Counting Perfect Matchings and the Eight-Vertex Model

Jin-Yi Cai

University of Wisconsin-Madison, Madison, WI, USA
http://pages.cs.wisc.edu/~jyc/
jyc@cs.wisc.edu

Tianyu Liu

University of Wisconsin-Madison, Madison, WI, USA
http://pages.cs.wisc.edu/~tl/
tl@cs.wisc.edu

Abstract

We study the approximation complexity of the partition function of the eight-vertex model on general 4-regular graphs. For the first time, we relate the approximability of the eight-vertex model to the complexity of approximately counting perfect matchings, a central open problem in this field. Our results extend those in [8].

In a region of the parameter space where no previous approximation complexity was known, we show that approximating the partition function is at least as hard as approximately counting perfect matchings via approximation-preserving reductions. In another region of the parameter space which is larger than the region that is previously known to admit Fully Polynomial Randomized Approximation Scheme (FPRAS), we show that computing the partition function can be reduced to counting perfect matchings (which is valid for both exact and approximate counting). Moreover, we give a complete characterization of nonnegatively weighted (not necessarily planar) 4-ary matchgates, which has been open for several years. The key ingredient of our proof is a *geometric* lemma.

We also identify a region of the parameter space where approximating the partition function on *planar* 4-regular graphs is feasible but on *general* 4-regular graphs is equivalent to approximately counting perfect matchings. To our best knowledge, these are the first problems that exhibit this dichotomic behavior between the planar and the nonplanar settings in approximate counting.

2012 ACM Subject Classification Theory of computation → Problems, reductions and completeness

Keywords and phrases Approximate complexity, the eight-vertex model, counting perfect matchings

Digital Object Identifier 10.4230/LIPIcs.ICALP.2020.23

Category Track A: Algorithms, Complexity and Games

Related Version A full version of the paper is available at https://arxiv.org/abs/1904.10493.

Funding Supported by NSF CCF-1714275.

1 Introduction

The eight-vertex model is defined over 4-regular graphs, the states of which are the set of *even orientations*, i.e. those with an even number of arrows into (and out of) each vertex. There are eight permitted types of local configurations around a vertex – hence the name eight-vertex model (see Figure 1).

Classically, the eight-vertex model is defined by statistical physicists on a square lattice region where each vertex of the lattice is connected by an edge to four nearest neighbors. In general, the eight configurations 1 to 8 in Figure 1 are associated with eight possible weights $w_1, \ldots, w_8$. By physical considerations, the total weight of a state remains unchanged if all arrows are flipped, assuming there is no external electric field. In this case we write $w_1 = w_2 = a$, $w_3 = w_4 = b$, $w_5 = w_6 = c$, and $w_7 = w_8 = d$. This complementary invariance is known as *arrow reversal symmetry* or *zero field assumption*.

47th International Colloquium on Automata, Languages, and Programming (ICALP 2020).
Editors: Artur Czumaj, Anuj Dawar, and Emanuela Merelli; Article No. 23; pp. 23:1–23:18
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

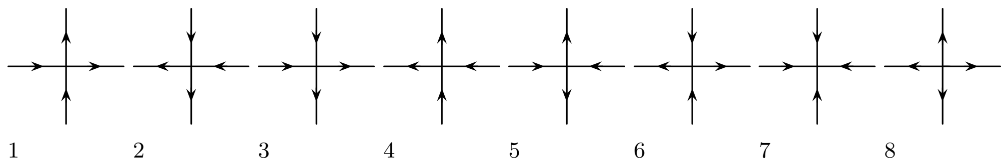


Figure 1 Valid configurations of the eight-vertex model.

Even in the zero-field setting, this model is already enormously expressive: the special case when $d = 0$ is the zero-field six-vertex model which has sub-models such as the ice $(a = b = c)$, KDP, and Rys F models; on the square lattice, some other important models such as the dimer and zero-field Ising models can be reduced to the eight-vertex model [2]. After it was introduced in 1970 by Sutherland [19] and by Fan and Wu [9], Baxter [1, 2] achieved a good understanding of the zero-field case in the thermodynamic limit on the square lattice (in physics it is called an "exactly solved model").

In this paper, we assume the arrow reversal symmetry and further assume that $a, b, c, d \geq 0$, as is the case in *classical* physics. Given a 4-regular graph G, we label four incident edges of each vertex from 1 to 4. The *partition function* of the eight-vertex model with parameters (a, b, c, d) on G is defined as

$$Z_{\mathrm{EV}}(G; a, b, c, d) = \sum_{\tau \in \mathcal{O}_{\mathbf{e}}(G)} a^{n_1+n_2} b^{n_3+n_4} c^{n_5+n_6} d^{n_7+n_8}, \tag{1}$$

where $\mathcal{O}_{\mathbf{e}}(G)$ is the set of all even orientations of G, and n_i is the number of vertices in type i in G ($1 \leq i \leq 8$, locally depicted as in Figure 1 where the 4 edges are oriented counterclockwise starting from the edge on the left) under an even orientation $\tau \in \mathcal{O}_{\mathbf{e}}(G)$.

In terms of the exact computational complexity, a complexity dichotomy is given for the eight-vertex model on 4-regular graphs for all eight parameters [6]. This is studied in the context of a classification program for the complexity of counting problems [5], where the eight-vertex model serves as important basic cases for Holant problems defined by not necessarily symmetric constraint functions. It is shown that every setting is either P-time computable (and some are surprising) or #P-hard. However, most cases for P-time tractability are due to nontrivial cancellations. In our setting where a, b, c, d are nonnegative, the problem of computing the partition function of the eight-vertex model exactly is #P-hard unless: (1) $a = b = c = d$ (this is equivalent to the unweighted case); (2) at least three of a, b, c, d are zero; or (3) two of a, b, c, d are zero and the other two are equal. In addition, on planar graphs it is also P-time computable for parameter settings (a, b, c, d) with $a^2 + b^2 = c^2 + d^2$, using the *FKT algorithm*.

Since exact computation is hard in most cases, one natural question is what is the approximate complexity of counting and sampling of the eight-vertex model. To our best knowledge, prior to [8], there is only one previous result in this regard due to Greenberg and Randall [12]. They showed that on square lattice regions a specific Markov chain (which flips the orientations of all four edges along a uniformly picked face at each step) is torpidly mixing when d is large. This means that when sinks and sources have large weights, this particular chain *cannot* be used to approximately count or sample eight-vertex configurations on the square lattice according to the Gibbs measure. Recently, similar torpid mixing results have been achieved for the six-vertex model on the square lattice [17].

The paper [8] gave the first classification results for the approximate complexity of the eight-vertex model on general and planar 4-regular graphs, and they conform to phase transition in physics. In order to state the results, we adopt the following notations assuming $a, b, c, d \geq 0$.

- $\mathcal{X} = \{\ (a, b, c, d) \mid a \le b + c + d,\ \ b \le a + c + d,\ \ c \le b + c + d,\ \ d \le a + b + c\}$;
- $\mathcal{Y} = \{\ (a, b, c, d) \mid a + d \le b + c,\ \ b + d \le a + c,\ \ c + d \le a + b\}$;
- $\mathcal{Z} = \{\ (a, b, c, d) \mid a^2 \le b^2+c^2+d^2,\ \ b^2 \le a^2+c^2+d^2,\ \ c^2 \le a^2+b^2+d^2,\ \ d^2 \le a^2+b^2+c^2\}$.

▶ Remark 1. $\mathcal{Y} \subset \mathcal{X}$. $\mathcal{Z} \subset \mathcal{X}$.

Physicists have shown an *order-disorder phase transition* for the eight-vertex model on the square lattice between parameter settings outside $\mathcal{X}$ and those inside $\mathcal{X}$ (see Baxter's book [3] for more details). In [8], it was shown that: (1) approximating the partition function of the eight-vertex model on general 4-regular graphs outside $\mathcal{X}$ is NP-hard, (2) there is an FPRAS[1] for general 4-regular graphs in the region $\mathcal{Y} \bigcap \mathcal{Z}$, and (3) there is an FPRAS for planar 4-regular graphs in the extra region $\{(a, b, c, d) \mid a + d \le b + c,\ \ b + d \le a + c,\ \ c + d \ge a + b\} \bigcap \mathcal{Z}$.

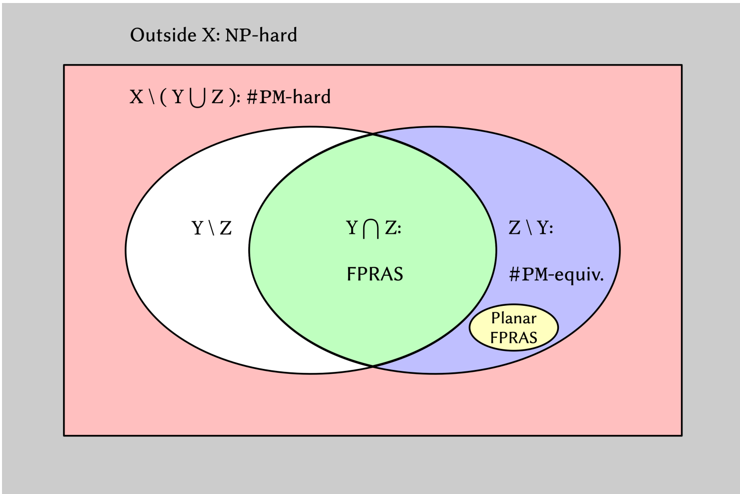


Figure 2 A Venn diagram of the approximation complexity of the eight-vertex model.

In this paper we make further progress in the classification of the approximate complexity of the eight-vertex model on 4-regular graphs in terms of the parameters (see Figure 2). For the first time, the complexity of approximating the partition function of the eight-vertex model ($\#\mathrm{EV}(a, b, c, d)$) is related to that of approximately counting perfect matchings ($\#\mathrm{PM}$).

▶ **Theorem 2.** *For any four positive numbers $a, b, c, d > 0$ such that $(a, b, c, d) \notin \mathcal{Y}$, the problem $\#\mathrm{EV}(a, b, c, d)$ is at least as hard to approximate as counting perfect matchings:*

$$\#\mathrm{PM} \le_{\mathrm{AP}} \#\mathrm{EV}(a, b, c, d).$$

▶ Remark 3. The theorem is stated for the case where all four parameters are positive. The same proof also works for the case when there is exactly one zero among the nonnegative values $\{a, b, c\}$. A complete account for four nonnegative values $\{a, b, c, d\}$ is given in the Table 1. There is a symmetry among a, b, c for the eight-vertex model on general (not necessarily planar) 4-regular graphs, so for simplicity in this table we assume $a \le b \le c$.

