

Approximability of the Eight-Vertex Model

Jin-Yi Cai

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

Tianyu Liu

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

Pinyan Lu

Shanghai University of Finance and Economics, China
<http://itcs.shufe.edu.cn/pinyan/>
lu.pinyan@mail.shufe.edu.cn

Jing Yu

Georgia Institute of Technology, Atlanta, GA, USA
jingyu@gatech.edu

Abstract

We initiate a study of the classification of approximation complexity of the eight-vertex model defined over 4-regular graphs. The eight-vertex model, together with its special case the six-vertex model, is one of the most extensively studied models in statistical physics, and can be stated as a problem of counting weighted orientations in graph theory. Our result concerns the approximability of the partition function on all 4-regular graphs, classified according to the parameters of the model. Our complexity results conform to the phase transition phenomenon from physics.

We introduce a *quantum decomposition* of the eight-vertex model and prove a set of *closure properties* in various regions of the parameter space. Furthermore, we show that there are extra closure properties on 4-regular planar graphs. These regions of the parameter space are concordant with the phase transition threshold. Using these closure properties, we derive polynomial time approximation algorithms via *Markov chain Monte Carlo*. We also show that the eight-vertex model is NP-hard to approximate on the other side of the phase transition threshold.

2012 ACM Subject Classification Theory of computation → Design and analysis of algorithms

Keywords and phrases Approximate complexity, the eight-vertex model, Markov chain Monte Carlo

Digital Object Identifier 10.4230/LIPIcs.CCC.2020.4

Related Version A full version of the paper is available at <https://arxiv.org/abs/1811.03126>.

Funding *Jin-Yi Cai*: Supported by NSF CCF-1714275.

Tianyu Liu: Supported by NSF CCF-1714275.

1 Introduction

The eight-vertex model is one of the most important vertex models in statistical physics [2]. Given a 4-regular graph G , an *even orientation* assigns a direction to every edge such that the number of arrows into (and out of) each vertex is even. In the *unweighted* case, the problem is to count the number of even orientations of G , and this is computable in polynomial time. In general there are *weights* associated with local configurations, and the problem is to compute a weighted sum called the partition function. This becomes a challenging problem, and the complexity picture becomes more intricate [6].



© Jin-Yi Cai, Tianyu Liu, Pinyan Lu, and Jing Yu;
licensed under Creative Commons License CC-BY
35th Computational Complexity Conference (CCC 2020).

Editor: Shubhangi Saraf; Article No. 4; pp. 4:1–4:18

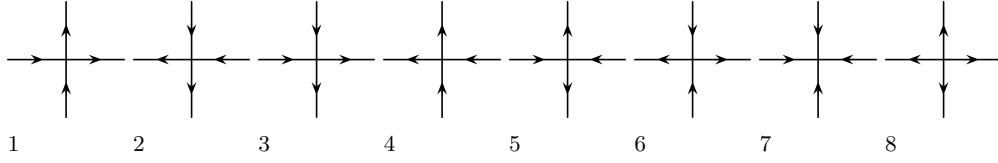
Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



4:2 Approximability of the Eight-Vertex Model



■ **Figure 1** Valid configurations of the eight-vertex model.

In physics, the eight-vertex model is defined on a square lattice region where each vertex of the lattice is connected by an edge to four nearest neighbors, with eight permitted local configurations (see Figure 1). They are associated with eight possible weights w_1, \dots, w_8 . In physics, it is typically assumed that the weight is unchanged if all arrows are flipped. 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 is called *arrow reversal symmetry*. In this paper, we make this assumption and further assume that $a, b, c, d \geq 0$, as 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

$$Z(G) = Z(G; a, b, c, d) = \sum_{\tau \in \mathcal{O}_e(G)} a^{n_1+n_2} b^{n_3+n_4} c^{n_5+n_6} d^{n_7+n_8}, \quad (1)$$

where $\mathcal{O}_e(G)$ 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) under $\tau \in \mathcal{O}_e(G)$. The famous *six-vertex model* is the special case $d = 0$, i.e., only Figure 1-1 to Figure 1-6 are allowed. In this case, states are *Eulerian orientations*. Further special cases include 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 it [2]. By any metric, these are among the most studied models in statistical physics.

As the problem is to compute the *partition function* $Z(G)$, naturally one should study its computational complexity. For the exact complexity, a dichotomy has been proved [6]. For most parameters, the problem is #P-hard. Regarding approximate complexity, to our best knowledge, there is only one previous result in this regard due to Greenberg and Randall [15]. They showed that on square lattice regions a specific Markov chain is torpidly mixing when d is large. It means that when sinks and sources have large weights, this particular chain cannot be used to approximately sample eight-vertex configurations on the square lattice according to the Gibbs measure.

In this paper we initiate a study toward a classification of the approximate complexity of $Z(G)$ on 4-regular graphs. Our results conform to the *order-disorder phase transitions* of the eight-vertex model in physics. (See the book [2] for more details; the full paper gives a brief description.)

To state our theorems and proofs, we adopt the following notations, for $a, b, c, d \in \mathbb{R}^+$.

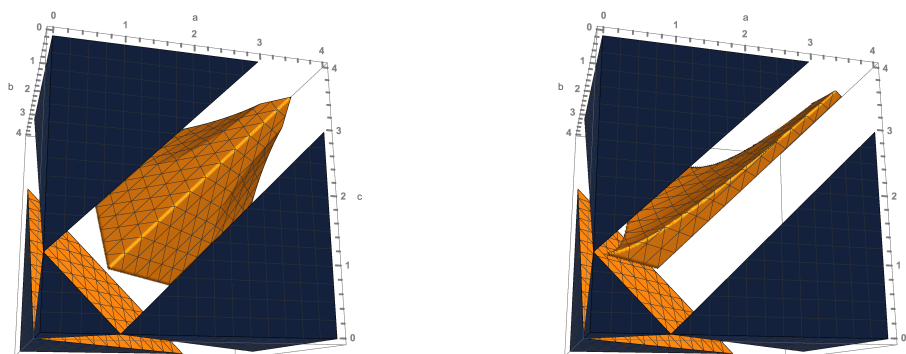
- $\mathcal{F}_{\leq 2} := \{(a, b, c, d) \mid a^2 \leq b^2 + c^2 + d^2, b^2 \leq a^2 + c^2 + d^2, c^2 \leq a^2 + b^2 + d^2, d^2 \leq a^2 + b^2 + c^2\}$;
- $\mathcal{F}_{>} := \{(a, b, c, d) \mid a > b + c + d \text{ or } b > a + c + d \text{ or } c > a + b + d \text{ or } d > a + b + c \text{ where at least two of } a, b, c, d > 0\}$;
- $\mathcal{A}_{\leq} := \{(a, b, c, d) \mid a + d \leq b + c\}$, $\mathcal{B}_{\leq} := \{(a, b, c, d) \mid b + d \leq a + c\}$, $\mathcal{C}_{\leq} := \{(a, b, c, d) \mid c + d \leq a + b\}$, $\mathcal{C}_{\geq} := \{(a, b, c, d) \mid c + d \geq a + b\}$, $\mathcal{C}_= := \{(a, b, c, d) \mid c + d = a + b\}$.

► **Remark 1.1.** We have $\mathcal{F}_{\leq 2} \subset \overline{\mathcal{F}_{>}}$, and $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq} \subset \overline{\mathcal{F}_{>}}$. Clearly $\mathcal{C}_= = \mathcal{C}_{\leq} \cap \mathcal{C}_{\geq}$. But $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\geq} \not\subset \overline{\mathcal{F}_{>}}$.

► **Theorem 1.2.** *There is an FPRAS* for $Z(a, b, c, d)$ if $(a, b, c, d) \in \mathcal{F}_{\leq 2} \cap \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$; there is no FPRAS for $Z(a, b, c, d)$ if $(a, b, c, d) \in \mathcal{F}_{>}$ unless $RP = NP$. In addition, for planar graphs there is an FPRAS for $Z(a, b, c, d)$ if $(a, b, c, d) \in \mathcal{F}_{\leq 2} \cap \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\geq}$.*

► **Remark 1.3.** The results in Theorem 1.2 are 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. After this work was done [10], the first two authors made a connection of the approximate complexity of the eight-vertex model to that of counting perfect matchings on general graphs (#PM) [8], a central open problem in this field. It was proved in [8] that approximating $Z(a, b, c, d)$ on general 4-regular graphs can be reduced to approximating #PM if $(a, b, c, d) \in \mathcal{F}_{\leq 2}$ and is at least as hard as approximating #PM if $(a, b, c, d) \notin \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$. The #PM-hardness result was proved by expressing the eight-vertex model partition function in the Holant framework and utilizing *holographic transformations*. At the end of this section we briefly describe the connection to Holant problems and a major motivation to study the complexity of the eight-vertex model, as part of the classification program of counting problems, quite apart from historical motivations in statistical physics.

► **Remark 1.4.** The relationship of the regions denoted by $\mathcal{F}_{\leq 2}$, $\mathcal{F}_{>}$, \mathcal{A}_{\leq} , \mathcal{B}_{\leq} , \mathcal{C}_{\leq} , \mathcal{C}_{\geq} , and $\mathcal{C}_{=}$ may not be easy to visualize, since they reside in 4-dimensional space. See Figure 2 (where we normalize $d = 1$)[†]. The roles of a , b , c , and d are not symmetric. In particular, d is the weight of sinks and sources and has a special role (e.g. see [15]). If $(a, b, c, d) \in \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$ then $d \leq a, b, c$. Surprisingly, by Theorem 1.2 FPRAS can still exist for planar graphs even when sinks/sources have large weights.



