now we consider connections between various $(=_k)$'s by $(\neq_2)$. Define an E_k-block as a connected component composed of $(=_k)$ and $(\neq_2)$. All dangling edges of each E_k-block are marked with $+$ or $-$ and this can be found by testing bipartiteness of an E_k-block where we treat $(=_k)$'s as vertices and $(\neq_2)$'s as edges. If any E_k-block is not bipartite, we have found a trivial E_k-block, and we return 0. Thus we may assume all E_k-blocks are nontrivial. We replace all E_k-blocks by equivalent E_k-blocks with all dangling edges on the outer face, and maintain planarity. If there is a nontrivial E_k-block of arity 2, by (9.4) its signature is $(\neq_2)$. We replace it with an edge labeled by $(\neq_2)$ to form an instance Ω', maintaining planarity, such that any pinned edge in Ω' corresponds to a pinned edge in Ω. So we may assume that the arity of any E_k-block is ≥ 4. Since $k \geq 6$, the only possible E_k-blocks of arity 4 functionally are those in Figure 9.12a up to a rotation. Since there is at least one ExactOne_d signature with $d \geq 3$, forming E_k-blocks does not consume all of Ω.

After these steps we may consider Ω a planar bipartite graph, with one side consisting of E_k-blocks of arity ≥ 4 and the other side ExactOne signatures in $\mathcal{EO}$. And they are now connected by edges labeled by $(=_2)$.

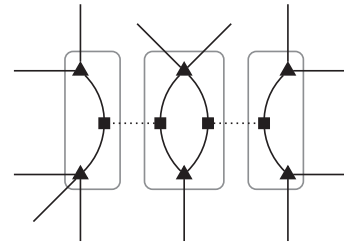
Suppose there are parallel edges between an E_k-block and an ExactOne_d signature. We show that this always leads to some pinned edges. If two parallel edges are marked by the same sign in the E_k-block, then they must be pinned to 0. If they are marked by different signs, then the remaining $d - 2 \geq 1$ edges of the ExactOne_d signature must be pinned to 0. Therefore, we may assume there are no parallel edges between any E_k-block and any ExactOne signature.

The next thing we do is to consider E_k-blocks of arity 4 with ExactOne signatures together. Call a connected component consisting of E_k-blocks of arity 4 and ExactOne signatures an *EO-Eq*-4-block. Figure 9.13a illustrates an example. As remarked earlier, the two possibilities of E_k-blocks of arity 4 can be viewed as two parallel $(\neq_2)$'s but with some correlation between them.

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press



(a) An *EO-Eq*-4-block. Triangles are assigned ExactOne signatures and circles are E_k-blocks of arity 4.



(b) Break the *EO-Eq*-4-block into three components. Squares are assigned $(\neq_2)$. The component in the middle contains a cycle, and hence is degenerate. The other two are equivalent to ExactOne signatures.

Figure 9.13. *EO-Eq*-4-blocks.

This is illustrated in Figure 9.12b. Note that the two dotted lines in Figure 9.12b represent different correlations.

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

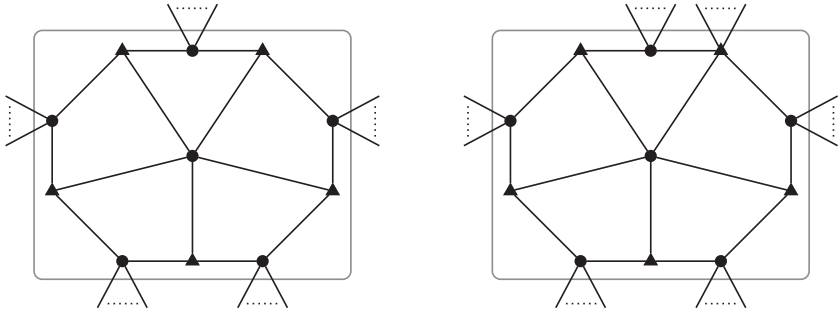
The *surprising* move of this proof is that we shall do so anyway!