[1] Suppose $f : \Sigma^* \to \mathbb{R}$ is a function mapping problem instances to real numbers. A *fully polynomial randomized approximation scheme (FPRAS)* [14] for a problem is a randomized algorithm that takes as input an instance x and $\varepsilon > 0$, running in time polynomial in n (the input length) and ε^{-1}, and outputs a number Y (a random variable) such that $\Pr[(1-\varepsilon)f(x) \le Y \le (1+\varepsilon)f(x)] \ge \frac{3}{4}$.

Table 1 Approximation complexity of the eight-vertex model with $(a,b,c,d) \notin d\text{-}\mathtt{SUM}^2$.

	$d = 0$	$d > 0$
$a = b = c = 0$	P-time computable (trivial)	P-time computable (trivial)
$a = b = 0, c > 0$	P-time computable (trivial)	$c = d$: P-time computable [6] $c \neq d$: NP-hard [8]
$a = 0, b, c > 0$	NP-hard [8]	#PM-hard (in this paper)
$a, b, c > 0$	NP-hard [7]	#PM-hard (in this paper)

The proof of Theorem 2 is in Section 3. Our proof for the hardness result has several ingredients:

1. We express the eight-vertex model on a 4-regular graph G as an edge-2-coloring problem on G using Valiant's *holographic transformation* [22].
2. We show that a modified version of the edge-2-coloring problem on G is equivalent to the zero-field Ising model on its *crossing-circuit graph* denoted by $\widetilde{G}$. Thus known #PM-equivalence result for the Ising model [11, Lemma 7] directly transfers to the modified version of the edge-2-coloring problem under certain parameter settings.
3. We further show that for any parameter setting outside $\mathcal{Y}$, approximating the partition function of the eight-vertex model is at least as hard as the #PM-equivalent modified edge-2-coloring problem via approximation-preserving reductions.

▶ **Theorem 4.** *For any $(a,b,c,d) \in \mathcal{Z}$,*

$$\#\mathrm{EV}(a,b,c,d) \leq_{\mathrm{AP}} \#\mathrm{PM}.$$

The proof of Theorem 4 is in Section 4. To prove the #PM-easiness result, we again express the eight-vertex model in the Holant framework (see Section 2) and show that the constraint functions of the eight-vertex model in $\mathcal{Z}$ can be implemented by constant-size *matchgates* with nonnegatively weighted edges (Definition 13). We note that allowing nonnegative edge-weights does not add more computational power to the unweighted #PM [18, Proposition 5]. The crucial ingredient of our proof is a *geometric* lemma (Lemma 18) in 3-dimensional space.

This matchgate expressibility is tight: no constraint functions of the eight-vertex model with parameter settings outside the region $\mathcal{Z}$ can be implemented by a matchgate (Lemma 19). Moreover, the general version of our result also works for the eight-vertex model without the arrow reversal symmetry. It is open if approximately computing the partition function in $\mathcal{X} \setminus (\mathcal{Y} \bigcup \mathcal{Z})$ is #PM-equivalent or not.

As part of this work, we give a complete characterization of the constraint functions that can be expressed by 4-ary matchgates in Theorem 15. This solves an important question that has been open for several years [18, 4]. We believe it is of independent interest.

▶ **Corollary 5.** *For any four positive numbers $a,b,c,d > 0$ such that $(a,b,c,d) \in \mathcal{Z} \setminus \mathcal{Y}$,*

$$\#\mathrm{EV}(a,b,c,d) \equiv_{\mathrm{AP}} \#\mathrm{PM}.$$

Note that for the eight-vertex model in the region[3] $\{(a,b,c,d) \mid a+d \le b+c, \quad b+d \le a+c, \quad c+d > a+b\} \bigcap \mathcal{Z}$, computing $Z_{\mathrm{EV}}(a,b,c,d)$ is (1) #P-complete in exact computation [6], (2) #PM-equivalent in approximate computation on general 4-regular graphs (Corollary 5), and (3) admits an FPRAS in approximate computation on planar 4-regular graphs [8]. To our best knowledge, these are the first identified problems having these three properties. Previously the combined results of [10] and [13] proved that counting k-colorings for certain range of parameters is FPRASable on general graphs but NP-hard in approximate complexity. The complexity result in this paper is different because the complexity we described in item (2) above is #PM-equivalent (neither harder nor easier). We note that the complexity status of #PM is a long standing open problem in the field (neither known to be FPRASable nor known to be NP-hard to approximate). These results contrast with, in the setting of approximately counting, the FKT algorithm for exact counting which shows that the #P-hard problem #PM can be computed in polynomial time on planar graphs.

2 Preliminaries

Given a 4-regular graph $G=(V,E)$, the *edge-vertex incidence graph* $G'=(U_E,U_V,E')$ is a bipartite graph where $(u_e,u_v) \in U_E \times U_V$ is an edge in E' iff $e \in E$ in G is incident to $v \in V$. We model an orientation $(w \to v)$ on an edge $e=\{w,v\} \in E$ from w into v in G by assigning 1 to $(u_e,u_w) \in E'$ and 0 to $(u_e,u_v) \in E'$ in G'. A configuration of the eight-vertex model on G is an *edge-2-coloring* on G', namely $\sigma : E' \to \{0,1\}$, where for each $u_e \in U_E$ its two incident edges are assigned 01 or 10, and for each $u_v \in U_V$ the sum of values $\sum_{i=1}^{4} \sigma(e_i) \equiv 0 \pmod 2$, over the four incident edges of u_v. Thus we model the even orientation rule of G on all $v \in V$ by requiring "two-0-two-1/four-0/four-1" locally at each vertex $u_v \in U_V$.

The "one-0-one-1" requirement on the two edges incident to a vertex in U_E is a binary Disequality constraint, denoted by $(\neq_2)$. The values of a 4-ary *constraint function*, or a *signature* f can be listed in a matrix $M(f) = \begin{bmatrix} f_{0000} & f_{0010} & f_{0001} & f_{0011} \\ f_{0100} & f_{0110} & f_{0101} & f_{0111} \\ f_{1000} & f_{1010} & f_{1001} & f_{1011} \\ f_{1100} & f_{1110} & f_{1101} & f_{1111} \end{bmatrix}$, called the *constraint matrix* of f. For the eight-vertex model satisfying the even orientation rule and arrow reversal symmetry, the signature f at every vertex $v \in U_V$ in G' has the form $M(f) = \begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix}$, if we draw a vertex with incident edges labeled 1, 2, 3, and 4 locally as the left, down, right, and up edges respectively according to Figure 1. Thus computing the partition function $Z_{\mathrm{EV}}(G;a,b,c,d)$ is equivalent to evaluating

$$Z'(G';f) := \sum_{\sigma:E'\to\{0,1\}} \prod_{u\in U_E} (\neq_2)\left(\sigma|_{E'(u)}\right) \prod_{u\in U_V} f\left(\sigma|_{E'(u)}\right).$$

where $E'(u)$ denotes the incident edges of $u \in U_E \cup U_V$. In fact, in this way we express the partition function of the eight-vertex model as the Holant sum in the framework for Holant problems:

$$Z_{\mathrm{EV}}(G;a,b,c,d) = \mathrm{Holant}\left(G'; \neq_2 \mid f\right)$$

where we use $\mathrm{Holant}(H; g \mid f)$ to denote the sum $\sum_{\sigma:E\to\{0,1\}} \prod_{u\in U} g\left(\sigma|_{E(u)}\right) \prod_{u\in V} f\left(\sigma|_{E(u)}\right)$ on a bipartite graph $H=(U,V,E)$. Each vertex in U (or V) is assigned the signature g (or f, respectively). The signature g is considered as a row vector (or covariant tensor),

[3] This region is the intersection of $\overline{\mathcal{Y}}$ and the extra region $\{(a,b,c,d) \mid a+d \le b+c, \quad b+d \le a+c, \quad c+d \ge a+b\}$ where there is an FPRAS for planar graphs. Note that the strict inequality $c+d > a+b$ is needed.

whereas the signature f is considered as a column vector (or contravariant tensor). (See [5] for more on Holant problems.) The following proposition says that an invertible holographic transformation does not change the complexity of the Holant problem in the bipartite setting.

▶ **Proposition 6** ([22]). *Suppose $T \in \mathbb{C}^2$ is an invertible matrix. Let $d_1 = \operatorname{arity}(g)$ and $d_2 = \operatorname{arity}(f)$. Define $g' = g\left(T^{-1}\right)^{\otimes d_1}$ and $f' = T^{\otimes d_2} f$. Then for any bipartite graph H, $\operatorname{Holant}(H; g \mid f) = \operatorname{Holant}(H; g' \mid f')$.*

We denote $\operatorname{Holant}(G; f) = \operatorname{Holant}(G'; {=_2} \mid f)$ and use $\operatorname{Holant}(f)$ to denote the problem whose input is a graph G and output is $\operatorname{Holant}(G; f)$; this is equivalent to the usual definition.

3 #PerfectMatchings-hardness

Our proof strategy for Theorem 2 is as follows. In Lemma 7, we express the eight-vertex model on a 4-regular graph G as a Holant problem; this is an equivalent form of the orientation problem expressed as an edge-2-coloring problem on G, and is achieved using a holographic transformation. In Lemma 8, we give an approximation-preserving reduction to show that this edge-2-coloring problem is at least as hard as a modified version of the edge-2-coloring problem where weights at some input originally in the support are dropped off. In Lemma 9, we establish the equivalence between this modified version of the edge-2-coloring problem and the zero-field Ising model. Thus a known result for the Ising model (Proposition 11) indicates the #PM-equivalence of this modified version of the edge-2-coloring problem under certain parameter settings (Corollary 12). It can be deduced from Lemma 7, Lemma 8, and Corollary 12 that for any (a, b, c, d) with $a + d > b + c$ (and symmetrically $b + d > a + c$ or $c + d > a + b$), approximately computing the partition function is at least as hard as the #PM-equivalent modified edge-2-coloring problem under approximation-preserving reductions.

