

**APPROXIMATE COMPLEXITY IN STATISTICAL MECHANICS:
COUNTING AND SAMPLING IN THE SIX- AND EIGHT-VERTEX MODELS**

by

Tianyu Liu

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Computer Sciences)

at the

UNIVERSITY OF WISCONSIN–MADISON

2020

Date of final oral examination: 05/01/2020

The dissertation is approved by the following members of the Final Oral Committee:

Jin-Yi Cai, Professor, Computer Sciences

Eric Bach, Professor, Computer Sciences

Shuchi Chawla, Professor, Computer Sciences

Vadim Gorin, Assistant Professor, Mathematics

© Copyright by Tianyu Liu 2020

All Rights Reserved

ACKNOWLEDGMENTS

First and foremost, I would like to express my heartfelt gratitude to my advisor, Jin-Yi Cai. During the past six years, I have been guided with expertise, encouraged with enthusiasm, tolerated with patience, and nurtured with care. His knowledge and vision helped me explore theoretical computer science and gain confidence as a researcher, and his zeal and perseverance always reminded me to stay humble as a person. Having had the opportunity to learn from and work with him is my greatest privilege.

I am grateful for having collaborated with many other brilliant people. My sincere thanks are extended to Pinyan Lu for his guidance and support during my doctoral study. The research work in this dissertation has largely benefited from his insights. It was a great pleasure working and exchanging ideas with Jing Yu, who also contributed to several results presented in this dissertation. I must also thank Surajit Chaudhuri, Vivek Narasayya, and Chi Wang for offering me the internship opportunity at Microsoft Research. The internship provided me with a unique opportunity to apply my knowledge to a field I had been less acquainted with.

I sincerely thank Eric Bach, Shuchi Chawla, and Vadim Gorin for being on my committee. Eric and Shuchi supported me generously whenever I needed help from them. I am appreciative of the discussion with Vadim on my research work, and I wish that he came to Madison earlier. In addition, I am indebted to numerous discussions with Zhiguo Fu, Artem Govorov, and Shuai Shao, both about research and about other things in life.

Of special importance during my final year was the Cisco Distinguished Graduate Fellowship, without which this dissertation could have taken longer to complete. Thanks also to the financial support from NSF CCF-1714275 and from the Department of Computer Sciences at UW–Madison.

My family provided me with encouragement and patience throughout the completion of this dissertation. I owe them for their unconditional love and support. I also thank all my friends, who have helped me over the years and made my spare time enjoyable.

CONTENTS

Contents ii

List of Tables iv

List of Figures v

Abstract vii

1 Introduction 1

1.1 *The six- and eight-vertex models* 1

1.2 *The even-coloring model* 8

1.3 *Holographic transformations* 9

2 On General 4-regular Graphs (the Six-Vertex Model) 13

2.1 *Closure properties* 16

2.2 *An FPRAS via Markov chains Monte Carlo* 20

2.3 *NP-hardness* 30

3 On General 4-regular Graphs (the Eight-Vertex Model) 37

3.1 *Closure properties* 40

3.2 *An FPRAS via Markov chains Monte Carlo* 59

3.3 *#PERFECTMATCHINGS-easiness* 72

3.4 *#PERFECTMATCHINGS-hardness* 83

3.5 *NP-hardness* 94

4 On Planar and on Bipartite 4-regular Graphs 101

4.1 *An FPRAS on planar graphs* 104

4.2 *An FPRAS on bipartite graphs* 108

5 On the Square Lattice 112

5.1 *Torpid mixing for the ferroelectric six-vertex model* 116

5.2 *Torpid mixing for the anti-ferroelectric six-vertex model* 121

5.3 *The even-coloring model revisited* 130

5.4 *An abstract polymer model* 135

5.5 *Convergence of the cluster expansion* 139

5.6 *An FPTAS at low temperatures* 144

5.7 *Extension to the torus* 147

Bibliography 151

LIST OF TABLES

3.1	Map from eight local configurations to signed pairings.	42
3.2	An even orientation and its quantum decomposition into acp 's.	42
3.3	Another even orientation and its quantum decomposition into acp 's.	43
3.4	Approximation complexity of the eight-vertex model with $(a, b, c, d) \notin \mathcal{Y}$	83
4.1	Elements of $\langle M_Z^{\text{Pl}}, M_{HZ}^{\text{Pl}} \rangle$ and preimages of $\mathcal{Y} = \mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD}$ under corresponding maps. A substantial subregion of each preimage admits FPRAS on planar 4-regular graphs. The last column lists the approximation complexity of the eight-vertex model on general 4-regular graphs in the corresponding preimage regions.	108
4.2	Elements of $\langle M_Z^{\text{Bi}}, M_{HZ}^{\text{Bi}} \rangle$ and preimages of $\mathcal{Y} = \mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD}$ under corresponding maps. A substantial subregion of each preimage admits FPRAS on bipartite 4-regular graphs. The last column lists the approximation complexity of the eight-vertex model on general 4-regular graphs in the corresponding preimage regions.	111