Suppose we ignore the correlation for the time being and replace every arity 4 E_k-block by two parallel $(\neq_2)$'s as in Figure 9.12b. This replacement produces a *planar* signature grid Ω_1. Every edge in Ω_1 corresponds to a unique edge in Ω. The set of satisfying assignments of Ω_1 is a superset of that of Ω. Moreover, if there is an edge pinned in Ω_1 to a known value, the corresponding edge is also pinned in Ω to the same value. Once we find that in Ω_1 we revert back to work in Ω and apply the pinning to the pinned edge.

All that remains to be shown is that pinning always happens in Ω_1. Each *EO-Eq*-4-block splits into some number of connected components in Ω_1. If any component contains a cycle, it must alternate between $(\neq_2)$, which are the newly created ones from the E_k-blocks of arity 4, and ExactOne_d signatures for $d \geq 3$. Note that ExactOne_d signatures on the cycle all have arity ≥ 3. The cycle has even length, and there are exactly two satisfying assignments, which assign exactly one 0 and one 1 to the two cycle edges incident to each ExactOne_d signature. Hence any edge not on the cycle but incident to some vertex (with a signature from $\mathcal{EO}$) belonging to the cycle is pinned to 0. Moreover such edges must exist by the arity of ExactOne_d signatures. This produces pinned edges.

Hence we may assume there are no cycles in these components. If we consider ExactOne functions as vertices, and $(\neq_2)$'s as edges, each component

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press



(a) Type 1 (b) Type 2

Figure 9.14. Two types of wheels. Each circle is an E_5-block and triangle an EXACTONE signature.

forms a tree. Suppose there are $n \geq 2$ vertices in such a tree. As discussed in item 1 above, the whole tree is an EXACTONE_t function for some arity t. Since each vertex in the tree has degree ≥ 3, $t \geq 3n - 2(n-1) = n + 2 \geq 4$. We replace these components by EXACTONE_t's. Since each connected component is a tree, we can contract it to a point, labeled by some $\text{EXACTONE}_t \in \mathcal{EO}$, and maintain planarity. This defines a planar signature grid Ω_2.

Thus, each connected component in the graph underlying Ω_2 is a planar bipartite graph with E_k-blocks of arity ≥ 6 on one side and EXACTONE_d signatures of arity ≥ 3 on the other. By Lemma 9.24, no component is simple, which means that there are parallel edges between some E_k-block and some EXACTONE_d signature. As discussed earlier, there must exist some pinned edge, and we can find a pinned edge with a known value in polynomial time. This gives a pinned edge with a known value in Ω_1, and then in Ω. This finishes the proof. □

Now we go on to consider Pl-Holant($\neq_2 \mid =_5, \mathcal{EO}$). Unlike the situation in Lemma 9.24, a planar (5, 3)-regular bipartite graph *can* be simple. However, we show that any *simple* bipartite graph where the degrees are ≥ 5 from the left and ≥ 3 from the right must have a special induced subgraph. We call this structure a "wheel," which is pictured in Figure 9.14. There is a vertex v of degree 5 in the middle, and all faces adjacent to this vertex are 4-gons (i.e., quadrilaterals). Moreover, at least four neighbors of v have degree 3. Depending on the degree of the fifth neighbor (whether it is 3 or not), we have two types of wheel, which are pictured in Figures 9.14a and 9.14b.

Lemma 9.26. *Let $G = (L \cup R, E)$ be a planar bipartite graph with parts L and R. Every vertex in L has degree ≥ 5 and every vertex in R has degree ≥ 3.*

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

If G is simple, then there exists one of the two wheel structures in Figure 9.14 in G as an induced subgraph.