▶ **Lemma 7.**

$$2^{|V(G)|} \cdot Z_{\mathrm{EV}}(G; a, b, c, d) = \operatorname{Holant}\left(G; \begin{bmatrix} a{+}b{+}c{+}d & 0 & 0 & -a{+}b{+}c{-}d \\ 0 & a{-}b{+}c{-}d & a{+}b{-}c{-}d & 0 \\ 0 & a{+}b{-}c{-}d & a{-}b{+}c{-}d & 0 \\ -a{+}b{+}c{-}d & 0 & 0 & a{+}b{+}c{+}d \end{bmatrix}\right).$$

Proof. Using the binary disequality function ($\neq_2$) for the orientation of any edge, we can express the partition function of the eight-vertex model G as a Holant problem on its edge-vertex incidence graph G',

$$Z_{\mathrm{EV}}(G; a, b, c, d) = \operatorname{Holant}\left(G'; {\neq_2} \mid f\right),$$

where f is the 4-ary signature with $M(f) = \left[\begin{smallmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{smallmatrix}\right]$. Note that, writing the truth table of $(\neq_2) = (0, 1, 1, 0)$ as a vector and multiplied by a tensor power of the matrix Z^{-1}, where $Z = \frac{1}{\sqrt{2}}\left[\begin{smallmatrix} 1 & 1 \\ i & -i \end{smallmatrix}\right]$ we get $(\neq_2)(Z^{-1})^{\otimes 2} = (1, 0, 0, 1)$, which is exactly the truth table of the binary equality function ($=_2$). Then according to Proposition 6, by the Z-transformation, we get

$$\begin{aligned} \operatorname{Holant}\left(G'; {\neq_2} \mid f\right) &= \operatorname{Holant}\left(G'; {\neq_2} \cdot \left(Z^{-1}\right)^{\otimes 2} \mid Z^{\otimes 4} \cdot f\right) \\ &= \operatorname{Holant}\left(G'; {=_2} \mid Z^{\otimes 4} f\right) \\ &= \operatorname{Holant}\left(G;\ Z^{\otimes 4} f\right), \end{aligned}$$

and a direct calculation shows that $M(Z^{\otimes 4} f) = \frac{1}{2}\begin{bmatrix} a{+}b{+}c{+}d & 0 & 0 & -a{+}b{+}c{-}d \\ 0 & a{-}b{+}c{-}d & a{+}b{-}c{-}d & 0 \\ 0 & a{+}b{-}c{-}d & a{-}b{+}c{-}d & 0 \\ -a{+}b{+}c{-}d & 0 & 0 & a{+}b{+}c{+}d \end{bmatrix}$. ◀

▶ **Lemma 8.** *Suppose $d > 0$ and at most one of a, b, c is zero. Then*

$$\text{Holant}\left(\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix}\right) \leq_{\text{AP}} \text{Holant}\left(\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & a-b+c-d & a+b-c-d & 0 \\ 0 & a+b-c-d & a-b+c-d & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix}\right).$$

Proof. This task can be reduced to

$$\text{Holant}\left(\neq_2 \mid \begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}\right) \leq_{\text{AP}} \text{Holant}\left(\neq_2 \mid \begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix}\right) \tag{2}$$

and the analysis can be found in the full version of this paper.

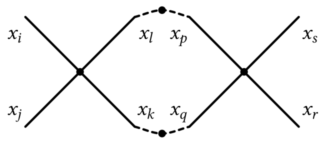


Figure 3 A gadget construction.

Next we show how to get (2). Given the signature f with matrix $\begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix}$ in #EV(a,b,c,d), we construct a 4-ary signature $\check{f}$ with constraint matrix $\begin{bmatrix} \check{d} & 0 & 0 & \check{a} \\ 0 & \check{b} & \check{c} & 0 \\ 0 & \check{c} & \check{b} & 0 \\ \check{a} & 0 & 0 & \check{d} \end{bmatrix}$ using a polynomial number of vertices and edges such that $\check{a}$, $\check{b}$, $\check{c}$, and $\check{d}$ are all exponentially close to 1 after normalization, i.e., to be 2^{-n^C} close to 1, for any $C > 0$, with a construction of $n^{O(1)}$ size in polynomial time.

We assume we start with the following condition:

$$0 < d \leq a \leq b \leq c. \tag{3}$$

If this is not the case, we can obtain a 4-ary construction that realizes this condition using constantly many vertices. With some preliminary construction we can further assume $1 \leq d \leq a \leq b \leq c \leq \frac{3}{2}d$ initially. (See the full version of this paper for details.) Note that starting with the signature f with matrix $M(f) = \begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix}$, we can arbitrarily permute a, b, c by relabeling the edges, and so we get signatures f_1 with $M(f_1) = \begin{bmatrix} d & 0 & 0 & b \\ 0 & a & c & 0 \\ 0 & c & a & 0 \\ b & 0 & 0 & d \end{bmatrix}$ and f_2 with $M(f_2) = \begin{bmatrix} d & 0 & 0 & c \\ 0 & a & b & 0 \\ 0 & b & a & 0 \\ c & 0 & 0 & d \end{bmatrix}$. There are two constructions G_1 and G_2 which we use as basic steps; both constructions start with a signature f with parameters satisfying (3).

1. $\boldsymbol{G_1}$: connect two vertices with signatures f_1 and f_2 respectively as in Figure 3. Since we are in the orientation view, we place the signature ($\neq_2$) on the two degree 2 vertices connecting the two degree 4 vertices. Then the signature g_1 of the construction G_1 is obtained by matrix multiplication $M(g_1) = M_{x_i x_j, x_s x_r}(g_1) = M(f_1) \cdot N \cdot M(f_2)$, where $N = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & 1 & & \\ 1 & & & \end{bmatrix}$. Thus

$$M(g_1) = \begin{bmatrix} (b+c)d & 0 & 0 & bc+d^2 \\ 0 & a(b+c) & a^2+bc & 0 \\ 0 & a^2+bc & a(b+c) & 0 \\ bc+d^2 & 0 & 0 & (b+c)d \end{bmatrix}.$$

The signature g_1 has four new parameters, denoted by

$$(a_1,\ b_1,\ c_1,\ d_1) = (a(b+c),\ bc+d^2,\ a^2+bc,\ (b+c)d).$$

We make the following observations and all of them can be easily verified using (3):

- d_1 is the weight on sink and source and $0 < d_1 \le a_1, b_1, c_1$.
- $c_1 = \max(a_1, b_1, c_1, d_1)$.
- $\frac{a_1}{d_1} = \frac{a}{d}, \frac{b_1}{d_1} \le \frac{b}{d}, \frac{c_1}{d_1} \le \frac{c}{d}$.
- $c_1 d_1 \le a_1 b_1$ because $c_1 d_1 - a_1 b_1 = -(b+c)(a-d)(bc-ad) \le 0$.

2. $\boldsymbol{G_2}$: connect two vertices with signatures f_2 as in Figure 3. Denote the signature of G_2 by g_2. We have

$$M(g_2) = M(f_2) \cdot N \cdot M(f_2) = \begin{bmatrix} 2cd & 0 & 0 & c^2+d^2 \\ 0 & 2ab & a^2+b^2 & 0 \\ 0 & a^2+b^2 & 2ab & 0 \\ c^2+d^2 & 0 & 0 & 2cd \end{bmatrix}.$$

The signature g_2 has four new parameters, denoted by

$$(a_2,\ b_2,\ c_2,\ d_2) = (2ab,\ a^2+b^2,\ c^2+d^2,\ 2cd).$$

The following observations can also be easily verified using (3):

- d_2 is the weight on sink and source and if $cd \le ab$, then $0 < d_2 \le a_2, b_2, c_2$.
- $\frac{a_2}{d_2} \le \frac{a}{d}, \frac{b_2}{d_2} \le \frac{b}{d}, \frac{c_2}{d_2} \le \frac{c}{d}$.
- $\frac{c_2-d_2}{d_2} = \frac{(c-d)^2}{2cd} \le \frac{1}{2}\left(\frac{c-d}{d}\right)^2 \le \left(\frac{c-d}{d}\right)^2$.

Based on the two basic constructions above, we construct the signature $\check{f}$ in logarithmically many rounds recursively, each of the $O(\log n)$ rounds uses the signature constructed in the previous round. We now describe a single round in this construction, which consists of two steps. In step 1 we use a signature with some parameter setting (a, b, c, d) satisfying (3) and apply G_1 to two copies of the signature. If the resulting parameter $b_1 < a_1$ we switch the roles of a_1 and b_1, and obtain $(a_1', b_1', c_1', d_1') = (b_1, a_1, c_1, d_1)$, again satisfying (3), as well as $c_1 d_1 \le a_1 b_1$. In step 2, we apply G_2 to two copies of the signature constructed in step 1 (with the switching of the roles of a_1 and b_1 if it is needed). Denote the parameters of the resulting signature by (a^*, b^*, c^*, d^*). Altogether each round uses four copies of the signature from the previous round, starting with the initial given signature. Therefore in polynomial time we can afford to carry out $C \log n$ rounds for any constant C. Note that, if we consider the normalized quantities $(\frac{a}{d}, \frac{b}{d}, \frac{c}{d}, \frac{d}{d})$, then the respective quantities in each step G_1 and G_2 do not increase their distances to 1, i.e.,

$$0 \le \frac{a^*}{d^*} - 1 \le \frac{a}{d} - 1, \quad 0 \le \frac{b^*}{d^*} - 1 \le \frac{b}{d} - 1, \quad 0 \le \frac{c^*}{d^*} - 1 \le \frac{c}{d} - 1.$$

This is true even if the G_2 construction in step 2 is applied in the case when the roles of a_1 and b_1 are switched for the signature from step 1, when that switch is required ($b_1 < a_1$) as described. More importantly, based on the properties of G_1 and G_2, we know that the (normalized) gap between d and the previous largest entry c shrinks quadratically fast, as measured by the new c^* normalized with d^*. More precisely,

$$0 \le \frac{c^*}{d^*} - 1 \le \left(\frac{c}{d} - 1\right)^2.$$

Note that c^* may no longer be the largest among a^*, b^*, c^*; however we will permute them to get $\widetilde{a}, \widetilde{b}, \widetilde{c}$ so that (3) is still satisfied before proceeding to the next round. This completes the description of our construction in one round which obtains $(\widetilde{a}, \widetilde{b}, \widetilde{c}, \widetilde{d})$ from (a, b, c, d).