(a) The four corner regions constitute $\mathcal{F}_{>}$. The non-corner region depicted is $\mathcal{F}_{\leq 2} \cap \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$. (b) An extra region that admits FPRAS on planar graphs.

■ **Figure 2** Regions of known complexity in the eight-vertex model.

To get these FPRAS, our most important contribution is a set of *closure properties*. See Section 2. We then use these closure properties to show that a Markov chain designed for the six-vertex model can be adapted to provide our FPRAS. The Markov chain we adapt is the *directed-loop algorithm* invented by Rahman and Stillinger [21]. The state space of

* Suppose $f : \Sigma^* \rightarrow \mathbb{R}$ is a function mapping problem instances to real numbers. A *fully polynomial randomized approximation scheme (FPRAS)* [19] 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 $\Pr[(1 - \varepsilon)f(x) \leq Y \leq (1 + \varepsilon)f(x)] \geq \frac{3}{4}$.

[†] Some 3D renderings of the parameter space can be found at <https://skfb.ly/6C9LE> and <https://skfb.ly/6C9MS>.

our Markov chain for the eight-vertex model consists of even orientations and near-even orientations, which is an extension of the space of valid configurations; the transitions of this algorithm are composed of creating, shifting, and merging of two “defective” edges.

This leads to a *Markov chain Monte Carlo* approximate counting algorithm by *sampling*. To prove that this is an FPRAS, we show that (1) the above Markov chain is rapidly mixing via a *conductance argument* [17, 12, 22, 16], (2) the valid configurations take a non-negligible proportion in the state space, and (3) there is a (not totally obvious) self-reduction (to reduce the computation of the partition function of a graph to that of a “smaller” graph) [18]. All three parts depend on the closure properties. Specifically, we show that when $(a, b, c, d) \in \mathcal{F}_{\leq 2}$, the conductance of the Markov chain can be polynomially bounded *if* the ratio of near-even orientations over even orientations can be polynomially bounded; when $(a, b, c, d) \in \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$, this ratio is indeed polynomially bounded according to the closure properties. Finally a self-reduction whose success in $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$ requires an additional closure property. Therefore, there is an FPRAS in the intersection of $\mathcal{F}_{\leq 2}$ and $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$ which is in the *disordered phase*.

A *4-ary construction* is a 4-regular graph Γ with four “dangling” edges. This defines a *constraint function* of arity 4. In Theorem 2.2 we show that the set of 4-ary constraint functions in $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$ is closed under 4-ary constructions. This is achieved by inventing a “*quantum decomposition*” of even-orientations. To define this, given an even orientation, a *plus pairing* groups the four edges around a vertex into two pairs such that both pairs satisfy “1-in-1-out”; a *minus pairing* groups the four edges around a vertex into two pairs such that both pairs independently satisfy either “2-in” or “2-out”. With weights, this leads to a weighted sum of $3^{|V|}$ “*annotated circuit partitions*”. (Details are in Section 2[‡].)

We use these tools to derive our FPRAS. Surprisingly, for planar graphs in the eight-vertex model we can show an additional region where FPRAS exists (also in the *disordered phase*). For planar graphs, in Theorem 2.3 we show that the extra regions $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\geq} \cap \overline{\mathcal{F}_{>}}$ and $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{=}$ also enjoy closure properties. This leads to an FPRAS on planar graphs when the parameter setting is in the intersection of $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\geq} \cap \overline{\mathcal{F}_{>}}$ and $\mathcal{F}_{\leq 2}$. And since $\mathcal{F}_{\leq 2} \subset \overline{\mathcal{F}_{>}}$, combined with the FPRAS on general graphs, we get an FPRAS for $\mathcal{F}_{\leq 2} \cap \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq}$ for all planar graphs. Considering the fact that the exact complexity for the eight-vertex model on planar graphs is not even understood, this is one of the very few cases where research on approximate complexity has advanced beyond that on exact complexity.

The NP-hardness of approximation in the whole *ordered phases* is shown by reductions from the problem of computing the maximum cut on a 3-regular graph. For the eight-vertex models not included in the six-vertex model ($d \neq 0$), both the reduction source and the “gadgets” we employ to prove the hardness are substantially different from those used in the hardness proof of the six-vertex model [9]. We note that the parameter settings in [15] where torpid mixing is proved are contained in our NP-hardness region.

In addition to the complexity result, we show that there is a fundamental difference in the behavior on the two sides separated by the phase transition threshold, in terms of closure properties. In Theorem 2.2, we show that the set of 4-ary constraint functions lying in the complement of $\mathcal{F}_{>}$ is closed under 4-ary constructions. We prove in this paper that approximation is hard on $\mathcal{F}_{>}$. It is not known if the eight-vertex model in the full region of $\overline{\mathcal{F}_{>}}$ admits FPRAS or not.

[‡] We use the term “quantum” is emphasize that these are linear combinations, or superpositions, of objects called annotated circuit partitions. There are similarities to holographic transformations, where cancellations occur in the analysis. However, currently we do not have any direct link to quantum computing.

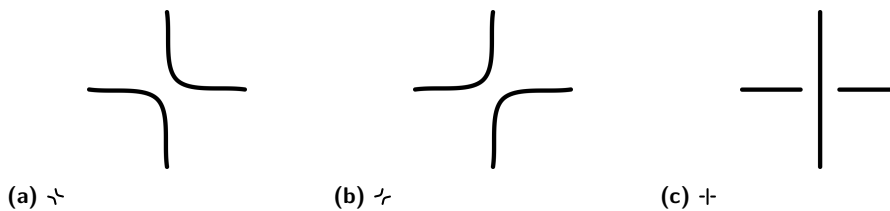
Aside from statistical physics, perhaps a more direct link of the study of partition functions of the eight-vertex model and complexity theory is the classification program of counting problems [4]. The eight-vertex model fits into the wider class of Holant problems. Holant problems are a broad class of counting problems that are more general and expressive than counting constraint satisfaction problems [14, 7], for which complexity dichotomies have been proved [3, 13, 11, 5]. Previous complexity dichotomy theorems have achieved a complete classification for the exact complexity of Holant problems for any set of *symmetric* constraint functions [4]. It turns out that the eight-vertex model is the special case of a Holant problem. In fact, the eight-vertex model can be expressed precisely as a Holant problem in the orientation setting with a single arity-4 *non-symmetric* constraint function that satisfies a parity condition. Under a suitable holographic transformation, it can be expressed as a Holant problem in a standard form. The worst case complexity of the eight-vertex model has been proved in [6]. It is hoped that this set of Holant problems serves as a base case for a future complete worst case complexity classification of Holant problems. The results in this paper are instead on the approximate complexity of the eight-vertex model, and are technically distinct. But a major motivation comes from the classification program on counting problems.

Previous results in approximate counting are mostly about spin systems. The present paper, together with [9], is probably the first fruitful attempt in the Holant literature to make connections to phase transitions. While there is still a gap in the complexity picture for the six-vertex and eight-vertex models, we believe the framework set in this paper gives a starting point for studying the approximate complexity of a broader class of counting problems.

2 Closure Properties

We introduce a “quantum decomposition” for the eight-vertex model, in which every configuration on G is expressed as a “superposition” of $3^{|V|}$ annotated circuit partitions.

Let v be a vertex of G , and e_1, e_2, e_3, e_4 the four labeled edges incident to v . A *pairing* ϱ at v is a partition of $\{e_1, e_2, e_3, e_4\}$ into two pairs. There are exactly three distinct pairings at v (Figure 3) which we denote by three special symbols: \curvearrowright , \curvearrowleft , \dashv , respectively. A *circuit partition* of G is a partition of the edges of G into edge-disjoint circuits (in such a circuit vertices may repeat but edges may not). It is in 1-1 correspondence with a family of pairings $\varphi = \{\varrho_v\}_{v \in V}$, where $\varrho_v \in \{\curvearrowright, \curvearrowleft, \dashv\}$ is a pairing at v —once the pairing at each vertex is fixed, then the two edges paired together at each vertex is also adjacent in the same circuit.



■ **Figure 3** (Unsigned) pairings at a degree 4 vertex.




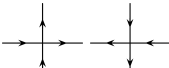
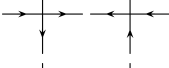

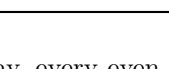
A *signed pairing* ϱ_v at v is a pairing with a sign, either *plus* (+) or *minus* (−). We denote a signed pairing by ϱ_+ or ϱ_- if the pairing is ϱ and the sign is plus or minus, respectively. An *annotated circuit partition* of G , or *acp* for short, is a circuit partition of G together with

4:6 Approximability of the Eight-Vertex Model

a map $V \rightarrow \{+, -\}$ such that along every circuit one encounters an even number of $-$ (a repeat vertex with $-$ counts twice on the circuit). Thus, it is in 1-1 correspondence with a family of signed pairings for all $v \in V$, with the restriction that there is an *even* number of $-$ along each circuit. Each circuit C in an *acp* has exactly two directed *states* – starting at an arbitrary edge in C with one of the two orientations on this edge, one can uniquely orient every edge in C such that for every vertex v on C , two edges incident at v paired up by $+$ have consistent orientations at v (i.e., they form “1-in-1-out” at v), whereas two edges paired up by $-$ have contrary orientations at v (i.e., they form “2-in” or “2-out” at v). These two directed states of C are well-defined because cyclically the direction of edges along C changes an even number of times, precisely at the minus signs. A *directed annotated circuit partition* (*dacp*) is an *acp* with each circuit in a directed state. If an *acp* has k circuits, then it defines 2^k *dacp*’s.