LIST OF FIGURES

1.1	Valid configurations of the six- and eight-vertex models.	1
1.2	A Venn diagram of the approximation complexity of the eight-vertex model on general 4-regular graphs.	6
1.3	Valid configurations of the even-coloring model.	9
2.1	Phase diagram and complexity diagram of the zero-field six-vertex model.	13
2.2	Pairings at a degree 4 vertex.	17
2.3	A 4-ary construction using vertices with constraint function in $\mathcal{X}$	18
2.4	Examples of the states in the directed-loop algorithm.	21
2.5	A vertex of degree-2 in η	27
2.6	The correspondence between configurations before and after deleting two vertices with $(=_2)$	29
2.7	A construction implementing a single edge in an independent set.	31
2.8	A locking device implementing a degree three vertex and its incident edges.	32
2.9	2-chain.	32
3.1	(Unsigned) pairings at a degree 4 vertex.	41
3.2	A 4-ary construction in the eight-vertex model.	45
3.3	Possible ways of deleting a vertex. The vertex (not explicitly shown) at the center of part (a) is removed in part (b) and (c).	51
3.4	Possible ways of deleting two vertices with intersections of type $\neg\vdash$	54
3.5	Examples of the states in the directed-loop algorithm.	60
3.6	The correspondence between configurations before and after cutting open two edges.	64
3.7	A vertex of degree-2 in η	71
3.8	4-ary matchgates.	75
3.9	A geometric argument.	80

3.10	2-chain.	86
3.11	A four-way connection implementing a single edge in 3-MAX CUT.	95
3.12	A locking device implementing a vertex of degree 3 in 3-MAX CUT.	96
3.13	A 4-ary construction that amplifies the maximum among a, b, c, d	96
3.14	Modifying the four-way connection for the case when $d > a + b + c$	99
4.1	A proper 2-coloring of the faces of a planar 4-regular graph and its canonical orientation τ	104
4.2	An even orientation τ' and its corresponding even coloring.	105
4.3	Local configurations at one type of vertices in a planar graph.	105
4.4	Local configurations at the other type of vertices in a planar graph.	106
5.1	Some states in the state space of $\mathcal{M}_D$	119
5.2	Some states in the state space of $\mathcal{M}_G$	122
5.3	A topological argument.	123
5.4	An auxiliary diagonal lattice.	125
5.5	A fault line argument.	127
5.6	A step in $\mathcal{M}_G$	128
5.7	A state in $\mathcal{M}_D$	129
5.8	A state ς in the even-coloring model from two different states in the eight-vertex model under M_{III} and M_{IV} , respectively.	133
5.9	$\mathbb{Z}^2$ and L	136
5.10	The deformation rule at $L \setminus \mathbb{Z}^2$ -vertices.	136
5.11	An example of the correspondence between an even coloring and compatible polymers.	137
5.12	The three valid local configurations of a polymer at $\mathbb{Z}^2$ -vertices.	138
5.13	Two possible ways to extend a tree at a $\mathbb{Z}^2$ -vertex.	142
5.14	Three possible ways to extend a tree at an $L \setminus \mathbb{Z}^2$ -vertex.	142
5.15	A polymer configuration on $(\mathbb{Z}/4\mathbb{Z})^2$ that does not correspond to any even-colorings.	147

ABSTRACT

The six- and eight-vertex models originate in statistical mechanics for crystal lattices with hydrogen bonds. The first such model was introduced by Linus Pauling in 1935 to account for the residual entropy of water ice. The family of models not only are among the most extensively studied topics in physics, but also have fascinated chemists, mathematicians, theoretical computer scientists, and others, with thousands of papers studying their properties and connections to other fields. In this dissertation, we study the computational complexity of approximately counting and sampling in the six- and eight-vertex models on various classes of underlying graphs.