We will construct the final signature $\check{f}$ by $O(\log n)$ rounds of this construction. Also we will follow each value a, b, c individually as they get transformed through each round. To state it formally, starting with the normalized triple $(\frac{a}{d}, \frac{b}{d}, \frac{c}{d})$, we define a successor triple $(\frac{a^*}{d^*}, \frac{b^*}{d^*}, \frac{c^*}{d^*})$, so that each entry has the respective successor (e.g., the entry $\frac{a}{d}$ has successor $\frac{a^*}{d^*}$). This is well-defined because (a_1, b_1, c_1, d_1) and (a_2, b_2, c_2, d_2) are homogeneous functions of (a, b, c, d). Note that even though from one round to the next, we may have to rename a^*, b^*, c^* so that the permutated triple $\widetilde{a}, \widetilde{b}, \widetilde{c}$ satisfies (3), the successor sequence as the rounds progress stays with an individual value. E.g., starting from (a, b, c, d), if after one round $a^* = \max(a^*, b^*, c^*, d^*) = \widetilde{c}$, then the successor of $\frac{a}{d}$ after two rounds is $\frac{(\widetilde{c})^*}{(d^*)^*}$. Now define $(\alpha_k, \beta_k, \gamma_k)$ to be the (ordered) triple $(\frac{a}{d}, \frac{b}{d}, \frac{c}{d})$ for $k = 1$, or its successor triple, at the beginning of the k-th round for $k > 1$.

Let $\check{f}$ be the 4-ary signature constructed after $3(k+1)$ rounds. By the *Pigeonhole Principle*, after $3(k+1)$ rounds, at least one of a, b, c has the property that in at least $k+1$ many rounds (let $1 \le i_0 < i_1 < \ldots < i_k \le 3(k+1)$ be $k+1$ such rounds) the corresponding $\frac{a}{d}, \frac{b}{d}, \frac{c}{d}$ or its successors are the maximum (normalized) value in that round, and thus its next successor gets shrunken quadratically in that round. Suppose this is a; the same proof works if it is b or c. Let α_i be the maximum (normalized) value at the beginning of round i in $k+1$ rounds, where $i \in \{i_0, \ldots, i_k\}$. Since initially we have $1 \le d \le a \le b \le c \le \frac{3}{2}d$,

$$0 \le \alpha_{i_1} - 1 \le \alpha_{i_0+1} - 1 \le (\alpha_{i_0} - 1)^2 \le \frac{1}{2^2}.$$

Then

$$0 \le \alpha_{i_2} - 1 \le \alpha_{i_1+1} - 1 \le (\alpha_{i_1} - 1)^2 \le \frac{1}{2^{2^2}}.$$

By induction $0 \le \alpha_{i_k} - 1 \le \frac{1}{2^{2^k}}$. At the end of $3(k+1)$ rounds, if $\check{f}$ has parameters $(\check{a}, \check{b}, \check{c}, \check{d})$, then

$$0 \le \frac{\max(\check{a}, \check{b}, \check{c})}{\check{d}} - 1 \le \alpha_{i_k} - 1 \le \frac{1}{2^{2^k}}.$$

Therefore, after logarithmically many rounds, using polynomially many vertices, we can get a 4-ary construction with parameters $\check{a}$, $\check{b}$, $\check{c}$, and $\check{d}$ that are exponentially close to 1 after normalizing by $\check{d}$. Thus (2) is proved. ◀

Problem : $\textsc{Ising}(\beta)$.
Instance : Graph $G = (V, E)$.
Output: $Z_{\textsc{Ising}}(G; \beta) := \sum_{\sigma: V \to \{0,1\}} \beta^{\text{mono}(\sigma)}$, where $\text{mono}(\sigma)$ denotes the number of edges $\{u, v\}$ such that $\sigma(u) = \sigma(v)$.

▶ **Lemma 9.** *The Ising problem* $\textsc{Ising}\left(\frac{w}{x}\right)$ *is equivalent to the Holant problem* $\text{Holant}\left(\left[\begin{smallmatrix} w & 0 & 0 & x \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ x & 0 & 0 & w \end{smallmatrix}\right]\right)$. *In particular,* $\textsc{Ising}\left(\frac{w}{x}\right) \equiv_{\text{AP}} \text{Holant}\left(\left[\begin{smallmatrix} w & 0 & 0 & x \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ x & 0 & 0 & w \end{smallmatrix}\right]\right)$.

▶ Remark 10. A non-homogenized form of the Ising model is $\widetilde{Z}_{\textsc{Ising}}(G; x, w) := \sum_{\sigma: V \to \{0,1\}} w^{\text{mono}(\sigma)} x^{|E| - \text{mono}(\sigma)}$. If $x \neq 0$ then $\widetilde{Z}_{\textsc{Ising}}(G; x, w) = x^{|E|} Z_{\textsc{Ising}}(G; \frac{w}{x})$. If $x = 0$ then in $\widetilde{Z}_{\textsc{Ising}}$ all vertices in each component must take the same assignment (all 0 or all 1). In this case both $\widetilde{Z}_{\textsc{Ising}}(G; x, w)$ and the Holant problem in Lemma 9 are trivially solvable in polynomial time.

Proof. For the problem $\operatorname{Holant}\left(\begin{bmatrix} w&0&0&x\\ 0&y&z&0\\ 0&z&y&0\\ x&0&0&w \end{bmatrix}\right)$, the roles of x, y, z are interchangeable by relabeling the edges. For example, if the signature $f(x_1, x_2, x_3, x_4)$ has the constraint matrix $\begin{bmatrix} w&0&0&x\\ 0&y&z&0\\ 0&z&y&0\\ x&0&0&w \end{bmatrix}$, then the signature $f(x_1, x_3, x_2, x_4)$ has the constraint matrix $\begin{bmatrix} w&0&0&z\\ 0&y&x&0\\ 0&x&y&0\\ z&0&0&w \end{bmatrix}$. It follows that

$$\operatorname{Holant}\left(\begin{bmatrix} w&0&0&x\\ 0&0&0&0\\ 0&0&0&0\\ x&0&0&w \end{bmatrix}\right) \quad \text{and} \quad \operatorname{Holant}\left(\begin{bmatrix} w&0&0&0\\ 0&0&x&0\\ 0&x&0&0\\ 0&0&0&w \end{bmatrix}\right)$$

are exactly the same problem. So to prove the lemma it suffices to prove the equivalence of

$$\textsc{Ising}\left(\frac{w}{z}\right) \quad \text{and} \quad \operatorname{Holant}\left(\begin{bmatrix} w&0&0&0\\ 0&0&z&0\\ 0&z&0&0\\ 0&0&0&w \end{bmatrix}\right).$$

First we show that $\operatorname{Holant}\left(\begin{bmatrix} w&0&0&0\\ 0&0&z&0\\ 0&z&0&0\\ 0&0&0&w \end{bmatrix}\right)$ can be expressed as $\textsc{Ising}\left(\frac{w}{z}\right)$.

Given a 4-regular graph $G = (V, E)$ as an instance of $\operatorname{Holant}\left(\begin{bmatrix} w&0&0&0\\ 0&0&z&0\\ 0&z&0&0\\ 0&0&0&w \end{bmatrix}\right)$, we can partition E into a set $\mathcal{C}$ of *circuits* (in which vertices may repeat but edges cannot) in the following way: at every vertex $v \in V$, denote the four edges incident to v by e_1, e_2, e_3, e_4 in a cyclic order according to the local labeling of the signature function; we make e_1 and e_3 into adjacent edges in a single circuit, and similarly we make e_2 and e_4 into adjacent edges in a single circuit (note that these may be the same circuit). We say each circuit in $\mathcal{C}$ is a *crossing circuit* of G. For the graph G, we define its *crossing-circuit graph* $\widetilde{G} = (\mathcal{C}, \widetilde{E})$, with possible multiloops and multiedges, as follows: its vertex set $\mathcal{C}$ consists of the crossing circuits; for every $v \in V$, if circuits C_1 and C_2 intersect at v, then there is an edge $\widetilde{e}_v \in \widetilde{E}$ labeled by v. Note that it is possible that $C_1 = C_2$, and for such a self-intersectison point the edge $\widetilde{e}_v$ is a loop. Each $C \in \mathcal{C}$ may have multiple loops, and for distinct circuits C_1 and C_2 there may be multiple edges between them. The edge set $\widetilde{E}$ of $\widetilde{G}$ is in 1-1 correspondence with V of G.

Observe that the problem $\operatorname{Holant}\left(\begin{bmatrix} w&0&0&0\\ 0&0&z&0\\ 0&z&0&0\\ 0&0&0&w \end{bmatrix}\right)$ requires that every valid configuration σ (that contributes a non-zero term) obeys the following rule at each vertex v:

- Assuming e_1, e_2, e_3, e_4 are the four edges incident to v in cyclic order, then $\sigma(e_1) = \sigma(e_3)$ (denoted by b_1) and $\sigma(e_2) = \sigma(e_4)$ (denoted by b_2). That is to say, all edges in a crossing circuit must have the same assignment (either all 0 or all 1). Therefore, the valid configurations σ on the edges of G are in 1-1 correspondence with $0,1$-assignments σ' on the vertices of $\widetilde{G}$.
- Under σ, the local weight on v is w if $b_1 = b_2$ and is z otherwise. Suppose crossing circuits C_1 and C_2 intersect at v (they could be identical). Then in $\widetilde{G}$, σ' has local weight w on the edge $\widetilde{e}_v$ if $\sigma'(C_1) = \sigma'(C_2)$ and has local weight z otherwise.