Next we describe an association between even orientations and *acp*’s as well as *dacp*’s. Given an even orientation τ of G , every local configuration of τ at a vertex defines exactly three signed pairings at this vertex according to Table 1. Note that, given τ and a pairing at a vertex v , the two pairs have either *both* consistent or *both* contrary orientations. Thus the same sign, $+$ or $-$, works for both pairs, although this depends on the pairing at v .

■ **Table 1** Map from eight local configurations to signed pairings.

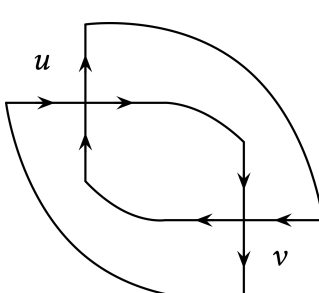
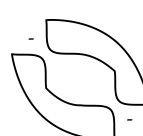



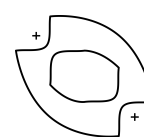
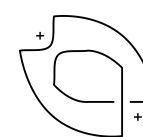
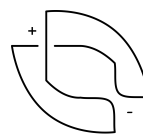
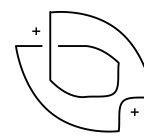
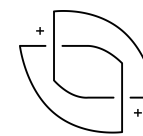
Configurations	Weight	Sign		
				
	a	-	+	+
	b	+	-	+
	c	+	+	-
	d	-	-	-

In this way, every even orientation τ defines $3^{|V|}$ *acp*’s, denoted by $\Phi(\tau)$. See Table 2 and Table 3 for two examples. Moreover, for any *acp* $\varphi \in \Phi(\tau)$, every circuit in φ is in one of the two well-defined directed states under the orientation τ . Thus each even orientation τ defines $3^{|V|}$ *dacp*’s.

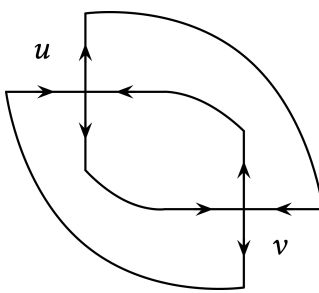
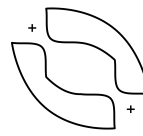
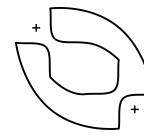
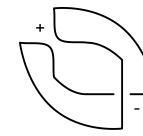
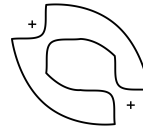
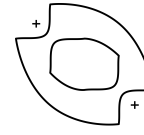
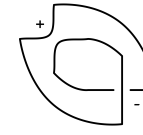
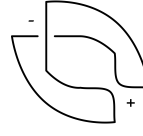
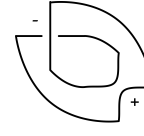
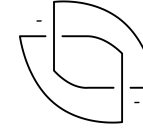
Conversely, for any *dacp*, if we ignore the signs at all vertices we get a valid even orientation (because each sign applies to both pairs). If a *dacp* comes from $\Phi(\tau)$ then we get back the even orientation τ . Therefore, the association from even orientations to *dacp*’s is 1-to- $3^{|V|}$, non-overlapping, and surjective. For every vertex v with the constraint function parameters (a, b, c, d) , we define a local weight function w that assigns a local weight to the six signed pairings at v , such that

$$\begin{cases} a = w(\swarrow_) + w(\nearrow_) + w(\dashv_) \\ b = w(\swarrow_) + w(\nwarrow_) + w(\dashv_) \\ c = w(\swarrow_) + w(\nearrow_) + w(\dashv_) \\ d = w(\swarrow_) + w(\nwarrow_) + w(\dashv_) \end{cases} \quad (2)$$

■ **Table 2** An even orientation and its quantum decomposition into *acp*'s.

τ	$\Phi(\tau)$		
			
			
			

■ **Table 3** Another even orientation and its quantum decomposition into *acp*'s.

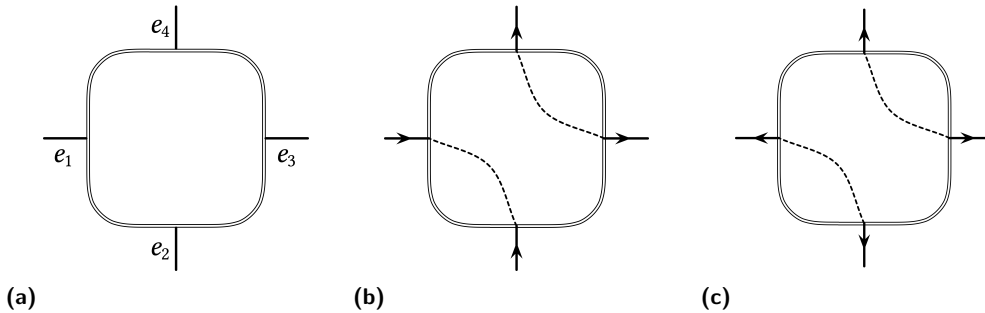
τ	$\Phi(\tau)$		
			
			
			

Note that for any a, b, c, d this is a linear system of rank 4 in six variables, and there is a solution space of dimension 2 (Lemma 2.5 discusses this freedom). Define the weight $\tilde{w}(\varphi)$ of an annotated circuit partition φ , either undirected (*acp*) or directed (*dacp*), be the product of weights at each vertex. Then the weight of an eight-vertex model configuration τ is equal to $\sum_{\varphi \in \Phi(\tau)} \tilde{w}(\varphi)$. This is obtained by writing a term in the summation in (1), which is a product of sums by (2), as a sum of products. Note that a single *acp* has the same weight when it becomes directed regardless which directed state the *dacp* is in.

We will illustrate the above in detail by the examples in Table 2 and Table 3. We assume the same constraint (a, b, c, d) is applied at u and v . The orientation at one vertex determines the other in this graph G . There are a total 8 valid configurations, 4 of which are total reversals of the other 4. $Z(G) = 2[a^2 + b^2 + c^2 + d^2]$. When we expand $Z(G)$ using (2) we

get a total of 72 terms. These correspond to 72 *dacp*'s. There are 9 ways to assign a pairing at u and at v . If we consider the configuration in Table 2, these 9 ways are listed under $\Phi(\tau)$, where the local orientation also determines a sign \pm at both u and v . These are 9 *acp*'s (without direction). (Here we take the view of decomposing an orientation into *acp*'s rather than *dacp*'s to illustrate the idea of quantum decomposition which will be exploited later in proofs.) For each *acp* φ , the weight $\tilde{w}(\varphi)$ is defined (without referring to the *dacp*, or the state of orientation on these circuits). Three of the *acp*'s (in the diagonal positions) define two distinct circuits while the other six define one circuit each. For each 2-tuple of pairings (ρ_u, ρ_v) that results in two circuits, the only valid annotations assign $(+, +)$ or $(-, -)$ at (u, v) , giving a total of 6 *acp*'s. And since each has two circuits, there are a total of 24 *dacp*'s. For the other six (off-diagonal) 2-tuples of pairings (ρ_u, ρ_v) that results in a single circuit, each has 4 valid annotations, giving a total of 24 *acp*'s. But these have only one circuit and thus give 48 *dacp*'s. To appreciate the “quantum superposition” of the decomposition, note that the same *acp* that has $(\mathcal{J}_+, \mathcal{J}_+)$ at (u, v) appears in both decompositions for the distinct configurations in Table 2 and Table 3[§].

► Remark 2.1. While a weight function w satisfying (2) is not unique, there are some regions of (a, b, c, d) that can be specified directly in terms of w by any weight function w satisfying (2), and the specification is independent of the choice of the weight function. E.g., the region $\overline{\mathcal{F}}_{>}$ is specified by $\begin{cases} w(\mathcal{N}_+) + w(\mathcal{J}_-) + w(\mathcal{I}_-) \geq 0 \\ w(\mathcal{N}_-) + w(\mathcal{J}_+) + w(\mathcal{I}_-) \geq 0 \\ w(\mathcal{N}_-) + w(\mathcal{J}_-) + w(\mathcal{I}_+) \geq 0 \\ w(\mathcal{N}_+) + w(\mathcal{J}_+) + w(\mathcal{I}_+) \geq 0 \end{cases}$. Also \mathcal{A}_{\leq} is specified by $w(\mathcal{N}_-) \leq w(\mathcal{N}_+)$, \mathcal{B}_{\leq} by $w(\mathcal{J}_-) \leq w(\mathcal{J}_+)$, and \mathcal{C}_{\leq} by $w(\mathcal{I}_-) \leq w(\mathcal{I}_+)$. In Lemma 2.5 (stated later), we will show that a nonnegative weight function w satisfying (2) exists iff $(a, b, c, d) \in \overline{\mathcal{F}}_{>}$.



■ Figure 4 A 4-ary construction in the eight-vertex model.