First, we study the approximability of the partition function on general 4-regular graphs, classified according to the parameters of the models. Our complexity results conform to the phase transition phenomenon from physics due to the change in temperature. We introduce a quantum decomposition of the six- and eight-vertex models and prove a set of closure properties in various regions of the parameter space. 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 in some parameter space in the high temperature regime. In some other parameter space in the high temperature regime, we prove that the problem is (at least) as hard as approximately counting perfect matchings, a central open problem in this field. We also show that the six- and eight-vertex models are NP-hard to approximate in the whole low temperature regime on general 4-regular graphs.

We then study the six- and eight-vertex models on more restricted classes of 4-regular graphs, including planar graphs and bipartite graphs. We give the first polynomial time approximation algorithm for the partition function in the low temperature regime on planar and on bipartite graphs. Our results show that the six- and eight-vertex models are the first problems with the provable property that while NP-hard to approximate on general graphs (even #P-hard for planar graphs in exact complexity), they possess efficient approximation schemes on both bipartite graphs and planar graphs in substantial regions of the parameter space.

Finally, we study the square lattice six- and eight-vertex models. We prove that natural Markov

chains for these models are mixing torpidly in the low temperature regime. Moreover, we give the first efficient approximate counting and sampling algorithms for the six- and the eight-vertex models on the square lattice at sufficiently low temperatures.

1 INTRODUCTION

1.1 The six- and eight-vertex models

The six-vertex model originates in statistical mechanics for crystal lattices with hydrogen bonds. It is a vertex model family with various parameter settings. Classically it is defined on a planar lattice region where each vertex of the lattice is connected by an edge to four nearest neighbors. States of the six-vertex model are orientations of the edges on the underlying undirected graph satisfying the “ice rule” – every vertex has two incoming edges and two outgoing edges, i.e., they are *Eulerian orientations*. The name of six-vertex model comes from the fact that there are six ways of arranging directions of the edges around a vertex (see the first six configurations in Figure 1.1).

Its generalization, the *eight-vertex model*, has two more permissible local configurations, namely those in Figure 1.1-7 and Figure 1.1-8 – hence the name eight-vertex model. The states are *even orientations*, i.e. those with an even number of arrows into (and out of) each vertex.

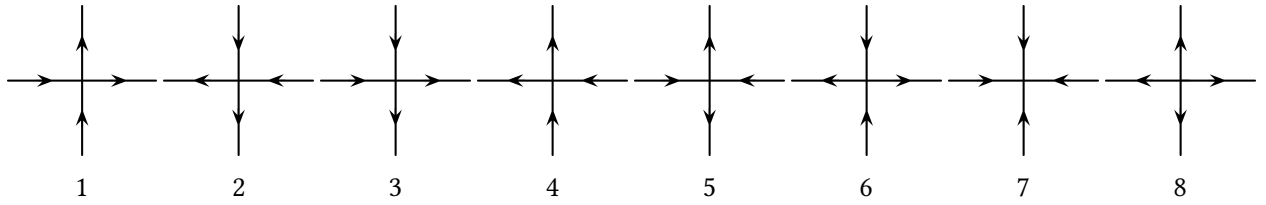


Figure 1.1: Valid configurations of the six- and eight-vertex models.

In general, the eight configurations 1 to 8 in Figure 1.1 are associated with eight possible energy parameters $\epsilon_1, \dots, \epsilon_8$. On G , a finite region in $\mathbb{Z}^2$ (with certain *boundary conditions*) or a 4-regular graph where the four incident edges of each vertex are labeled from 1 to 4, the *partition function* is defined as

$$Z_{8V}(G) = \sum_{\tau \in \mathcal{O}_e(G)} \exp(-\beta(n_1 \epsilon_1 + \dots + n_8 \epsilon_8)) \quad (1.1.1)$$

where $\mathcal{O}_e(G)$ is the set of all even orientations of G , n_i is the number of vertices in type i in G ($1 \leq i \leq 8$, locally depicted as in Figure 1.1) under an even orientation $\tau \in \mathcal{O}_e(G)$, $\beta > 0$ is the inverse temperature. Low (respectively, high) temperatures refer to large (respectively, small) β

values.