This means

$$\operatorname{Holant}\left(G; \begin{bmatrix} w&0&0&0\\ 0&0&z&0\\ 0&z&0&0\\ 0&0&0&w \end{bmatrix}\right) = z^{|V(G)|} \cdot Z_{\textsc{Ising}}\left(\widetilde{G}; \frac{w}{z}\right).$$

Next we show that $\textsc{Ising}\left(\frac{w}{z}\right)$ can be expressed as $\operatorname{Holant}\left(\begin{bmatrix} w&0&0&0\\ 0&0&z&0\\ 0&z&0&0\\ 0&0&0&w \end{bmatrix}\right)$. Note that every graph $G = (V, E)$ (without isolated vertices) is the crossing-circuit graph of some 4-regular graph $\overline{G}$. To define $\overline{G}$ from G, one only needs to do the following: (1) transform each vertex $v \in V$ into a closed cycle C_v; (2) for each loop at $v \in V$, make a self-intersection on C_v; and (3) for each non-loop edge $\{u, v\} \in E$ (u and v are two distinct vertices), make C_u and C_v

intersect in a "crossing" way at a vertex in $\overline{G}$ (by first creating a vertex p on C_u and another vertex p' on C_v, then merging p and p' with local labeling $1,3$ on C_u and $2,4$ on C_v). Then the above proof holds for the reverse direction. ◀

▶ **Proposition 11** ([11, Lemma 7]). *Suppose $\beta < -1$. Then $\#\mathrm{PM} \equiv_{\mathrm{AP}} \textsc{Ising}(\beta)$.*

▶ **Corollary 12.** *Suppose $x \neq 0$ and $\frac{w}{x} < -1$. Then $\#\mathrm{PM} \equiv_{\mathrm{AP}} \mathrm{Holant}\left(\left[\begin{smallmatrix} w & 0 & 0 & x \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ x & 0 & 0 & w \end{smallmatrix}\right]\right)$.*

When $a+d > b+c$ we have $\frac{a+b+c+d}{-a+b+c-d} < -1$, so by Corollary 12, Lemma 8, and Lemma 7,

$$\begin{aligned}
\#\mathrm{PM} &\equiv_{\mathrm{AP}} \mathrm{Holant}\left(\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix}\right) \\
&\leq_{\mathrm{AP}} \mathrm{Holant}\left(\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & a-b+c-d & a+b-c-d & 0 \\ 0 & a+b-c-d & a-b+c-d & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix}\right) \\
&\equiv_{\mathrm{AP}} \#\mathrm{EV}(a,b,c,d).
\end{aligned}$$

By the symmetry of a, b, c, we have proved Theorem 2.

4 #PerfectMatchings-easiness

In this section, we address two problems:

1. What are the signatures that can be realized by 4-ary matchgates (Definition 13)? Although the set of signatures that can be realized by planar matchgates with complex edge weights have been completely characterized [5], the set of signatures that can be realized by general (not necessarily planar) matchgates with nonnegative real edge weights is not fully understood, even for matchgates of arity 4. This type of matchgates plays a crucial role in the study of the approximate complexity of counting problems [18, 4], as we will see in this paper.
   In Theorem 15, we give a complete characterization of signatures of arity 4 that can be realized by matchgates with nonnegative real edges. Our method is primarily geometric.
2. Theorem 2 shows that for positive parameters $(a,b,c,d) \notin \mathcal{Y}$ the problem $\#\mathrm{EV}(a,b,c,d)$ is at least as hard as counting perfect matchings approximately. Here we ask the reverse question: For what parameter settings (a,b,c,d) does $\#\mathrm{EV}(a,b,c,d) \leq_{\mathrm{AP}} \#\mathrm{PM}$?
   We know that

   $$Z_{\mathrm{EV}}(G; a,b,c,d) = \mathrm{Holant}\left(G'; \neq_2 \mid f\right),$$

   where f is the 4-ary signature with $M(f) = \left[\begin{smallmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{smallmatrix}\right]$. Considering the fact that $(\neq_2)$ can be easily realized by a matchgate (a vertex with two dangling edges), Theorem 4 is a direct consequence of Lemma 17 which says that any signature in $\mathcal{Z}$ is realizable by some *4-ary matchgate* of constant size (with nonnegative edge weights, but not necessarily planar) (see Definition 13). Our theorem works for the eight-vertex model with parameter settings $\mathcal{S}^{\mathrm{E}}_{\leq 2}$ (defined below) not necessarily satisfying the arrow reversal symmetry.
   Moreover, Lemma 19 indicates that our result is tight in the sense that $\mathcal{S}^{\mathrm{E}}_{\leq 2}$ captures precisely the set of all signatures that can be realized by 4-ary matchgates (with even support, i.e., nonzero only on inputs of even Hamming weight). A similar statement holds for $\mathcal{S}^{\mathrm{O}}_{\leq 2}$. the corresponding set with odd support.

▶ **Definition 13.** *We use the term* a k-ary matchgate *to denote a graph* Γ *having* k *"dangling" edges, labelled* $i_1, \ldots, i_k$*. Each dangling edge has weight* 1 *and each non-dangling edge* e *is equipped with a nonnegative weight* w_e*. A configuration is a* $0,1$*-assignment to the edges. A configuration is a perfect matching if every vertex has exactly one incident edge assigned* 1*. A* k*-ary matchgate implements the signature* f*, where* $f(b_1, \ldots, b_k)$ *for* $(b_1, \ldots, b_k) \in \{0,1\}^k$ *is the sum, over perfect matchings, of the product of the weight of edges with assignment* 1*, where the dangling edge* i_j *is assigned* b_j*, and the empty product has weight* 1*.*

▶ Remark 14. Contrary to Definition 13 which does not require planarity, planar matchgates with complex edge weights has been completely characterized [21, 5]. As computing the weighted sum of perfect matchings is in polynomial time over planar graphs by the *FKT algorithm* [20, 15, 16], problems that can be locally expressed by planar matchgates are tractable over planar graphs.

▶ **Notation.**

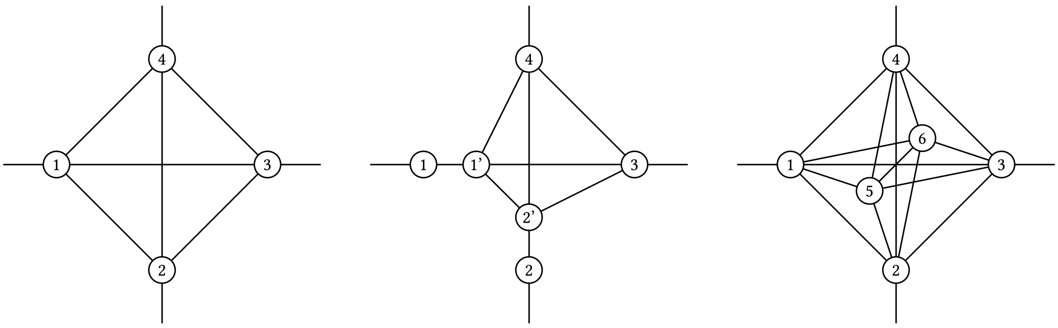
$$\mathcal{S}^{\mathrm{E}}_{\leq^2} = \{f \mid M(f) = \begin{bmatrix} d_1 & 0 & 0 & a_1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ a_2 & 0 & 0 & d_2 \end{bmatrix} \textit{ satisfying } \begin{cases} a_1a_2 \leq b_1b_2+c_1c_2+d_1d_2 \\ b_1b_2 \leq a_1a_2+c_1c_2+d_1d_2 \\ c_1c_2 \leq a_1a_2+b_1b_2+d_1d_2 \\ d_1d_2 \leq a_1a_2+b_1b_2+c_1c_2 \end{cases}, a_1, \cdots, d_2 \geq 0.\},$$

$$\mathcal{S}^{\mathrm{O}}_{\leq^2} = \{f \mid M(f) = \begin{bmatrix} 0 & d_1 & a_1 & 0 \\ b_1 & 0 & 0 & c_1 \\ c_2 & 0 & 0 & b_2 \\ 0 & a_2 & d_2 & 0 \end{bmatrix} \textit{ satisfying } \begin{cases} a_1a_2 \leq b_1b_2+c_1c_2+d_1d_2 \\ b_1b_2 \leq a_1a_2+c_1c_2+d_1d_2 \\ c_1c_2 \leq a_1a_2+b_1b_2+d_1d_2 \\ d_1d_2 \leq a_1a_2+b_1b_2+c_1c_2 \end{cases}, a_1, \cdots, d_2 \geq 0.\}.$$

▶ **Theorem 15.** *Denote by* $\mathcal{M}$ *the set of signatures that can be realized by 4-ary matchgates. Then* $\mathcal{M} = \mathcal{S}^{\mathrm{E}}_{\leq^2} \bigcup \mathcal{S}^{\mathrm{O}}_{\leq^2}$*.*

▶ Remark 16. Note that any signature in $\mathcal{M}$ must satisfy either *even parity* (nonzero only on inputs of even Hamming weight) or *odd parity* (nonzero only on inputs of odd Hamming weight). Theorem 15 for the even parity part ($\mathcal{S}^{\mathrm{E}}_{\leq^2}$) is a combination of Lemma 17 and Lemma 19. The odd parity part can be proved similarly.

▶ **Lemma 17.** *Suppose* $f \in \mathcal{S}^{\mathrm{E}}_{\leq^2}$*. Then there is a 4-ary matchgate of constant size whose signature is* f*.*



(a) **(b)** **(c)**

Figure 4 4-ary matchgates.

Proof. We first note that if any of the four inequalities in the definition of $\mathcal{S}_{\leq^2}$ is an equality, then the remaining three inequalities automatically hold, since the 8 values $a_1, \ldots, d_2$ are all nonnegative.

Given a signature $\begin{bmatrix} d_1 & 0 & 0 & a_1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ a_2 & 0 & 0 & d_2 \end{bmatrix}$, first we construct a matchgate for $d_1 d_2 = a_1 a_2 + b_1 b_2 + c_1 c_2$. If $d_1 d_2 = 0$ then all four products $a_1 a_2 = b_2 b_2 = c_1 c_2 = d_1 d_2 = 0$, and one can easily adapt from the following proof to show that the signature is realizable as a matchgate signature. So it suffices to implement the normalized version $\begin{bmatrix} a_1a_2+b_1b_2+c_1c_2 & 0 & 0 & a_1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ a_2 & 0 & 0 & 1 \end{bmatrix}$. Our construction is a weighted K_4 depicted in Figure 4a. Let e_1, e_2, e_3, e_4 be the dangling edges incident to vertices $1, 2, 3, 4$, respectively. Denote by w_{ij} the weight on the edge between vertex i and vertex j. One can check that the following weight assignment meets our need: $w_{12} = a_1, w_{34} = a_2, w_{14} = b_1, w_{23} = b_2, w_{13} = c_1, w_{24} = c_2$.