Proof. Let $V = L \cup R$ be the set of vertices and let F be the set of faces. We assign a *score* s_v to each vertex $v \in V$. We will define s_v so that $\sum_{v \in V} s_v = |V| - |E| + |F| = 2 > 0$. The score s_v will be the sum of values assigned to v locally. These values include a base score for the vertex v itself, and values from the $\deg(v)$ edges incident to v, and from the $\deg(v)$ faces containing v. The base score is $+1$ for each vertex, which accounts for $|V|$. For each k-gon face, we assign $\frac{1}{k}$ to each of its vertex. This accounts for $|F|$. As G is a simple bipartite graph, $k \geq 4$ and a score coming from a face to a vertex is at most $\frac{1}{4}$.

For $-|E|$, we separate two cases. For any edge if one of the two endpoints has degree 3, we give the degree 3 vertex a score of $-\frac{7}{12}$, and the other one $-\frac{5}{12}$. This is well defined because all degree 3 vertices are in R. If the endpoints are not of degree 3, we give each endpoint $-\frac{1}{2}$. This accounts for $-|E|$.

Claim 1: For any $v \in V$, $s_v \leq 0$ unless $v \in L$ and has degree 5.

To prove this claim, suppose $v \in L$ and has degree $d \geq 6$, then

$$s_v \leq 1 + \frac{d}{4} - \frac{5}{12}d = 1 - \frac{d}{6} \leq 0.$$

Now suppose $v \in R$ and v has degree $d \geq 4$. Then every edge adjacent to v gives a score $-\frac{1}{2}$. Hence,

$$s_v \leq 1 + \frac{d}{4} - \frac{1}{2}d = 1 - \frac{d}{4} \leq 0.$$

The remaining case is that $v \in R$ and v has degree 3. Then,

$$s_v \leq 1 + \frac{d}{4} - \frac{7}{12}d = 1 - \frac{d}{3} \leq 0.$$

Claim 1 is proved.

Since the total score is positive, there must exist $v \in L$, such that $\deg(v) = 5$ and $s_v > 0$.

Claim 2: There exists $v \in L$, such that $\deg(v) = 5$, $s_v > 0$, and all adjacent faces are 4-gons.

Suppose otherwise. Then any $v \in L$ with $\deg(v) = 5$ and $s_v > 0$ is adjacent to at least one k-gon for some $k \geq 6$. In this case,

$$s_v \leq 1 + \frac{1}{4} \cdot 4 + \frac{1}{6} - \frac{5}{12} \cdot 5 = \frac{1}{12}.$$

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

Moreover, if v is adjacent to more than one k-gon with $k \geq 6$, then

$$s_v \leq 1 + \frac{1}{4} \cdot 3 + \frac{1}{6} \cdot 2 - \frac{5}{12} \cdot 5 = 0,$$

contrary to the assumption that $s_v > 0$. Hence v is adjacent to exactly one k-gon with $k \geq 6$. Call this face F_v.

Along the boundary of F_v, v has two neighbors in R. We match each vertex v that has a positive score to the vertex that is on the boundary of F_v and next to v in clockwise order. By bipartiteness, v is matched to a vertex in R. We do this matching in all faces containing at least one positively scored vertex. It is possible that more than one such v are matched to the same $u \in R$. Suppose a vertex $u \in R$ is matched to from ℓ different such vertices of positive score. This means that u is adjacent to at least ℓ many k-gons with $k \geq 6$. The total number of faces that contain u is exactly $\deg(u)$ and therefore $\ell \leq \deg(u)$. Thus, if u has degree 3 then u has score

$$s_u \leq 1 + \frac{1}{4} \cdot (3 - \ell) + \frac{1}{6} \cdot \ell - \frac{7}{12} \cdot 3 = -\frac{\ell}{12}.$$

If u has degree $d \geq 4$ then u has score

$$s_u \leq 1 + \frac{1}{4} \cdot (d - \ell) + \frac{1}{6} \cdot \ell - \frac{1}{2} \cdot d \leq -\frac{\ell}{12}.$$

Hence in any case, we have $s_u \leq -\frac{\ell}{12}$. It implies that the sum of s_u and the scores of all positively scored vertices matched to u is ≤ 0. However each positively scored vertex is matched to a vertex in R. Hence the sum of all scores cannot be positive. This is a contradiction.