It is clear from (1.1.1) that Z_{8V} is a function of $\exp(-\beta\epsilon_i)$ for $1 \leq i \leq 8$. By physical considerations, the weight of a configuration remains unchanged if all arrows are flipped, assuming there is no external electric field. In this case we have $\epsilon_1 = \epsilon_2$, $\epsilon_3 = \epsilon_4$, $\epsilon_5 = \epsilon_6$, $\epsilon_7 = \epsilon_8$. Thus we can write

$$\begin{aligned} \exp(-\beta\epsilon_1) &= \exp(-\beta\epsilon_2) = a, & \exp(-\beta\epsilon_3) &= \exp(-\beta\epsilon_4) = b, \\ \exp(-\beta\epsilon_5) &= \exp(-\beta\epsilon_6) = c, & \exp(-\beta\epsilon_7) &= \exp(-\beta\epsilon_8) = d. \end{aligned} \quad (1.1.2)$$

This complementary invariance is known as the *arrow reversal symmetry* or the *zero field assumption*. Therefore, the partition function can be alternatively written as

$$Z_{8V}(G; a, b, c, d) = \sum_{\tau \in \mathcal{O}_e(G)} a^{n_1+n_2} b^{n_3+n_4} c^{n_5+n_6} d^{n_7+n_8}. \quad (1.1.3)$$

In physics $a, b, c, d > 0$ in the (strict) eight-vertex model. In most part of this dissertation we allow nonnegative weights (except for Chapter 4 where we allow negative weights), and by setting $d = 0$ we get the six-vertex model. In this case, we write $Z_{6V}(a, b, c)$ instead of $Z_{8V}(a, b, c, 0)$.

The six-vertex model was first introduced by Linus Pauling [Pau35] in 1935 to describe the properties of ice. In 1967, Elliot Lieb [Lie67c, Lie67a, Lie67b] famously showed that, for parameters $(a, b, c) = (1, 1, 1)$ on the square lattice graph, as the side N of the square approaches ∞ , the value of the “partition function per vertex” $W = Z^{1/N^2}$ approaches $\left(\frac{4}{3}\right)^{3/2} \approx 1.5396007 \dots$ (this is called Lieb’s square ice constant). This is called an “exact solution” of the model, and is considered a triumph. After that, “exact solutions” for other lattice type graphs (such as [Sut67, FW70]) have been obtained in the limiting sense.

The eight-vertex model was first introduced in 1970 by Sutherland [Sut70], and Fan and Wu [FW70] as a generalization of the six-vertex model for certain more desirable properties on the square lattice. However, in contrast to the six-vertex model which has been “exactly solved” under various parameter settings and external fields [Lie67c, Lie67a, Lie67b, Sut67, FW70], the eight-vertex model was “exactly solved” only in the zero-field case [Bax71, Bax72]. This model is

enormously expressive even in the zero-field setting: its special case when $d = 0$, the zero-field six-vertex model, 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 it [Bax82]. Therefore, insight to the eight-vertex model is also much sought-after in statistical physics.

For half a century, the six-vertex model and the eight-vertex model have fascinated physicists, chemists, mathematicians and others, with thousands of papers studying their properties and connections to other fields. This includes the proof of the alternating sign matrix conjecture [Zei96, Kup96], the discovery of the arctic curve phenomena [SZ04, AR05], and under different parameter settings the equivalence between the six-vertex model on the one hand, and proper three-colorings [Bax82], the Tutte polynomial [Tut54, Ver88], and tilings of the Aztec diamond [Elo99] on the other.

Not until recently did we fully understand the exact computational complexity of the six- and eight-vertex models (for all eight parameters) on 4-regular graphs [CFX18, CFS17, CF17]. This is studied in the context of a classification program for the complexity of counting problems [CC17], where the six- and 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 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*. Previously, Mihail and Winkler first proved that computing the number of unweighted Eulerian orientations is #P-complete over general graphs [MW96]; Huang and Lu proved that it remains #P-complete for even degree regular (but not necessarily planar) graphs [HL16]; Guo and Williams improved it to planar 4-regular graphs [GW13]. The latter is equivalent to computing the partition

function of the six-vertex model on planar graphs with the parameter setting $(1, 1, 1)$.

In this dissertation, we study the approximate complexity of counting and sampling in the six- and eight-vertex models on various classes of graphs. An interesting discovery is that most of our results conform to the *phase transition* phenomena in physics — in statistical mechanics setting, the square lattice six- and eight-vertex models exhibit different macroscopic behaviors (in the thermodynamic limit) under different parameters a, b, c, d due to the change in (inverse) temperatures (see Baxter’s book [Bax82] for details). There are five phases classified into three types, which can be roughly thought of as either one of the parameters a, b, c, d dominates or none does^{*}:

- Disordered: $a, b, c, d < \frac{1}{2}(a + b + c + d)$,
- Ferroelectric: $a > b + c + d$, or $b > a + c + d$,
- Anti-ferroelectric: $c > a + b + d$, or $d > a + b + c$.