For $a_1 a_2 = b_1 b_2 + c_1 c_2 + d_1 d_2$, without loss of generality we assume $a_1 a_2 \neq 0$ and we normalize $a_1 = 1$. Then our construction is shown in Figure 4b where we set $w_{11'} = 1, w_{22'} = 1, w_{1'2'} = d_2, w_{34} = d_1, w_{1'4} = c_2, w_{2'3} = c_1, w_{1'3} = b_2, w_{2'4} = b_1$. One can verify that it realizes the normalized signature $\begin{bmatrix} d_1 & 0 & 0 & 1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ b_1b_2+c_1c_2+d_1d_2 & 0 & 0 & d_2 \end{bmatrix}$. The construction for $b_1 b_2 = a_1 a_2 + c_1 c_2 + d_1 d_2$ and $c_1 c_2 = a_1 a_2 + b_1 b_2 + d_1 d_2$ are symmetric to the above case.

It remains to show that the interior

$$\begin{cases} a_1a_2 & < & b_1b_2 + c_1c_2 + d_1d_2 \\ b_1b_2 & < & a_1a_2 + c_1c_2 + d_1d_2 \\ c_1c_2 & < & a_1a_2 + b_1b_2 + d_1d_2 \\ d_1d_2 & < & a_1a_2 + b_1b_2 + c_1c_2 \end{cases} \tag{4}$$

can all be reached. We first deal with the case when all eight parameters are strictly positive and leave the other cases to the end of this proof. We use a weighted K_6 to be our matchgate depicted in Figure 4c, and set $w_{12} = r_1, w_{34} = r_2, w_{14} = s_1, w_{23} = s_2, w_{13} = t_1, w_{24} = t_2, w_{15} = p_1, w_{25} = p_2, w_{35} = p_3, w_{45} = p_4, w_{16} = q_1, w_{26} = q_2, w_{36} = q_3, w_{46} = q_4, w_{56} = 1$. Then the matchgate has a singature with the following parameters

$$\begin{aligned} a_1' &= r_1 + p_1q_2 + p_2q_1, & a_2' &= r_2 + p_3q_4 + p_4q_3, \\ b_1' &= s_1 + p_1q_4 + p_4q_1, & b_2' &= s_2 + p_2q_3 + p_3q_2, \\ c_1' &= t_1 + p_1q_3 + p_3q_1, & c_2' &= t_2 + p_2q_4 + p_4q_2, \\ d_1' &= (r_1r_2 + s_1s_2 + t_1t_2) + & d_2' &= 1, \\ &\quad (p_3q_4 + p_4q_3)r_1 + (p_1q_2 + p_2q_1)r_2 + \\ &\quad (p_2q_3 + p_3q_2)s_1 + (p_1q_4 + p_4q_1)s_2 + \\ &\quad (p_2q_4 + p_4q_2)t_1 + (p_1q_3 + p_3q_1)t_2. \end{aligned}$$

Note that all the edge weights have to be nonnegative. By properly setting the edge weights in the matchgate, we show that we can achieve any relative ratios among the eight given positive values $a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2$ that satisfy (4). Our first step is to achieve any relative ratios among the four product values a_1a_2, b_1b_2, c_1c_2, d_1d_2 satisfying (4); and the second step is to adjust the relative ratio within the pairs $\{a_1, a_2\}$, $\{b_1, b_2\}$, $\{d_1, d_2\}$ and $\{c_1, c_2\}$ without affecting the product values. This can be justified by the observation that, by a scaling a global positive constant can be easily achieved, and all appearances of a_1 and a_2 in (4) are as a product a_1a_2, and similarly for $b_1, b_2, c_1, c_2, d_1, d_2$.

For the fourteen edge weights $r_1, \ldots, q_4$ to be determined, let

$$\begin{cases} A' = p_1p_2q_3q_4 + p_3p_4q_1q_2, & R = r_1r_2 + r_1(p_3q_4 + p_4q_3) + r_2(p_1q_2 + p_2q_1), \\ B' = p_1p_4q_2q_3 + p_2p_3q_1q_4, & S = s_1s_2 + s_1(p_2q_3 + p_3q_2) + s_2(p_1q_4 + p_4q_1), \\ C' = p_1p_3q_2q_4 + p_2p_4q_1q_3, & T = t_1t_2 + t_1(p_2q_4 + p_4q_2) + t_2(p_1q_3 + p_3q_1), \end{cases} \tag{5}$$

and define

$$\begin{cases} A &= & A' + S + T \\ B &= & B' + R + T \\ C &= & C' + R + S \\ D &= & A' + B' + C'. \end{cases} \tag{6}$$

Note that A', B', C', R, S, T are all nonnegative and so are A, B, C, D.

Our goal is to choose the fourteen edge weights $r_1, \ldots, q_4$ so that A, B, C, D are all positive and satisfy

$$\begin{cases} A &= \frac{1}{2}(b_1b_2 + c_1c_2 + d_1d_2 - a_1a_2) \\ B &= \frac{1}{2}(a_1a_2 + c_1c_2 + d_1d_2 - b_1b_2) \\ C &= \frac{1}{2}(a_1a_2 + b_1b_2 + d_1d_2 - c_1c_2) \\ D &= \frac{1}{2}(a_1a_2 + b_1b_2 + c_1c_2 - d_1d_2). \end{cases} \tag{7}$$

Note that, by definition, the left-side of (7) is precisely the right-side of (7) when $a_1, \ldots, d_2$ are replaced by $a'_1, \ldots, d'_2$ respectively. Denote the products $a_1a_2, b_1b_2, c_1c_2, d_1d_2$ by a^{**}, b^{**}, c^{**}, d^{**} respectively. Then (7) is a set of four linear equations $M \cdot \begin{bmatrix} a^{**} \\ b^{**} \\ c^{**} \\ d^{**} \end{bmatrix} = \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix}$, where $M = \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$. Note that M is invertible and $M^{-1} = M$, so (7) is equivalent to $M \cdot \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} a^{**} \\ b^{**} \\ c^{**} \\ d^{**} \end{bmatrix}$, having an identical form. Since the requirement (4) in terms of $a^{**}, b^{**}, c^{**}, d^{**}$ translates into the requirement A, B, C, D being strictly positive via M, it is not surprising that the requirement $a^{**}, b^{**}, c^{**}, d^{**}$ being strictly positive translates into the requirement

$$\begin{cases} A < B + C + D \\ B < A + C + D \\ C < A + B + D \\ D < A + B + C, \end{cases} \tag{8}$$

and that A, B, C, D are positive is the same as (4).

Furthermore, let $\begin{cases} X=S+T \\ Y=R+T \\ Z=R+S \end{cases}$, then the requirement R, S, T being positive is equivalent to the requirement $\begin{cases} Y+Z>X \\ X+Z>Y \\ X+Y>Z \end{cases}$. This is because $\begin{bmatrix} 0&1&1 \\ 1&0&1 \\ 1&1&0 \end{bmatrix} \cdot \begin{bmatrix} R \\ S \\ T \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$ is the same as $\frac{1}{2}\begin{bmatrix} -1&1&1 \\ 1&-1&1 \\ 1&1&-1 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} R \\ S \\ T \end{bmatrix}$.

The crucial ingredient of our proof is a *geometric* lemma in 3-dimensional space. Suppose $a^{**}, b^{**}, c^{**}, d^{**}$ are positive and they satisfy (4). This defines $\begin{bmatrix} \tilde{A} \\ \tilde{B} \\ \tilde{C} \\ \tilde{D} \end{bmatrix} = M \cdot \begin{bmatrix} a^{**} \\ b^{**} \\ c^{**} \\ d^{**} \end{bmatrix}$. By a scaling we may assume $\tilde{D} = 1$. Hence $(\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ are positive and satisfy (8). Thus $(\tilde{A}, \tilde{B}, \tilde{C})$ belongs to the set U in the statement of Lemma 18.

By Lemma 18, there exist (strictly) positive tuples $(\tilde{A}', \tilde{B}', \tilde{C}')$ and $(\tilde{X}, \tilde{Y}, \tilde{Z})$ such that

$$(\tilde{A}, \tilde{B}, \tilde{C}) = (\tilde{A}', \tilde{B}', \tilde{C}') + (\tilde{X}, \tilde{Y}, \tilde{Z}),$$

satisfying $\tilde{A}' + \tilde{B}' + \tilde{C}' = 1$ and $\begin{cases} \tilde{Y}+\tilde{Z}>\tilde{X} \\ \tilde{X}+\tilde{Z}>\tilde{Y} \\ \tilde{X}+\tilde{Y}>\tilde{Z} \end{cases}$. By the previous observation this indicates that there exist (strictly) positive $\tilde{A}', \tilde{B}', \tilde{C}', \tilde{R}, \tilde{S}, \tilde{T}$ such that $\begin{cases} \tilde{A}=\tilde{A}'+\tilde{S}+\tilde{T} \\ \tilde{B}=\tilde{B}'+\tilde{R}+\tilde{T} \\ \tilde{C}=\tilde{C}'+\tilde{R}+\tilde{S} \\ \tilde{D}=\tilde{A}'+\tilde{B}'+\tilde{C}' \end{cases}$.

We first set p_i, q_i $(1 \le i \le 4)$ so that $(A', B', C') = c \cdot (\tilde{A}', \tilde{B}', \tilde{C}')$ for some constant c. To achieve this, set $q_1 = q_2 = q_3 = q_4 = 1$, and let o_1, o_2, o_3 be positive, and then set $p_1 = \sqrt{\frac{o_2 o_3}{o_1}}$, $p_2 = \sqrt{\frac{o_3 o_1}{o_2}}$, $p_3 = \sqrt{\frac{o_1 o_2}{o_3}}$, and $p_4 = \frac{1}{p_1 p_2 p_3}$. We have $p_1 p_2 p_3 p_4 = 1$, and $p_2 p_3 = o_1, p_3 p_1 = o_2, p_1 p_2 = o_3$. Then set $\begin{cases} A' = p_1 p_2 + \frac{1}{p_1 p_2} = o_3 + \frac{1}{o_3} \\ B' = p_2 p_3 + \frac{1}{p_2 p_3} = o_1 + \frac{1}{o_1} \\ C' = p_3 p_1 + \frac{1}{p_3 p_1} = o_2 + \frac{1}{o_2} \end{cases}$, which can be independently any positive numbers at least 2, by choosing o_1, o_2, o_3 to be suitable positive numbers. This allows us to get A', B', C' such that $(A', B', C') = c \cdot (\tilde{A}', \tilde{B}', \tilde{C}')$ for some constant c. Then it is obvious that $r_1, r_2, s_1, s_2, t_1, t_2$ can be set so that $(R, S, T) = c \cdot (\tilde{R}, \tilde{S}, \tilde{T})$. Compute A, B, C, D according to (6). As a consequence, $(A, B, C, D) = c \cdot (\tilde{A}, \tilde{B}, \tilde{C}, \tilde{D})$ is a valid solution.