Therefore there exists $v \in L$ such that $s_v > 0$, $\deg(v) = 5$, and all adjacent faces are 4-gons. This proves claim 2.

We further note that at most one neighbor of v can have degree ≥ 4, for otherwise,

$$s_v \leq 1 + \frac{5}{4} - \frac{1}{2} \cdot 2 - \frac{5}{12} \cdot 3 = 0.$$

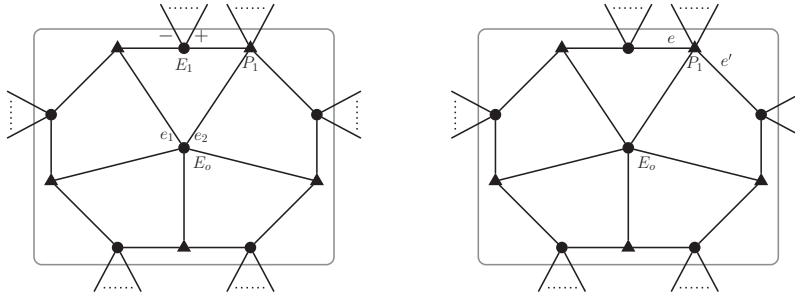
If all five neighbors of v have degree 3, that is a wheel of type 1 as in Figure 9.14a. If exactly one neighbor of v has degree ≥ 4, and the other four neighbors all have degree 3, that is a wheel of type 2 as in Figure 9.14b. □

As we shall see, either structure in Figure 9.14 leads to pinned edges.

Lemma 9.27. Pl-Holant($\neq_2 \mid =_5, \mathcal{EO}, \neq_2, \Delta_0, \Delta_1$) *is tractable.*

Proof. We proceed as in Lemma 9.25 up until the point of getting Ω_2. Note that due to (9.4) the only nontrivial E_5-blocks of arity ≤ 4 are ($\neq_2$) and those

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press



(a) Different signs of an E_5-block along the cycle lead to pinning

(b) Edges e and e' are pinned in wheels of type 2

Figure 9.15. Degeneracies in the wheel structure.

in Figure 9.12a, so that all the steps and transformations from Ω to Ω_1 and to Ω work exactly as in Lemma 9.25. Now each connected component of Ω_2 is planar and bipartite with vertices on one side having degree ≥ 5 and those on the other side ≥ 3. If it has some parallel edges, then we can find a pinned edge just as in Lemma 9.25. So we may assume it has no parallel edges. We only need to show that there are pinned edges in Ω_2.

Therefore each connected component of Ω_2 satisfies the condition of Lemma 9.26. By Lemma 9.26, the wheel structure in Figure 9.14 appears. All we need to show is that wheel structures of either type contain pinned edges.

First we claim that if a wheel of either type has a E_5-block, call it E_1, on the outer cycle which has different signs on the two edges incident to it along the cycle, then the middle ($=_5$), denoted by E_o, is pinned. This is pictured in Figure 9.15a. It does not matter whether the wheel is type 1 or 2, or the position of E_1 relative to the special triangle P_1 in type 2. Because E_o is an equality, both e_1 and e_2, the two edges incident to E_o that are connected to the two ExactOne signatures flanking E_1, must take the same value. If both e_1 and e_2 are assigned 1, then the two incoming wires of E_1 along the cycle have to be both assigned 0, whereas they are marked by different signs. This is a contradiction. Hence both e_1 and e_2 are pinned to 0 as well as all edges of E_o.

We may therefore assume that each E_5-block has same signs along the outer cycle, either $++$ or $--$. If the wheel is of type 1, then there is no valid assignment such that E_o is assigned 0 because the cycle has odd length. In fact if E_o is assigned 0, then we can remove E_o and its incident edges, and effectively the five ExactOne signatures are now ($\neq_2$)'s forming a 5-cycle linked by binary equalities. Hence all incident edges of E_o are pinned to 1.