Among these phases, the disordered phase represents the *high temperature* regime and the *ordered phases* (ferroelectric and anti-ferroelectric phases) represent the *low temperature* regimes. For any parameters $\epsilon_1, \epsilon_3, \epsilon_5, \epsilon_7$ (when the minimum value is unique), there exists a critical inverse temperature $\beta_c = \beta_c(\epsilon_1, \epsilon_3, \epsilon_5, \epsilon_7)$ when the weights a, b, c, d are on the phase transition threshold between the disordered phase ($\beta < \beta_c$) and the corresponding ordered phase ($\beta > \beta_c$).

In this dissertation, we adopt the following notations assuming $a, b, c, d \geq 0$.

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

We note that $\mathcal{Y} \subset \mathcal{X}$ and $\mathcal{Z} \subset \mathcal{X}$. In physics terms, $\mathcal{X}$ (almost, except for the boundaries) corresponds to the disordered phase and its complement $\overline{\mathcal{X}}$, consisting of 4 disjoint regions in which one of (a, b, c, d) dominates, correspond to the ordered phases.

^{*}For the six-vertex model, there are four phases as $d = 0$ and thus the phase with $d > a + b + c$ does not exist.

The first class of graphs we study is general 4-regular graphs. Mihail and Winkler's pioneering work [MW96] gave the first *fully polynomial randomized approximation scheme (FPRAS)*[†] for the number of Eulerian orientations on a general graph. On 4-regular graphs, this corresponds to the single point $(1, 1, 1)$ in the disordered phase $\mathcal{X}$. In [CLL19, CLLY20, CL20], we studied the classification of the approximate complexity of the six- and eight-vertex models on general 4-regular graphs in terms of the parameters. In some parameter space, we proved that the problem admits an FPRAS via the method of *Markov chain Monte Carlo (MCMC)*; in some other parameter space, we proved that the problem is NP-hard to approximate via *approximation-preserving reductions*. In addition to the above regions, there exist parameter settings under which we related the problem to that of approximately counting perfect matchings on general graphs (denoted by $\#PERFECTMATCHINGS$ or $\#PM$). The approximate complexity of counting perfect matchings on general (not necessarily bipartite) graphs is a central open problem in this field. It is also called the dimer model in statistical physics. On a bipartite graph with n vertices, the number of perfect matchings is equal to the *permanent* of a $n \times n$ 0, 1 matrix. Jerrum, Sinclair, and Vigoda [JS89, JSV04] gave an FPRAS for estimating the permanent of an arbitrary $n \times n$ matrix with nonnegative entries via MCMC, hence solving the problem in the bipartite case. This was a major achievement in approximate counting.

In particular, we showed the following results for general 4-regular graphs (see Figure 1.2):

- (1) there is an FPRAS for $Z_{8V}(a, b, c, d)$ if $(a, b, c, d) \in \mathcal{Y} \cap \mathcal{Z}$;
- (2) approximating $Z_{8V}(a, b, c, d)$ can be reduced to $\#PM$ if $(a, b, c, d) \in \mathcal{Z}$;
- (3) approximating $Z_{8V}(a, b, c, d)$ is at least as hard as $\#PM$ if $(a, b, c, d) \notin \mathcal{Y}$; and
- (4) approximating $Z_{8V}(a, b, c, d)$ for $(a, b, c, d) \in \overline{\mathcal{X}}$ is NP-hard.

[†]Suppose $f : \Sigma^* \rightarrow \mathbb{R}$ is a function mapping problem instances to real numbers. A *fully polynomial randomized approximation scheme (FPRAS)* [KL83] 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 $\Pr[(1 - \varepsilon)f(x) \leq Y \leq (1 + \varepsilon)f(x)] \geq \frac{3}{4}$.