A *4-ary construction* is a 4-regular graph Γ having four “external” edges (Figure 4a), and a constraint function on each node. It defines a 4-ary constraint function with four input variables as the partial sum in $Z(\Gamma)$ with a given assignment on the dangling edges. If we imagine the graph Γ is shrunk to a single point except the 4 external edges, then a 4-ary construction can be viewed as a virtual vertex with parameters (a', b', c', d') in the eight-vertex model, for some $a', b', c', d' \geq 0$ (satisfying the even orientation rule and arrow reversal symmetry). A *planar 4-ary construction* is a 4-regular plane graph with four dangling edges on the outer face ordered counterclockwise e_1, e_2, e_3, e_4 .

We say a set of constraint functions S is closed under 4-ary constructions if the constraint function of any 4-ary construction using functions from S also belongs to S .

[§] There is a superficial similarity between the quantum decomposition and skein relations in knot theory [20]. This remains to be explored further.

► **Theorem 2.2.** *Constraint function sets $\overline{\mathcal{F}}_>$ and $\mathcal{A}_\leq \cap \mathcal{B}_\leq \cap \mathcal{C}_\leq$ are closed under 4-ary constructions.*

► **Theorem 2.3.** *Constraint function sets $\mathcal{A}_\leq \cap \mathcal{B}_\leq \cap \mathcal{C}_\geq \cap \overline{\mathcal{F}}_>$ and $\mathcal{A}_\leq \cap \mathcal{B}_\leq \cap \mathcal{C}_=$ are closed under 4-ary plane constructions.*

A *trail & circuit partition (tcp)* for a 4-ary construction Γ is a partition of the edges in Γ into edge-disjoint circuits and exactly two *trails* (walks with no repeated edges) which end in the four external edges. An *annotated trail & circuit partition (atcp)* for Γ is a *tcp* with a valid annotation, which assigns an *even* number of $-$ sign along each circuit. Like circuits, each trail in an *atcp* has exactly two directed states. If an *atcp* φ has k circuits (and 2 trails), then φ defines 2^{k+2} *directed annotated trail & circuit partitions (datcp)*. The weight $\tilde{w}(\varphi)$ of an annotated trail & circuit partition φ , either an *atcp* or *datcp*, can be similarly defined. Again set the weight function as in (2).

For each dangling edge e_i ($1 \leq i \leq 4$) of Γ , let us describe the state of e_i by 0 (or 1) if it is coming into (respectively going out of) Γ . Denote the constraint function of Γ by f and consider $f(0011)$. Under the eight-vertex model, if a configuration τ of the 4-ary construction with constraint function f has a nonzero contribution to $f(0011)$, it has e_1, e_2 coming in and e_3, e_4 going out. The contribution by τ is a weighted sum over a set $\Phi_{0011}(\tau)$ of *datcp*. Each *datcp* in $\Phi_{0011}(\tau)$ is captured in exactly one of the following three types, according to how e_1, e_2, e_3, e_4 are connected by the two trails:

(1) $\{ \overset{e_1}{\rightarrow} \square \overset{e_2}{\leftarrow}, \overset{e_3}{\leftarrow} \square \overset{e_4}{\rightarrow} \}$ and on both trails the numbers of minus pairings are odd; or

(2) $\{ \overset{e_1}{\rightarrow} \square \overset{e_4}{\rightarrow}, \overset{e_2}{\rightarrow} \square \overset{e_3}{\rightarrow} \}$ and on both trails the numbers of minus pairings are even

(Figure 4b); or

(3) $\{ \overset{e_1}{\rightarrow} \square \overset{e_3}{\rightarrow}, \overset{e_2}{\rightarrow} \square \overset{e_4}{\rightarrow} \}$ and on both trails the numbers of minus pairings are even.

Let $\Phi_{0011, \mathcal{N}_-}$, $\Phi_{0011, \mathcal{J}_+}$ and $\Phi_{0011, \mathcal{I}_+}$ be the subsets of *datcp* contributing to $f(0011)$ defined in case (1), (2) and (3) respectively. The value $f(0011)$ is a weighted sum of contributions according to \tilde{w} from these three disjoint sets. Defining the weight of a set Φ of *datcp* by $W(\Phi) = \sum_{\varphi \in \Phi} \tilde{w}(\varphi)$ yields $f(0011) = W(\Phi_{0011, \mathcal{N}_-}) + W(\Phi_{0011, \mathcal{J}_+}) + W(\Phi_{0011, \mathcal{I}_+})$. Similarly we can define $\Phi_{1100, \mathcal{N}_-}$, $\Phi_{1100, \mathcal{J}_+}$ and $\Phi_{1100, \mathcal{I}_+}$, and get $f(1100) = W(\Phi_{1100, \mathcal{N}_-}) + W(\Phi_{1100, \mathcal{J}_+}) + W(\Phi_{1100, \mathcal{I}_+})$. Note that there is a bijective weight-preserving map between $\Phi_{0011, \mathcal{N}_-}$ and $\Phi_{1100, \mathcal{N}_-}$ by reversing the direction of every circuit and trail of a *datcp*. Thus, $W(\Phi_{0011, \mathcal{N}_-}) = W(\Phi_{1100, \mathcal{N}_-})$, $W(\Phi_{0011, \mathcal{J}_+}) = W(\Phi_{1100, \mathcal{J}_+})$, and $W(\Phi_{0011, \mathcal{I}_+}) = W(\Phi_{1100, \mathcal{I}_+})$. Consequently $f(0011) = f(1100)$. Similarly we have $f(0110) = f(1001)$, $f(0101) = f(1010)$ and $f(0000) = f(1111)$. For any pairing ϱ , and for every 4-bit pattern $b_1 b_2 b_3 b_4 \in \{0, 1\}^4$, we can define $\Phi_{b_1 b_2 b_3 b_4, \varrho_+}$ if (both) paired $b_i \neq b_j$, and $\Phi_{b_1 b_2 b_3 b_4, \varrho_-}$ if (both) paired $b_i = b_j$. We can prove the following and call the common value $W(\mathcal{N}_-)$:

$$\blacksquare \quad W(\Phi_{0011, \mathcal{N}_-}) = W(\Phi_{1100, \mathcal{N}_-}) = W(\Phi_{0000, \mathcal{N}_-}) = W(\Phi_{1111, \mathcal{N}_-}).$$

We can prove 5 other sets of similar equalities, and call them $W(\mathcal{J}_-)$, $W(\mathcal{I}_-)$, $W(\mathcal{N}_+)$,

$$W(\mathcal{J}_+), \text{ and } W(\mathcal{I}_+). \text{ Consequently, } f \text{ has parameters } \begin{cases} a' = W(\mathcal{N}_-) + W(\mathcal{J}_+) + W(\mathcal{I}_+) \\ b' = W(\mathcal{N}_+) + W(\mathcal{J}_-) + W(\mathcal{I}_+) \\ c' = W(\mathcal{N}_+) + W(\mathcal{J}_+) + W(\mathcal{I}_-) \\ d' = W(\mathcal{N}_-) + W(\mathcal{J}_-) + W(\mathcal{I}_-) \end{cases}.$$

Proof Sketch for the closure of $\mathcal{A}_\leq \cap \mathcal{B}_\leq \cap \mathcal{C}_\leq$. By definition $(a, b, c, d) \in \mathcal{A}_\leq \cap \mathcal{B}_\leq \cap \mathcal{C}_\leq$

means that $\begin{cases} a+d \leq b+c \\ b+d \leq a+c \\ c+d \leq a+b \end{cases}$. By the weight function w defined in (2) this is equivalent to

$$\begin{cases} w(\mathcal{N}_+) \geq w(\mathcal{N}_-) \\ w(\mathcal{J}_+) \geq w(\mathcal{J}_-) \\ w(\mathcal{I}_+) \geq w(\mathcal{I}_-) \end{cases}. \text{ Since } \mathcal{A}_\leq \cap \mathcal{B}_\leq \cap \mathcal{C}_\leq \subset \overline{\mathcal{F}}_>, \text{ by Lemma 2.5 (stated later) we can assume } w$$

4:10 Approximability of the Eight-Vertex Model

is nonnegative. Therefore, we only need to establish $\begin{cases} W(\mathfrak{S}_+) \geq W(\mathfrak{S}_-) \\ W(\mathfrak{S}_+) \geq W(\mathfrak{S}_-) \\ W(+_+) \geq W(+_-) \end{cases}$. We prove $W(\mathfrak{S}_+) \geq W(\mathfrak{S}_-)$. Proof for the other two inequalities is symmetric.

An *atcp* is a *tcp* together with a valid annotation. Consider the set Ψ of *tcp* such that the two (unannotated) trails connect e_1 with e_2 , and e_3 with e_4 . Denote by χ_{12} (respectively χ_{34}) the trail in ψ connecting e_1 and e_2 (respectively e_3 and e_4). Each *tcp* $\psi \in \Psi$ may have many valid annotations.

Since Γ is 4-regular, any vertex inside Γ appears exactly twice counting multiplicity in a *tcp* ψ . It appears either as a self-intersection point of a trail or a circuit, or alternatively in exactly two distinct trails/circuits. So when traversed, in total one encounters an even number of $-$ among all circuits and the two trails in any valid annotation of ψ , and since one encounters an even number of $-$ along each circuit, the numbers of $-$ along χ_{12} and χ_{34} have the same parity. We say a valid annotation of ψ is *positive* if there is an even number of $-$ along χ_{12} (and χ_{34}), and *negative* otherwise.