To adjust the relative ratio between $\{d_1, d_2\}$, say increasing $\frac{d_2}{d_1}$ by δ, while keeping all product values and the relative ratios within the other three pairs, just increase $r_1, r_2, s_1, s_2, t_1, t_2$ by $\delta^{1/2}$ and increase p_i, q_i $(1 \le i \le 4)$ by $\delta^{1/4}$. Similarly, to increase $\frac{a_2}{a_1}$ by δ alone without affecting the other products and ratios, just increase r_2 by $\delta^{1/2}$ and p_3, p_4, q_3, q_4 by $\delta^{1/4}$, and decrease r_1 by $\delta^{1/2}$ and p_1, p_2, q_1, q_2 by $\delta^{1/4}$. The other cases are symmetric.

Finally we deal with the cases when there are zeros among the eight parameters. Note that at most one of $a_1 a_2, b_1 b_2, c_1 c_2, d_1 d_2$ is zero, because if at least two products are zero, say $a_1 a_2 = b_1 b_2 = 0$, then (4) forces a contradiction that $c_1 c_2 < d_1 d_2$ and $d_1 d_2 < c_1 c_2$. In the case $d_1 d_2 = 0$:

- $d_1 = 0, d_2 \neq 0$: We make the modification that $w_{12} = w_{34} = w_{14} = w_{23} = w_{13} = w_{24} = 0$, i.e. $r_1, r_2, s_1, s_2, t_1, t_2 = 0$.
- $d_1 = d_2 = 0$: We make the further modification that $w_{56} = 0$.
- $d_1 \neq 0, d_2 = 0$: We connect the four dangling edges in Figure 4c to four degree 2 vertices, respectively. This switches the role of d_1 and d_2 in the previous proof.

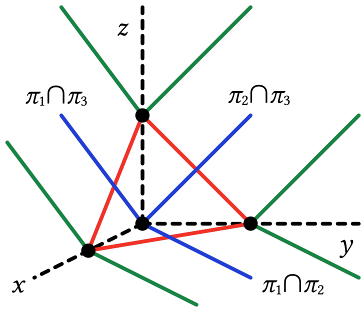
One can check our proof is still valid in the above three cases. If $a_1 a_2 = 0$, then we connect the dangling edges on vertices 1, 2 to two degree 2 vertices (similar to the operation from Figure 4a to Figure 4b). This switches the role of d_1, d_2 with a_2, a_1 and the proof folllows. The proofs for $b_1 b_2 = 0$ and $c_1 c_2 = 0$ are symmetric. ◀

Now we give the crucial geometric lemma.

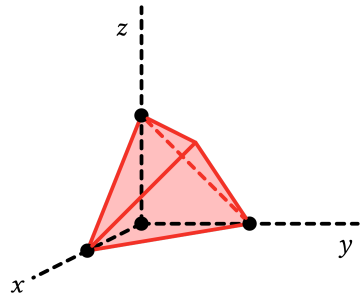
▶ **Lemma 18.** *Let* $U = \{(x, y, z) \in \mathbb{R}^3_{>0} \mid x < y+z+1, y < x+z+1, z < x+y+1, 1 < x+y+z\}$, $V = \{(x, y, z) \in \mathbb{R}^3_{>0} \mid x + y + z = 1\}$, *and* $W = \{(x, y, z) \in \mathbb{R}^3_{>0} \mid y + z > x, x + z > y, x + y > z\}$. *Then U is the Minkowski sum of V and W, namely, U consists of precisely those points $\boldsymbol{u} \in \mathbb{R}^3$, such that $\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w}$ for some $\boldsymbol{v} \in V$ and $\boldsymbol{w} \in W$. The same statement is true for the closures of U, V and W (in the topology of Euclidean space).*

Proof. Observe that U, V, and W are the interiors of a polyhedron with 7 facets, a regular triangle, and a polyhedron with 3 facets, respectively.

The polyhedron for W is the intersection of three half spaces bounded by three planes, $\begin{cases} (\pi_1): y+z \ge x \\ (\pi_2): x+z \ge y \\ (\pi_3): x+y \ge z \end{cases}$, where the planes are defined by equalities. Note that these inequalities imply that $x, y, z \ge 0$, thus this polyhedron has only three facets. We can find the intersection of each pair of the three planes for W as $\begin{cases} \pi_1 \cap \pi_2 : x = y \ge 0, z = 0 \\ \pi_1 \cap \pi_3 : x = z \ge 0, y = 0 \\ \pi_2 \cap \pi_3 : y = z \ge 0, x = 0 \end{cases}$. Note that these intersections lie on the planes $z = 0$, $y = 0$, and $x = 0$, respectively.



(a) The blue rays are the intersections of the three facets for W. The red triangle is the boundary for V. The green rays together with the red triangle are the intersections of the seven facets for U.



(b) The tetrahedron in U that is left uncovered by sliding W along the boundary of V, but is covered by the rays from the simplex V in the direction of $(1, 1, 1)$.

Figure 5

Let $\boldsymbol{e}_1 = (1, 0, 0), \boldsymbol{e}_2 = (0, 1, 0), \boldsymbol{e}_3 = (0, 0, 1)$. Then the triangle for V is just the convex hull of $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$. Suppose we shift the origin of W from $\mathbf{0}$ to $\boldsymbol{e}_1$ and denote the resulting (interior of a) polyhedron by $W^{\boldsymbol{e}_1}$, then we have the defining inequalities $\begin{cases} (\pi_1^{\boldsymbol{e}_1}): y+z \geq (x-1) \\ (\pi_2^{\boldsymbol{e}_1}): (x-1)+z \geq y \\ (\pi_3^{\boldsymbol{e}_1}): (x-1)+y \geq z \end{cases}$, where the shifted planes are defined by the corresponding equalities. By symmetry, if we shift the origin of W to $\boldsymbol{e}_2$ and to $\boldsymbol{e}_3$, we have respectively $W^{\boldsymbol{e}_2}$ with $\begin{cases} (\pi_1^{\boldsymbol{e}_2}): (y-1)+z \geq x \\ (\pi_2^{\boldsymbol{e}_2}): x+z \geq (y-1) \\ (\pi_3^{\boldsymbol{e}_2}): x+(y-1) \geq z \end{cases}$, and $W^{\boldsymbol{e}_3}$ with $\begin{cases} (\pi_1^{\boldsymbol{e}_3}): y+(z-1) \geq x \\ (\pi_2^{\boldsymbol{e}_3}): x+(z-1) \geq y \\ (\pi_3^{\boldsymbol{e}_3}): x+y \geq (z-1) \end{cases}$. Note that the shifted planes $\pi_1^{\boldsymbol{e}_1}$, $\pi_2^{\boldsymbol{e}_2}$, and $\pi_3^{\boldsymbol{e}_3}$ contain three distinct facets of U, and they coincide exactly with a facet of $W^{\boldsymbol{e}_1}$, $W^{\boldsymbol{e}_2}$, and $W^{\boldsymbol{e}_3}$, respectively.

By sliding W with its origin along the line $x + y = 1, z = 0$ from $\boldsymbol{e}_1$ to $\boldsymbol{e}_2$, we have a partial coverage of U by the shifting copies of W from $W^{\boldsymbol{e}_1}$ and $W^{\boldsymbol{e}_2}$:

- The shifted ray of $\pi_1 \cap \pi_2 : x = y \geq 0, z = 0$ moves from $\pi_1^{\boldsymbol{e}_1} \cap \pi_2^{\boldsymbol{e}_1} : (x - 1) = y \geq 0, z = 0$ to $\pi_1^{\boldsymbol{e}_2} \cap \pi_2^{\boldsymbol{e}_2} : x = (y - 1) \geq 0, z = 0$. Notice that this is a parallel transport, and stays on the plane $z = 0$, and thus it swipes another facet of U on $z = 0$ bounded by the two lines $x - y = -1$ and $x - y = 1$.
- The shifted ray of $\pi_1 \cap \pi_3 : x = z \geq 0, y = 0$ moves from $\pi_1^{\boldsymbol{e}_1} \cap \pi_3^{\boldsymbol{e}_1} : (x - 1) = z \geq 0, y = 0$ to $\pi_1^{\boldsymbol{e}_2} \cap \pi_3^{\boldsymbol{e}_2} : x = z \geq 0, (y - 1) = 0$; the shifted ray of $\pi_2 \cap \pi_3 : y = z \geq 0, x = 0$ moves from $\pi_2^{\boldsymbol{e}_1} \cap \pi_3^{\boldsymbol{e}_1} : y = z \geq 0, (x - 1) = 0$ to $\pi_2^{\boldsymbol{e}_2} \cap \pi_3^{\boldsymbol{e}_2} : (y - 1) = z \geq 0, x = 0$. Notice that both stay on the plane $x + y - z = 1$ which is $\pi_3^{\boldsymbol{e}_1} = \pi_3^{\boldsymbol{e}_2}$.

It follows that the part of U satisfying $x + y - z > 1$ is covered by the Minkowski sum of W and the line segment on $x + y = 1, z = 0$ from $\boldsymbol{e}_1$ to $\boldsymbol{e}_2$ (which is a side of the triangle V).