Otherwise the wheel is of type 2, and each E_5-block has signs $++$ or $--$ along the outer cycle. We denote by P_1 the special ExactOne$_d$ function that has

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

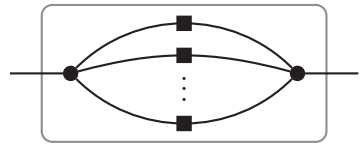


Figure 9.16. A binary gadget with $(=_r)$ assigned to circle vertices and $(\neq_2)$ assigned to square vertices.

arity $d > 3$. We claim that the two edges e and e' incident to P_1 along the cycle are both pinned to 0. This is illustrated in Figure 9.15b. As P_1 is EXACTONE_d, at most one of e and e' is 1. If one of e and e' is 1, the other is 0, and as P_1 is an EXACTONE_d function its edge to E_o is also 0, and thus all edges incident to E_o are 0. As all five neighbors of E_o are EXACTONE functions, the four EXACTONE_3 functions effectively become $(\neq_2)$ functions along the wheel, and we can remove E_o and its incident edges. This becomes the same situation as in the previous case of type 1, where effectively a cycle of five binary equalities are linked by five binary disequalities, which has no valid assignment. It implies that both e and e' are pinned to 0. This finishes the proof. □

Now we discuss the implications of Lemma 9.25 and Lemma 9.27 for planar hypergraph perfect matching.

Let $\mathcal{G} = \{=_k \mid k \in S\}$ be a set of EQUALITY signatures, where S is a set of positive integers containing at least one $r \geq 3$. Moreover let $\mathcal{EO}^+ := \{\text{EXACTONE}_d \mid d \in \mathbb{Z}^+\} = \mathcal{EO} \cup \{(\neq_2), [0, 1]\}$. Then Pl-Holant$(\mathcal{G} \mid \mathcal{EO}^+)$ is the problem of counting perfect matchings over hypergraphs with planar incidence graphs, where the hyperedge sizes are prescribed by S. In the incidence graph, vertices assigned signatures in $\mathcal{G}$ on the left represent hyperedges, and vertices assigned signatures in $\mathcal{EO}^+$ on the right represent vertices of the hypergraph. Let $t = \gcd(S)$. We show that this problem is tractable if $t \geq 5$ and #P-hard if $t \leq 4$. The tractability when $t \geq 5$ follows from Lemma 9.25 and 9.27, since we can reduce Pl-Holant$(\mathcal{G} \mid \mathcal{EO}^+)$ to Pl-Holant$(\neq_2 \mid =_t, \mathcal{EO}, (\neq_2), [0, 1])$. The reduction goes as follows. With $(\neq_2)$ on the LHS and $=_t$ on the RHS, we can construct all E_t-blocks and hence all of $\mathcal{EQ}_t$ on the right. Note that $\mathcal{G} \subseteq \mathcal{EQ}_t$. Then we move all signatures in $\mathcal{G}$ to the left via $(\neq_2)$.

The hardness of Pl-Holant$(\mathcal{G} \mid \mathcal{EO}^+)$ for $t \leq 4$ follows from Corollary 9.23. The reason is as follows. We construct $(\neq_2)$ on the left using the gadget pictured in Figure 9.16 with $(=_r) \in \mathcal{G}$ on the left side assigned to circle vertices and $(\neq_2)$ on the right side assigned to square vertices. Then we move $\mathcal{G}$ to the right side via $(\neq_2)$ on the right side. We construct $(=_t)$ on the right side in the same Euclidean process using $\mathcal{G}$ of the right side and $(\neq_2)$ of the left side. This gives us a reduction from Pl-Holant$(\neq_2 \mid =_t, \mathcal{EO})$, which is

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

#P-hard by Corollary 9.23 if $t = 3, 4$. Otherwise $t = 1, 2$. Recall that $(=_r) \in \mathcal{G}$ for some $r \geq 3$. We use $(=_t)$ to reduce the arity of $(=_r)$ to 3 or 4, if necessary. Again we are done by Corollary 9.23.