To prove $W(\mathfrak{S}_+) \geq W(\mathfrak{S}_-)$, it suffices to prove that for each *tcp* $\psi \in \Psi$, the total weight W_+ contributed by the set of positive annotations of ψ is at least the total weight W_- contributed by the set of negative annotations of ψ . We prove this nontrivial statement by induction on the number N of vertices shared by any two distinct circuits in ψ .

Base case: The base case is $N = 0$. In the base case, we can deal with self-intersections on trails and circuits easily, so let us assume that no trail or circuit is self-intersecting. Then every vertex on any circuit C of ψ is shared by C and exactly one trail, χ_{12} or χ_{34} . Also, every vertex on χ_{12} or χ_{34} is shared with some circuit or the other trail.

We will account for the product values of $w(\mathbf{g}_v)$ according to how v is shared. We first consider shared vertices of a circuit $C \in \psi$ with the trails. Let $s, t \geq 0$ be the numbers of vertices C shares with χ_{12} and χ_{34} , respectively. Let x_i ($1 \leq i \leq s$) (if $s > 0$) and y_j ($1 \leq j \leq t$) (if $t > 0$) be these shared vertices respectively (for $s = 0$ or $t = 0$, the statements below are vacuously true). For any v , if ρ is the pairing at v according to ψ , then let $w_+(v) = w(\rho_+)$, and $w_-(v) = w(\rho_-)$, both at v . In any valid annotation of ψ (either positive or negative), one encounters an even number of $-$ on the vertices along C , each of which is shared with exactly one of χ_{12} and χ_{34} . Hence the number of $-$ in x_i ($1 \leq i \leq s$) has the same parity as the number of $-$ in y_j ($1 \leq j \leq t$). Other than having the same parity, the annotation for x_i ($1 \leq i \leq s$) is independent from the annotation for y_j ($1 \leq j \leq t$) for a valid annotation, and from the annotations on other circuits. Let $S_+(C)$ (respectively $S_-(C)$) be the sum of products of $w(\mathbf{g}_v)$ over $v \in \{x_i \mid 1 \leq i \leq s\}$, summed over valid annotations such that the number of $-$ in x_i ($1 \leq i \leq s$) is even (respectively odd). Similarly let $T_+(C)$ (respectively $T_-(C)$) be the corresponding sums for y_j ($1 \leq j \leq t$). We have

$$S_+(C) - S_-(C) = \prod_{i=1}^s (w_+(x_i) - w_-(x_i)) \geq 0,$$

$$T_+(C) - T_-(C) = \prod_{j=1}^t (w_+(y_j) - w_-(y_j)) \geq 0.$$

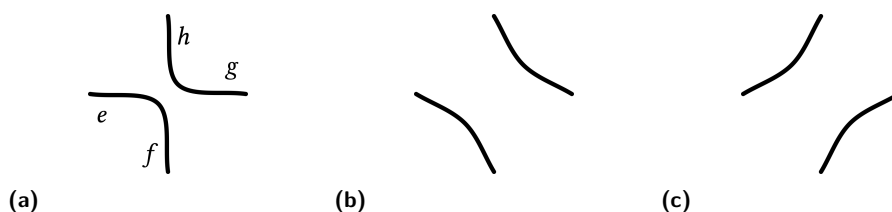
Both differences are nonnegative by the hypothesis.

The product $S_+(C)T_+(C)$ is the sum over all valid annotations of vertices on C such that the numbers of $-$ on vertices shared by χ_{12} and C and by χ_{34} and C are both even. Similarly $S_-(C)T_-(C)$ is the sum over all valid annotations of vertices on C such that the numbers of $-$ on vertices shared by χ_{12} and C and by χ_{34} and C are both odd. We have $S_+(C)T_+(C) \geq S_-(C)T_-(C)$.

Next we also account for the vertices shared by χ_{12} and χ_{34} in ψ . Let p be this number and if $p > 0$ let z_k ($1 \leq k \leq p$) be these vertices. Let q be the number of circuits in ψ , denoted by C_l ($1 \leq l \leq q$) (if $q > 0$). Then we claim that

$$W_+ - W_- = \prod_{k=1}^p (w_+(z_k) - w_-(z_k)) \prod_{l=1}^q (S_+(C_l)T_+(C_l) - S_-(C_l)T_-(C_l)),$$

and in particular $W_+ - W_- \geq 0$. To prove this claim we only need to expand the product, and separately collect terms that have a + sign and a - sign. In a product term in the fully expanded sum, let p' be the number of $-w_-(z_k)$, and q' be the number of $-S_-(C_l)T_-(C_l)$. Then a product term has a + sign (and thus included in W_+) iff $p' + q' \equiv 0 \pmod{2}$.



■ **Figure 5** Possible ways of deleting a vertex. The vertex (not explicitly shown) at the center of part (a) is removed in part (b) and (c).

Induction step: Suppose v is a shared vertex between two distinct circuits C_1 and C_2 , and let $\{e, f, g, h\}$ be its incident edges in Γ . We may assume the pairing ϱ_v in ψ is $\{e, f\}$ and $\{g, h\}$, and thus e, f are in one circuit, say C_1 , while g, h are in another circuit C_2 (Figure 5a). Define Γ' to be the 4-ary construction obtained from Γ by deleting v and merging e with f , and g with h (Figure 5b). Define Γ'' to be the 4-ary construction obtained from Γ by deleting v and merging e with h , and f with g (Figure 5c). Note that in Γ' , we have two circuits C'_1 and C'_2 (each has one fewer vertex v from C_1 and C_2), but in Γ'' the two circuits are merged into one C^* . Define W'_+ and W'_- (respectively W''_+ and W''_-) similarly for Γ' (respectively Γ'') with tcp being $\psi' = \psi \setminus \{\varrho_v\}$.

We can decompose $W_+ - W_-$ according to whether the sign on ϱ_v is + or -. Recall that for any valid annotation of ψ , one encounters an even number of - along C_1 and C_2 . If the sign on ϱ_v is +, the number of - along C_1 (and C_2) at all vertices other than v in any valid annotation is always even; if the sign on ϱ_v is -, this number (for both C_1 and C_2) is always odd. $W_+ - W_-$ can be decomposed into two parts, corresponding to terms with ϱ_v being + or - respectively. All terms of the first (and second) part have a factor $w_+(v)$ (and $w_-(v)$ respectively). And so we can write

$$W_+ - W_- = w_+(v)[W_+ - W_-]_e + w_-(v)[W_+ - W_-]_o, \tag{3}$$

where $[W_+ - W_-]_e$ and $[W_+ - W_-]_o$ collect terms in $W_+ - W_-$ in the first and second part respectively, but without the factor at v . However by considering valid annotations for Γ' we also have

$$W'_+ - W'_- = [W_+ - W_-]_e, \tag{4}$$

because a valid annotation on both C'_1 and C'_2 is equivalent to a valid annotation on both C_1 and C_2 with v assigned +. Similarly, by considering valid annotations for Γ'' we also have

$$W''_+ - W''_- = [W_+ - W_-]_e + [W_+ - W_-]_o, \tag{5}$$

4:12 Approximability of the Eight-Vertex Model

because depending on whether ϱ_v is assigned $+$ or $-$, a valid annotation on both C_1 and C_2 gives either both an even or both an odd number of $-$ on $C_1 \setminus \{v\}$ and $C_2 \setminus \{v\}$, which is equivalent to an even number of $-$ on the merged circuit C^* . From (3,4,5) we have

$$W_+ - W_- = (w_+(v) - w_-(v))(W'_+ - W'_-) + w_-(v)(W''_+ - W''_-).$$

By induction, $W'_+ \geq W'_-$ and $W''_+ \geq W''_-$. Since $w_+(v) \geq w_-(v)$ is given by hypothesis, we get $W_+ \geq W_-$. ◀

The proof of planar closures is similar but more intricate, and uses the *Jordan Curve Theorem*.

► **Lemma 2.4.** *Suppose $x, x', y, y', z, z' \in \mathbb{R}$ satisfy the eight inequalities: $X + Y + Z \geq 0$ where $X \in \{x, x'\}, Y \in \{y, y'\}, Z \in \{z, z'\}$. Then there exist nonnegative $\tilde{x}, \tilde{x}', \tilde{y}, \tilde{y}', \tilde{z}, \tilde{z}'$ such that all eight sums $X + Y + Z$ are unchanged when x, x', y, y', z, z' are substituted by the respective values $\tilde{x}, \tilde{x}', \tilde{y}, \tilde{y}', \tilde{z}, \tilde{z}'$.*

► **Lemma 2.5.** *The parameter setting (a, b, c, d) belongs to $\overline{\mathcal{F}}_{>}$ iff there exists a nonnegative weight function w satisfying (2).*

► **Notation.** *Fix for each vertex v in a 4-regular graph G a weight function w on signed pairings (satisfying (2) at v). Let $Z_v(\varrho)$ be the weighted sum of the set of all dacp's having the signed pairing ϱ at v .*

► **Corollary 2.6.** *If at each vertex in a 4-regular graph G we have a nonnegative weight function w such that $w(\swarrow_+) \geq w(\swarrow_-)$, $w(\nearrow_+) \geq w(\nearrow_-)$, and $w(\dashv_+) \geq w(\dashv_-)$, then $Z_v(\swarrow_+) \geq Z_v(\swarrow_-)$, $Z_v(\nearrow_+) \geq Z_v(\nearrow_-)$, and $Z_v(\dashv_+) \geq Z_v(\dashv_-)$ at each vertex v in G .*