Symmetrically, after sliding W with its origin from $\boldsymbol{e}_2$ to $\boldsymbol{e}_3$ along the line $y+z = 1, x = 0$ we get the parallel tranport from $W^{\boldsymbol{e}_2}$ to $W^{\boldsymbol{e}_3}$. Also after sliding W with its origin from $\boldsymbol{e}_3$ back to $\boldsymbol{e}_1$ along the line segment $x + z = 1, y = 0$ we get the parallel tranport from $W^{\boldsymbol{e}_3}$ to $W^{\boldsymbol{e}_1}$. After these, the only subset in U that is left uncovered by shifting copies of W is $U \cap \{(x, y, z) \mid \begin{cases} -x+y+z \leq 1 \\ x-y+z \leq 1 \\ x+y-z \leq 1 \end{cases} \} = \{(x, y, z) \in \mathbb{R}^3_{>0} \mid \begin{cases} x+y+z \geq 1 \\ -x+y+z \leq 1 \\ x-y+z \leq 1 \\ x+y-z \leq 1 \end{cases} \}$ – a *tetrahedron* (Figure 5b). However this subset can be covered by the rays $\{\boldsymbol{v} + \lambda(1, 1, 1) \mid \boldsymbol{v} \in V, \lambda > 0\}$. Note that $\lambda(1, 1, 1) \in W$ for all $\lambda > 0$.

Finally regarding the closures $\overline{U}, \overline{V}$ and $\overline{W}$, for $v_n \in V$ and $w_n \in W$, if $v_n \to v \in \overline{V}$ and $w_n \to w \in \overline{W}$, then $u_n = v_n + w_n \in U$, and $u_n \to v + w$. So $v + w \in \overline{U}$. Conversely, if $u_n \to u \in \overline{U}$, where $u_n \in U$, then $u_n = v_n + w_n$ for some $v_n \in V$ and $w_n \in W$. As V is bounded, there is a convergent subsequence $\{v_{n_k}\}$, such that $v = \lim_{k\to\infty} v_{n_k} \in \overline{V}$. Then $w_{n_k} = u_{n_k} - v_{n_k}$ also converges to some $w \in \overline{W}$, and then $u = \lim_{k\to\infty}(v_{n_k} + w_{n_k}) = v + w$, is a sum of points from $\overline{V}$ and $\overline{W}$.

This completes the proof. ◀

▶ **Lemma 19.** *Suppose f is the signature of a 4-ary matchgate with $M(f) = \begin{bmatrix} d_1 & 0 & 0 & a_1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ a_2 & 0 & 0 & d_2 \end{bmatrix}$. Then $f \in \mathcal{S}^{\mathrm{E}}_{\leq 2}$. In particular, if f satisfies arrow reversal symmetry, $f \in \mathcal{Z}$.*

▶ Remark 20. The last part $d_1d_2 \leq a_1a_2 + b_1b_2 + c_1c_2$ was proved in [4, Lemma 56]. The proofs for other three parts are symmetric and similar to the proof for the last part. For completeness, here we give the proof for the first part $a_1a_2 \leq b_1b_2 + c_1c_2 + d_1d_2$.

Proof. Consider a 4-ary matchgate Γ with signature f. Given that $M(f) = \begin{bmatrix} d_1 & 0 & 0 & a_1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ a_2 & 0 & 0 & d_2 \end{bmatrix}$, $a_1a_2 \leq b_1b_2 + c_1c_2 + d_1d_2$ is equivalent as

$$f(0011)f(1100) \leq f(0110)f(1001) + f(0101)f(1010) + f(0000)f(1111). \tag{9}$$

Let $I = \{i_1, i_2, i_3, i_4\}$ be the set of dangling edges of Γ. For $X \subseteq I$, let M_X denote the set of perfect matchings that include dangling edges in X (by assigning them 1) and exclude dangling edges in $I \setminus X$ (by assigning them 0). We exhibit an injective map

$$\mu : M_{\{i_1,i_2\}} \times M_{\{i_3,i_4\}} \to [M_{\{i_2,i_3\}} \times M_{\{i_1,i_4\}}] \bigcup [M_{\{i_2,i_4\}} \times M_{\{i_1,i_3\}}] \bigcup [M_\emptyset \times M_I]$$

which is weight-preserving in the sense that for matchings m_1, m_2, m_3, m_4 with $\mu(m_1, m_2) = (m_3, m_4)$, we have $w(m_1)w(m_2) = w(m_3)w(m_4)$. The existence of μ implies (9).

Given $(m_1, m_2) \in M_{\{i_1,i_2\}} \times M_{\{i_3,i_4\}}$, consider $m_1 \oplus m_2$ and note that this is a collection of cycles together with two paths. Let π be the path connecting the dangling edge i_1 to some other dangling edge; let π' be the path connecting the remaining two dangling edges. Let $m_3 := m_1 \oplus \pi$ and $m_4 := m_2 \oplus \pi$. Then we have the following

- If π connects i_1 to i_2, then $m_3 \in M_\emptyset$ and $m_4 \in M_I$;
- If π connects i_1 to i_3, then $m_3 \in M_{\{i_2,i_3\}}$ and $m_4 \in M_{\{i_1,i_4\}}$;
- If π connects i_1 to i_4, then $m_3 \in M_{\{i_2,i_4\}}$ and $m_4 \in M_{\{i_1,i_3\}}$.

The construction is invertible, since if (m_3, m_4) is in the image of the above mapping, then $m_3 \oplus m_4 = m_1 \oplus m_2$. From $m_1 \oplus m_2$, we can recover π (as the unique path that connects i_1 to one of the other dangling edges in $\{i_2, i_3, i_4\}$). Then we can recover m_1 and m_2 as $m_3 \oplus \pi$ and $m_4 \oplus \pi$ respectively. Therefore, $\mu : (m_1, m_2) \to (m_3, m_4)$ is an injection.

To see that μ is weight-preserving, observe that the each of the edges in π appears in exactly one of m_1 and m_2 and in exactly one of m_3 and m_4 and that $m_i \setminus \pi = m_{i+2} \setminus \pi$ for $i \in \{1, 2\}$. Hence,

$$w(m_1)w(m_2) = \prod_{e\in m_1\setminus\pi} w_e \cdot \prod_{e\in m_2\setminus\pi} w_e \cdot \prod_{e\in\pi} w_e = \prod_{e\in m_3\setminus\pi} w_e \cdot \prod_{e\in m_4\setminus\pi} w_e \cdot \prod_{e\in\pi} w_e = w(m_3)w(m_4).$$

◀

References

1 R. J. Baxter. Eight-vertex model in lattice statistics. *Phys. Rev. Lett.*, 26:832–833, April 1971. `doi:10.1103/PhysRevLett.26.832`.

2 R. J. Baxter. Partition function of the eight-vertex lattice model. *Annals of Physics*, 70(1):193–228, 1972. `doi:10.1016/0003-4916(72)90335-1`.

3 R. J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Academic Press Inc., San Diego, CA, USA, 1982.

4 Andrei Bulatov, Leslie Ann Goldberg, Mark Jerrum, David Richerby, and Stanislav Živný. Functional clones and expressibility of partition functions. *Theoretical Computer Science*, 687:11–39, 2017. doi:10.1016/j.tcs.2017.05.001.

5 Jin-Yi Cai and Xi Chen. *Complexity Dichotomies for Counting Problems*, volume 1. Cambridge University Press, 2017. doi:10.1017/9781107477063.

6 Jin-Yi Cai and Zhiguo Fu. Complexity classification of the eight-vertex model. *CoRR*, abs/1702.07938, 2017. arXiv:1702.07938.

7 Jin-Yi Cai, Tianyu Liu, and Pinyan Lu. Approximability of the six-vertex model. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '19, pages 2248–2261, 2019. doi:10.1137/1.9781611975482.136.

8 Jin-Yi Cai, Tianyu Liu, Pinyan Lu, and Jing Yu. Approximability of the eight-vertex model. *CoRR*, abs/1811.03126, 2018. arXiv:1811.03126.

9 Chungpeng Fan and F. Y. Wu. General lattice model of phase transitions. *Phys. Rev. B*, 2:723–733, August 1970. doi:10.1103/PhysRevB.2.723.

10 Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. *J. ACM*, 62(6), December 2015. doi:10.1145/2785964.

11 Leslie Ann Goldberg and Mark Jerrum. Inapproximability of the Tutte polynomial. *Information and Computation*, 206(7):908–929, 2008. doi:10.1016/j.ic.2008.04.003.

12 Sam Greenberg and Dana Randall. Slow mixing of Markov chains using fault lines and fat contours. *Algorithmica*, 58(4):911–927, December 2010. doi:10.1007/s00453-008-9246-3.

13 Thomas P. Hayes, Juan C. Vera, and Eric Vigoda. Randomly coloring planar graphs with fewer colors than the maximum degree. *Random Structures & Algorithms*, 47(4):731–759, 2015. doi:10.1002/rsa.20560.

14 Richard M. Karp and Michael Luby. Monte-Carlo algorithms for enumeration and reliability problems. In *Proceedings of the 24th Annual Symposium on Foundations of Computer Science*, SFCS '83, pages 56–64, Washington, DC, USA, 1983. IEEE Computer Society. doi:10.1109/SFCS.1983.35.

15 P.W. Kasteleyn. The statistics of dimers on a lattice: I. The number of dimer arrangements on a quadratic lattice. *Physica*, 27(12):1209–1225, 1961. doi:10.1016/0031-8914(61)90063-5.

16 P.W. Kasteleyn. Graph theory and crystal physics. In F. Harary, editor, *Graph Theory and Theoretical Physics*, pages 43–110. Academic Press, 1967.

17 Tianyu Liu. Torpid mixing of Markov chains for the six-vertex model on $\mathbb{Z}^2$. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, APPROX/RANDOM 2018, pages 52:1–52:15, 2018. doi:10.4230/LIPIcs.APPROX-RANDOM.2018.52.

18 Colin McQuillan. Approximating Holant problems by winding. *CoRR*, abs/1301.2880, 2013. arXiv:1301.2880.

19 Bill Sutherland. Two dimensional hydrogen bonded crystals without the ice rule. *Journal of Mathematical Physics*, 11(11):3183–3186, 1970. doi:10.1063/1.1665111.

20 H. N. V. Temperley and Michael E. Fisher. Dimer problem in statistical mechanics-an exact result. *The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics*, 6(68):1061–1063, 1961. doi:10.1080/14786436108243366.

21 Leslie Valiant. Quantum circuits that can be simulated classically in polynomial time. *SIAM Journal on Computing*, 31(4):1229–1254, 2002. doi:10.1137/S0097539700377025.

22 Leslie G. Valiant. Holographic algorithms. *SIAM J. Comput.*, 37(5):1565–1594, February 2008. doi:10.1137/070682575.