If there is no hyperedge of size ≥ 3 in Pl-Holant$(\mathcal{G} \mid \mathcal{EO}^+)$, then for $S = \{1\}$, the problem is tractable trivially, and for $S = \{2\}$, the problem is tractable by Kasteleyn's algorithm. The problem is #P-hard if $S = \{1, 2\}$, because we can realize nonperfect matching signatures $[1, 1, 0, \ldots, 0]$. In summary, we have the following theorem.

Theorem 9.28. *The problem* Pl-Holant$(\mathcal{G} \mid \mathcal{EO}^+)$ *counts perfect matchings over hypergraphs with planar incidence graphs, where the hyperedge sizes are prescribed by a set S of positive integers. Let* $t = \gcd(S)$. *If* $t \geq 5$ *or* $S = \{1\}$ *or* $\{2\}$, *then the problem is computable in polynomial time. Otherwise* $S = \{1, 2\}$, *or* $t \leq 4$ *and* $S \not\subseteq \{1, 2\}$, *then the problem is* #P*-hard.*

By looking at the proof of Lemma 9.19 and Lemma 9.21 closely, one can improve the #P-hardness part of Theorem 9.28 further.

Exercise: Prove that counting perfect matchings in 3-uniform 3-regular planar hypergraphs is also #P-hard.

Exercise: Prove the same result for 4-uniform 3-regular planar hypergraphs.

9.7 Dichotomy for Planar Holant

With the new tractable problems related to hypergraph perfect matchings found, together with $\mathscr{M}$-transformable signatures due to holographic reductions to Kasteleyn's algorithm, it turns out that there are no more planar tractable problems in the Pl-Holant framework for any set of symmetric signatures $\mathcal{F}$. The following is the main dichotomy theorem for planar Holant problems for complex valued symmetric signatures on Boolean variables. The proof of this theorem is omitted here; the interested readers can find it in [CFGW15].

For a set $\mathcal{F}$ of signatures, $\mathcal{F}_{nd}^{\geq 3}$ denotes the set of nondegenerate signatures in $\mathcal{F}$ of arity ≥ 3, and $\mathcal{F}^*$ denotes $\mathcal{F}$ with all degenerate signatures $[a, b]^{\otimes m}$ replaced by unary $[a, b]$.

Theorem 9.29. *Let* $\mathcal{F}$ *be any set of symmetric, complex-valued signatures in Boolean variables. Then* Pl-Holant$(\mathcal{F})$ *is* #P*-hard unless* $\mathcal{F}$ *satisfies one of the following conditions:*

1. *All nondegenerate signatures in* $\mathcal{F}$ *are of arity* ≤ 2;
2. $\mathcal{F}$ *is* $\mathscr{A}$*-transformable.*

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press

3. $\mathcal{F}$ *is* $\mathscr{P}$*-transformable.*
4. $\mathcal{F} \subseteq \mathscr{V}^{\sigma} \cup \{f \in \mathscr{R}_2^{\sigma} \mid \text{arity}(f) = 2\}$ *for some* $\sigma \in \{+, -\}$.
5. *All nondegenerate signatures in* $\mathcal{F}$ *are in* $\mathscr{R}_2^{\sigma}$ *for some* $\sigma \in \{+, -\}$.
6. $\mathcal{F}$ *is* $\mathscr{M}$*-transformable.*
7. $\mathcal{F} \subseteq Z\mathscr{P} \cup \mathscr{M}_4^{\sigma}$ *for some* $\sigma \in \{+, -\}$*, and the greatest common divisor of the arities of the signatures in* $\mathcal{F}^* \cap \mathscr{P}_2$ *is* ≥ 5.

In each exceptional case, Pl-Holant($\mathcal{F}$) *is computable in polynomial time. If* $\mathcal{F}$ *satisfies condition 1, 2, 3, 4, or 5, then* Holant($\mathcal{F}$) *is computable in polynomial time without planarity; otherwise* Holant($\mathcal{F}$) *is* #P*-hard.*

https://doi.org/10.1017/9781107477063.010 Published online by Cambridge University Press