► **Corollary 2.7.** *If at each vertex in a 4-regular plane graph G we have a nonnegative weight function w such that $w(\swarrow_+) \geq w(\swarrow_-)$, $w(\nearrow_+) \geq w(\nearrow_-)$, and $w(\dashv_+) \leq w(\dashv_-)$, then $Z_v(\swarrow_+) \geq Z_v(\swarrow_-)$, $Z_v(\nearrow_+) \geq Z_v(\nearrow_-)$, and $Z_v(\dashv_+) \leq Z_v(\dashv_-)$ at each vertex v in G .*

3 FPRAS

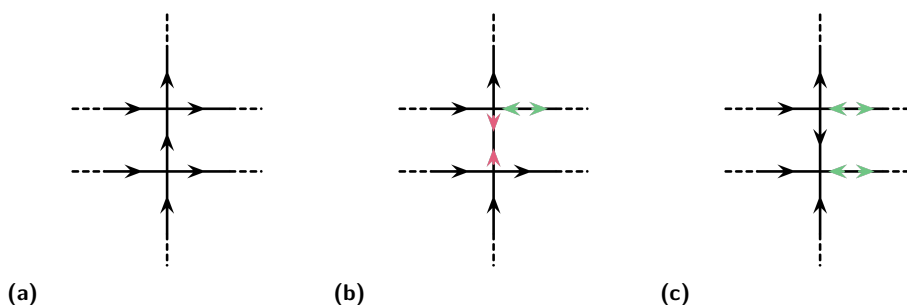
► **Theorem 3.1.** *There is an FPRAS for $Z(a, b, c, d)$ if $(a, b, c, d) \in \mathcal{F}_{\leq 2} \cap \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$.*

► **Theorem 3.2.** *There is an FPRAS for $Z(a, b, c, d)$ on planar graphs if $(a, b, c, d) \in \mathcal{F}_{\leq 2} \cap \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\geq}$.*

We prove the FPRAS results using the common approach of approximately counting via almost uniformly sampling [18, 17, 12, 22, 16] by showing that the *directed-loop algorithm*, a Markov chain algorithm designed for the six-vertex model[¶], can be adapted for the eight-vertex model. The state space Ω of our Markov chain \mathcal{MC} for the eight-vertex model consists of the set Ω_0 of *even orientations* (e.g. Figure 6a) and the set of *near-even orientations* with exactly two “defective” edges (e.g. Figure 6b and Figure 6c), which is an extension of the space of valid configurations; the transitions of this algorithm are composed of creating (from Figure 6a to Figure 6b), shifting (between Figure 6b and Figure 6c), and merging (from Figure 6b to Figure 6a) of the two defects on edges.

► **Notation.** *Let $\mathcal{Z}(S)$ be the weighted sum of states in the set S .*

[¶] The directed-loop algorithm was invented by Rahman and Stillinger [21] and is widely used for the six-vertex model (e.g., [24, 1, 23]).



■ **Figure 6** Examples of the states in the directed-loop algorithm.

Proof of Theorem 3.1. It is easy to show that \mathcal{MC} is *irreducible* and *aperiodic*, and it satisfies the *detailed balance condition* under the Gibbs distribution. By the theory of Markov chains, we have an almost uniform sampler of $\Omega_0 \cup \Omega_2$. This sampler is efficient if \mathcal{MC} is rapidly mixing. According to Lemma 3.5, when $(a, b, c, d) \in \mathcal{F}_{\leq 2}$, \mathcal{MC} is rapidly mixing if $\frac{Z(\Omega_2)}{Z(\Omega_0)}$ is polynomially bounded via a *conductance argument* [17, 12, 22, 16] in which the *paths* between any two states τ_1 and τ_2 and the amount of *flow* each of them takes are decided by a *quantum decomposition* of the “symmetric difference” $\tau_1 \oplus \tau_2$. According to Corollary 3.3 (a corollary of the *closure property* Theorem 2.2), $\frac{Z(\Omega_2)}{Z(\Omega_0)}$ is polynomially bounded. As a consequence, if all the constraint function comes from $\mathcal{F}_{\leq 2} \cap \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$, \mathcal{MC} is rapidly mixing and even orientations take a non-negligible proportion in the state space. Therefore, we are able to efficiently sample valid eight-vertex configurations according to the Gibbs measure on Ω_0 (almost uniformly).

In order for self-reduction, we need to extend the type of vertices a graph allows in the eight-vertex model. Previously, a graph can only have degree 4 vertices, on each of which the constraint function satisfies the even orientation rule and has arrow reversal symmetry. Now, a graph can also have degree 2 vertices, on each of which the constraint function satisfies the “1-in-1-out” rule and both valid local configurations have weight 1. Both Lemma 3.5 and Corollary 3.3 still hold with this extension, because such a degree 2 vertex and its two incident edges just work together as a single edge.

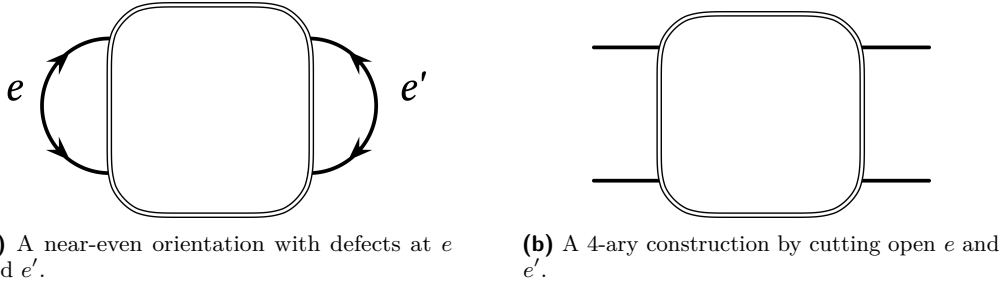
We design the following algorithm to approximately computing $Z(G)$ via sampling with \mathcal{MC} . As we have argued in Section 2, the partition function of the eight-vertex models can be viewed as the weighted sum of the set of *dacp*’s (the weight of a *dacp* and its underlying *acp* are the same). Since every constraint function belongs to $\overline{\mathcal{F}}_{>}$, by Lemma 2.5 for each vertex $v \in V$ we can choose a nonnegative weight function w on signed pairings at v . Thus the ratios among different signed pairings $\{\curvearrowright, \curvearrowleft, \dashv\} \times \{+, -\}$ showing up at v in weighted *dacp*’s can be uniquely determined by the ratios among different local orientations (represented by a, b, c , and d) showing up at v . According to Corollary 2.6 (another corollary of Theorem 2.2), there must be a pairing $\varrho \in \{\curvearrowright_+, \curvearrowleft_+, \dashv_+\}$ showing up at v with probability at least $\frac{1}{6}$ among all six signed pairings, as long as the partition function is not zero (this can be easily tested in polynomial time). Therefore, running \mathcal{MC} on G , we can approximate, with a sufficient $1/\text{poly}(n)$ precision, the probability of having $\varrho \in \{\curvearrowright_+, \curvearrowleft_+, \dashv_+\}$ at v , denoted by $\text{Pr}_v(\varrho)$. Denote by $G_{v,\varrho}$ the graph with v being split into v_1 and v_2 (both satisfy the “1-in-1-out” rule) and the edges reconnected according to ϱ . Write the partition function of $G_{v,\varrho}$ as $Z(G_{v,\varrho})$, we have $\text{Pr}_v(\varrho) = w(\varrho)Z(G_{v,\varrho})/Z(G)$ which means $Z(G) = w(\varrho)Z(G_{v,\varrho})/\text{Pr}_v(\varrho)$. To approximate $Z(G)$ it suffices to approximate $Z(G_{v,\varrho})$, which can be done by running \mathcal{MC} on $G_{v,\varrho}$ and recursing. Repeating this process for $|V|$ steps we decompose the graph G into the base case, a set of disjoint cycles whose partition function is just 2^C where C is the number of cycles. By this self-reduction, the partition function $Z(G)$ can be approximated. ◀

Proof of Theorem 3.2. The proof is similar to that of Theorem 3.1. Given a plane graph G with the constraint function on every vertex from $\mathcal{F}_{\leq 2} \cap \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\geq}$, we can still efficiently sample even orientations according to the Gibbs measure by Lemma 3.5 and Corollary 3.4. However, in order to make our algorithm work, we need to extend the type of vertices in the eight-vertex model again, by allowing degree 2 vertices with the constraint functions satisfying the “2-in/2-out” rule and both valid local configurations have weight 1. One can check that Lemma 3.5 still hold even with this extension.

The self-reduction still looks at a vertex v at a time. According to Corollary 2.7, there must be a pairing $\varrho \in \{\curvearrowright_+, \curvearrowleft_+, \dashv_-, \dashv_+\}$ showing up at v with probability at least $\frac{1}{6}$ among all six signed pairings, as long as the partition function is not zero. If ϱ is \curvearrowright_+ or \curvearrowleft_+ , it can be handled as in the proof of Theorem 3.1. If ϱ is \dashv_- , let G_{v, \dashv_-} be the graph still with v being split into v_1 and v_2 and the edges reconnected according to \dashv_- , but this time both v_1 and v_2 satisfy the “2-in/2-out” rule. Observe that Corollary 3.4 hold for G_{v, \dashv_-} if and only if it holds for G'_{v, \dashv_-} where we replace v by a virtual vertex v' with parameter setting $(a, b, c, d) = (0, 0, 1, 1)$ (this is equivalent as fixing $w(\dashv_-) = 1$ and w on other five signed pairings being 0, if we require w to be nonnegative). Since $(0, 0, 1, 1) \in \mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\geq} \cap \overline{\mathcal{F}_{>}}$. Theorem 2.3 and consequently Corollary 3.4 still hold for G'_{v, \dashv_-} thus also for G_{v, \dashv_-} .

The base case is similar to that in the proof of Theorem 3.1. This time we decompose G into a set of disjoint cycles with an even number of degree 2 vertices that satisfy the “2-in/2-out” rule (by the *Jordan Curve Theorem* argued in the full version). The partition function of this cycle graph is just 2^C where C is the number of cycles. Again, the partition function $Z(G)$ can be approximated. ◀

► **Corollary 3.3.** Given a 4-regular graph $G = (V, E)$, if the constraint function on every vertex is from $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$, then $\frac{Z(\Omega_2)}{Z(\Omega_0)} \leq \binom{|E|}{2}$.



■ **Figure 7** A 4-ary construction made by cutting open the two edges with defects.

Proof. Let $\Omega_2^{\{e, e'\}} \subseteq \Omega_2$ be the set of near-even orientations in which e, e' are the two defective edges. We have $\frac{Z(\Omega_2)}{Z(\Omega_0)} = \sum_{\{e, e'\} \in \binom{E}{2}} \frac{Z(\Omega_2^{\{e, e'\}})}{Z(\Omega_0)}$. For any $\tau \in \Omega_2$, each of e and e' may have both half-edges going out (as in Figure 7a) or coming in, with 4 possibilities. If we “cut open” e and e' as shown in Figure 7b, we get a 4-ary construction Γ using degree 4 vertices with constraint functions in $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\leq}$. Denote the constraint function of Γ by (a', b', c', d') , with the input order being counter-clockwise starting from the upper-left edge. For this 4-ary construction we observe that near-even orientations in $\Omega_2^{\{e, e'\}}$ contribute a total weight $2(a' + d')$ while even orientations in Ω_0 contribute a total weight $2(b' + c')$. By Theorem 2.2 we know that for the 4-ary construction Γ , $a' + d' \leq b' + c'$. Therefore, $\frac{Z(\Omega_2^{\{e, e'\}})}{Z(\Omega_0)} \leq 1$. ◀

► **Corollary 3.4.** *Given a 4-regular plane graph $G = (V, E)$, if the constraint function on every vertex is from $\mathcal{A}_{\leq} \cap \mathcal{B}_{\leq} \cap \mathcal{C}_{\geq} \cap \overline{\mathcal{F}_{>}}$, then $\frac{Z(\Omega_2)}{Z(\Omega_0)} \leq \binom{|E|}{2}$.*

► **Lemma 3.5.** *Assume $Z(\Omega_0) > 0$ and the constraint function on every vertex belongs to $\mathcal{F}_{\leq 2}$. MC is rapidly mixing if $\frac{Z(\Omega_2)}{Z(\Omega_0)}$ is polynomially bounded.*

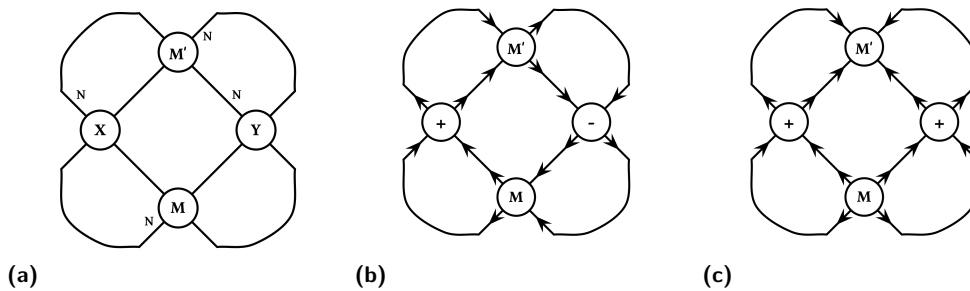
4 Hardness

► **Theorem 4.1.** *If $(a, b, c, d) \in \mathcal{F}_{>}$, then $Z(a, b, c, d)$ does not have an FPRAS unless $RP=NP$.*

► **Remark 4.2.** For any $(a, b, c, d) \in \mathcal{F}_{>}$, at least two of a, b, c , and d are nonzero. The case $d = 0$ and $a, b, c > 0$ was proved in [9]. The case $d = 0$ and one of a, b, c is zero can be proved by a reduction from computing the partition function of the anti-ferromagnetic Ising model on 3-regular graphs; we postpone this proof to an expanded version of this paper. In this section, we prove the theorem when $d > 0$ and $a > b + c + d$. Since the proof of NP-hardness for $Z(a, b, c, d)$ is for not necessarily planar graphs, we can permute the parameters a, b, c . Thus the proof for $b > a + c + d$ and $c > a + b + d$ is symmetric. The adaption that we make to prove the case when $d > a + b + c$ can be found in the full version.

► **Remark 4.3.** The construction for the proof when $a > b + c + d$, or $b > a + c + d$, or $c > a + b + d$, is in fact a bipartite graph. This means approximating $Z(a, b, c, d)$ in those cases is NP-hard even for *bipartite graphs*.

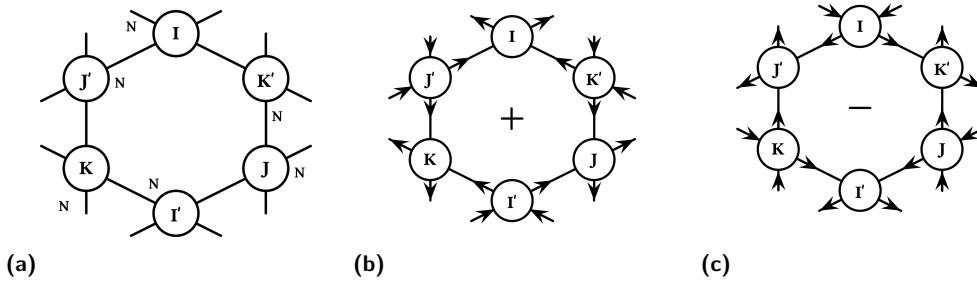
Proof. Let 3-MAX CUT denote the NP-hard problem of computing the cardinality of a maximum cut in a 3-regular graph [25]. We reduce 3-MAX CUT to approximating $Z(a, b, c, d)$. Before proving the theorem we briefly state our idea. Denote an instance of 3-MAX CUT by $G = (V, E)$. Given $V_+ \subseteq V$ and $V_- = V \setminus V_+$, an edge $\{u, v\} \in E$ is in the cut between V_+ and V_- if and only if $(u \in V_+, v \in V_-)$ or $(u \in V_-, v \in V_+)$. The maximum cut problem favors the partition of V into V_+ and V_- so that there are as many edges in $V_+ \times V_-$ as possible. We want to encode this local preference on each edge by a local fragment of a graph G' in terms of configurations in the eight-vertex model.



■ **Figure 8** A four-way connection implementing a single edge in 3-MAX CUT.

First we show how to implement a toy example—a single edge $\{u, v\}$ —by a construction in the eight-vertex model. Suppose there are four vertices X, Y, M, M' connected as in Figure 8a shows. The order of the 4 edges at each vertex is aligned to Figure 1 by a rotation so that the edge marked by “N” corresponds to the north edge in Figure 1. Let us impose the virtual constraint on X and Y so that the parameter setting on each of them is $\check{a} > \check{b} = \check{c} = \check{d} = 0$. (We will show how to implement this virtual constraint in the sense of approximation later.)

In other words, the four edges incident on X can only be in two possible configurations, Figure 1-1 or Figure 1-2. The same is true for Y . We say X (and similarly Y) is in state $+$ if its local configuration is in Figure 1-1 (with the “top” two edges going out and the “bottom” two edges coming in); it is in state $-$ if its local configuration is in Figure 1-2 (with the “top” two edges coming in and the “bottom” two edges going out). Hence there are a total of 4 valid configurations given the virtual constraints. When (X, Y) is in state $(+, -)$ (or $(-, +)$), M and M' have local configurations both being Figure 1-1 (or both being Figure 1-2), with weight a (Figure 8b); when (X, Y) is in state $(+, +)$ (or $(-, -)$), M and M' have local configurations both being Figure 1-7 or Figure 1-8, with weight $d < a$ (Figure 8c). This models how two adjacent vertices interact in 3-MAX CUT. We will call the connection pattern described in Figure 8a between the set of 4 dangling edges incident to X and the set of 4 dangling edges incident to Y (each with two on “top” and two on “bottom”) a *four-way connection*.



■ **Figure 9** A locking device implementing a vertex of degree 3 in 3-MAX CUT.

To model a vertex of degree 3 in a 3-MAX CUT instance, we use the *locking device* in Figure 9a. Let us assume we have the virtual constraint that each of I, I', J, J', K, K' can only be in two local configurations, Figure 1-1 or Figure 1-2. In fact, each locking device has two states, one shown in Figure 9b with every node in configuration Figure 1-1 (called the $+$ state) and the other shown in Figure 9c with every node in configuration Figure 1-2 (called the $-$ state). If we think of the external edges incident to I, J, K to serve as the “top” edges (with “N” aligned with the “N” at X or Y in Figure 8a), and the edges incident to I', J', K' as the “bottom” edges there, then we simulate the \pm state of a degree 3 vertex as follows: (1) top edges are going out and bottom edges are coming in if the device is in $+$ state, and top edges are coming in and bottom edges are going out if the device is in $-$ state; and (2) the top edges on I, J, K are going out or coming in at the same time.

Next we show how to enforce the virtual constraint in Figure 9a that each vertex has two contrary configurations, in the sense of approximation. The idea is to implement an *amplifier* as a 4-ary construction with parameter $(\hat{a}, \hat{b}, \hat{c}, \hat{d})$ such that $\hat{a} \gg \hat{b} + \hat{c} + \hat{d}$ using polynomially many vertices in the eight-vertex model. We obtain such an amplifier by an iteration of Γ shown in Figure 10. The parameter setting (a', b', c', d') of Γ is $\begin{cases} a' = \Lambda(a, b, c, d) \\ b' = \Lambda(b, c, d, a) \\ c' = \Lambda(c, d, a, b) \\ d' = \Lambda(d, a, b, c) \end{cases}$, where

$$\Lambda(\xi, x, y, z) = \xi^7 + (3x^4 + 3y^4 + 3z^4 + 4x^2y^2 + 4x^2z^2 + 4y^2z^2)\xi^3 + (2x^4y^2 + 2x^4z^2 + 2x^2y^4 + 2y^4z^2 + 2x^2z^4 + 2y^2z^4 + 30x^2y^2z^2)\xi. \quad (6)$$

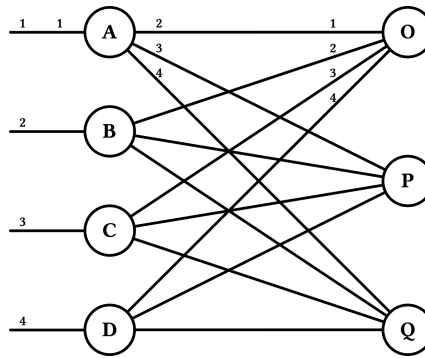


Figure 10 A 4-ary construction that amplifies the maximum among a, b, c, d .

This construction uses 7 vertices and is called a 1-amplifier. We obtain $(a_1, b_1, c_1, d_1) = (a', b', c', d')$ which amplifies the relative weight of configurations in Figure 1-1 or Figure 1-2. If we plug in the amplifier Γ into each vertex of Γ itself (called a 2-amplifier), we can obtain (a_2, b_2, c_2, d_2) using 7^2 vertices. Iteratively, we can construct a series of constraint functions with parameters (a_k, b_k, c_k, d_k) ($k \geq 1$) such that $\begin{cases} a_{k+1} = \Lambda(a_k, b_k, c_k, d_k) \\ b_{k+1} = \Lambda(b_k, c_k, d_k, a_k) \\ c_{k+1} = \Lambda(c_k, d_k, a_k, b_k) \\ d_{k+1} = \Lambda(d_k, a_k, b_k, c_k) \end{cases}$, using 7^k vertices for each k (called a k -amplifier). Lemma 4.4 shows that the asymptotic growth rate is exponential in the number of vertices used.

To reduce the problem 3-MAX CUT to approximating $Z(a, b, c, d)$, let $\kappa > \lambda \geq 1$ be two constants that are sufficiently large. For each 3-MAX CUT instance $G = (V, E)$ with $|V| = n$ and $|E| = m$, we construct a graph G' where a device in Figure 9a is created for each $v \in V$, and a four-way connection is made for every $\{u, v\} \in E(G)$, on the dangling edges corresponding to $\{u, v\}$ as in Figure 8a. For each 4-way connection in Figure 8a, each of the nodes M, M' is replaced by a $(\lambda \log n)$ -amplifier to boost the ratio of the configurations in Figure 1-1 or Figure 1-2 over other configurations. For each device in Figure 9a, each of the nodes I, I', J, J', K, K' is replaced by a $(\kappa \log n)$ -amplifier to lock in the configurations Figure 9b or Figure 9c. We can prove that the maximum size s of all cuts in G can be recovered from an approximate solution to $Z(G'; f)$. In fact, there is a valid configuration (at the granularity of nodes and edges shown in Figure 9a) of weight $(a_{\kappa \log n})^{6n} (a_{\lambda \log n})^{2s} (d_{\lambda \log n})^{2(m-s)}$, and the weighted sum of all configurations is smaller than $\frac{1}{2} (a_{\kappa \log n})^{6n} (a_{\lambda \log n})^{2(s+1)} (d_{\lambda \log n})^{2(m-(s+1))}$. ◀

► **Lemma 4.4.** *Let $(a_k, b_k, c_k, d_k) = \Lambda^{(k)}(a, b, c, d)$ given by (6). Assuming $a_0 > b_0 + c_0 + d_0$, $a_0, d_0 > 0$, and $b_0, c_0 \geq 0$, there exists some constants $\alpha > 0, \beta > 1$ depending only on a_0, b_0, c_0, d_0 such that for all $k \geq 1$, $\frac{a_k}{b_k + c_k + d_k} \geq \alpha \beta^{2^k}$.*

References

- 1 G. T. Barkema and M. E. J. Newman. Monte carlo simulation of ice models. *Phys. Rev. E*, pages 1155–1166, 1998.
- 2 R. J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Academic Press Inc., San Diego, CA, USA, 1982.
- 3 Andrei A. Bulatov. The complexity of the counting constraint satisfaction problem. *J. ACM*, 60(5), October 2013. doi:10.1145/2528400.
- 4 Jin-Yi Cai and Xi Chen. *Complexity Dichotomies for Counting Problems*, volume 1. Cambridge University Press, 2017. doi:10.1017/9781107477063.
- 5 Jin-Yi Cai and Xi Chen. Complexity of counting CSP with complex weights. *J. ACM*, 64(3), June 2017. doi:10.1145/2822891.

- 6 Jin-Yi Cai and Zhiguo Fu. Complexity classification of the eight-vertex model. *CoRR*, abs/1702.07938, 2017. [arXiv:1702.07938](https://arxiv.org/abs/1702.07938).
- 7 Jin-Yi Cai and Artem Govorov. Perfect matchings, rank of connection tensors and graph homomorphisms. In *Proceedings of the 2019 Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '19, pages 476–495, 2019. doi:10.1137/1.9781611975482.30.
- 8 Jin-Yi Cai and Tianyu Liu. Counting perfect matchings and the eight-vertex model. In *Proceedings of the 47th International Colloquium on Automata, Languages, and Programming*, ICALP '20, pages 23:1–23:19, 2020. [arXiv:1904.10493](https://arxiv.org/abs/1904.10493).
- 9 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.
- 10 Jin-Yi Cai, Tianyu Liu, Pinyan Lu, and Jing Yu. Approximability of the eight-vertex model, 2018. [arXiv:1811.03126](https://arxiv.org/abs/1811.03126).
- 11 Jin-Yi Cai, Pinyan Lu, and Mingji Xia. The complexity of complex weighted Boolean #CSP. *Journal of Computer and System Sciences*, 80(1):217–236, 2014. doi:10.1016/j.jcss.2013.07.003.
- 12 Martin Dyer, Alan Frieze, and Ravi Kannan. A random polynomial-time algorithm for approximating the volume of convex bodies. *J. ACM*, 38(1):1–17, January 1991. doi:10.1145/102782.102783.
- 13 Martin Dyer and David Richerby. An effective dichotomy for the counting constraint satisfaction problem. *SIAM Journal on Computing*, 42(3):1245–1274, 2013. doi:10.1137/100811258.
- 14 Michael Freedman, László Lovász, and Alexander Schrijver. Reflection positivity, rank connectivity, and homomorphism of graphs. *Journal of the American Mathematical Society*, 20:37–51, 2007. doi:10.1090/S0894-0347-06-00529-7.
- 15 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.
- 16 Mark Jerrum. *Counting, Sampling and Integrating: Algorithm and Complexity*. Birkhäuser, Basel, 2003.
- 17 Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM Journal on Computing*, 18(6):1149–1178, 1989. doi:10.1137/0218077.
- 18 Mark R. Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43(Supplement C):169–188, 1986. doi:10.1016/0304-3975(86)90174-X.
- 19 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.
- 20 W.B.Raymond Lickorish. *An Introduction to Knot Theory*. Springer New York, 1997. URL: https://books.google.com/books?id=PhHhw_kRvewC.
- 21 Aneesur Rahman and Frank H. Stillinger. Proton distribution in ice and the kirkwood correlation factor. *The Journal of Chemical Physics*, 57(9):4009–4017, 1972. doi:10.1063/1.1678874.
- 22 Alistair Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1:351–370, 1992.
- 23 Olav F. Syljuåsen and M. B. Zvonarev. Directed-loop monte carlo simulations of vertex models. *Phys. Rev. E*, 70:016118, July 2004. doi:10.1103/PhysRevE.70.016118.
- 24 A. Yanagawa and J.F. Nagle. Calculations of correlation functions for two-dimensional square ice. *Chemical Physics*, 43(3):329–339, 1979. doi:10.1016/0301-0104(79)85201-5.
- 25 Mihalis Yannakakis. Node- and edge-deletion NP-complete problems. In *Proceedings of the 10th Annual ACM Symposium on Theory of Computing*, STOC '78, pages 253–264, 1978. doi:10.1145/800133.804355.