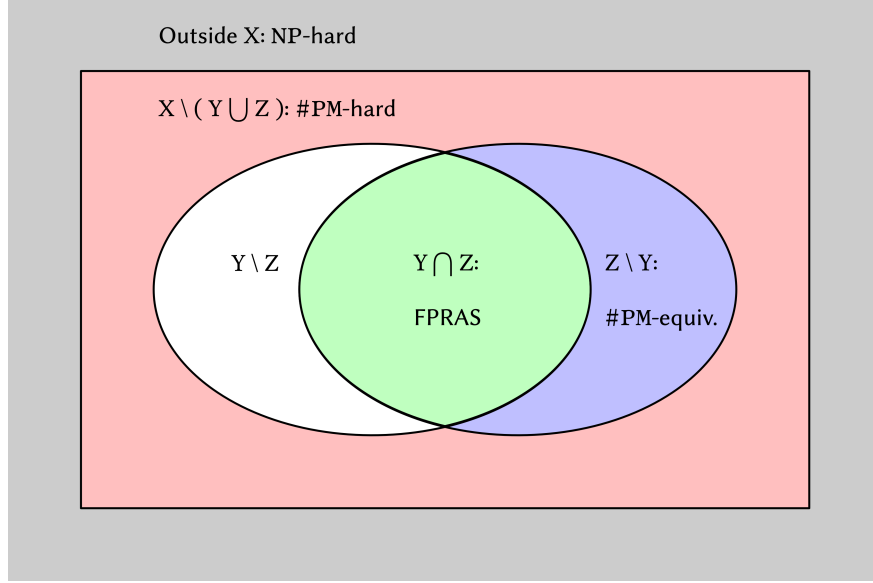


Figure 1.2: A Venn diagram of the approximation complexity of the eight-vertex model on general 4-regular graphs.

We also study the six- and eight-vertex models on more restricted classes of 4-regular graphs, including planar graphs and bipartite graphs. In [CLLY20], we showed that there is an FPRAS for planar 4-regular graphs in an additional subregion of $\overline{\mathcal{Y}} \cap \mathcal{Z}$, still via MCMC. Thus we 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. These are the first problems that exhibit this particular dichotomic behavior between the planar and the nonplanar settings in approximate counting.

So far, all previous positive results are confined within $\mathcal{Z}$, in particular within the disordered phase $\mathcal{X}$. In an unpublished work, we gave the first FPRAS for $Z_{8v}(a, b, c, d)$ for (a, b, c, d) outside $\mathcal{Z}$ and even in the ordered phases $\overline{\mathcal{X}}$ on planar and on bipartite graphs. This was done by simultaneously setting up two different kinds of relations between the partition functions of the eight-vertex model and a special edge-2-coloring model, called the *even-coloring model*. Our results showed that the eight-vertex model is the first problem with the provable property that while NP-hard to approximate on general graphs (even #P-hard for planar graphs in exact complexity), it possesses FPRAS on both bipartite graphs and planar graphs in substantial regions of its parameter space.

Finally we study the six- and eight-vertex models on square lattice graphs. We mentioned earlier that the square lattice six- and eight-vertex models have been widely studied in statistical mechanics as well as other disciplines. A central goal among these studies is to (at least approximately) count or sample states in the model according to the Gibbs measure and MCMC serves as the primary tool. A partial list includes [RS72, YN79, BN98, Elo99, SZ04, AR05, BCG16, KL17, LKV17, Dim18, LKRV18, KS18]. Despite significant effort, physicists were hampered by the inefficiency of such algorithms at low temperature regimes, as widely reported [BN98, SZ04, LKV17, LKRV18].

Observations that these natural Markov chains reach equilibrium fast in the high temperature regimes but extremely slowly in the low temperature regimes were partially (but rigorously) verified by theoretical computer scientists. Previous attempts in designing efficient approximate counting and sampling algorithms for the six-vertex and eight-vertex models are all by MCMC and success is restricted in (a subregion of) the disordered phase, i.e., the high temperature regime. In addition to [MW96] which counts unweighted Eulerian orientations on general graphs and our works mentioned above for not necessarily square lattice graphs, multiple Markov chains for the unweighted six-vertex model were proved to mix rapidly on rectangular regions of the square lattice with various boundary conditions [LRS01, RT00, GMP04]. These works relied on a correspondence between Eulerian orientations on the edges and proper three-colorings of the faces on the square lattice.

In the ordered phases, i.e., the low temperature regimes, Markov chains with local moves for the square lattice regions with free or periodic boundary conditions are believed to achieve equilibrium exponentially slowly (torpid mixing). When the temperature is sufficiently low, this was proved by Greenberg and Randall [GR10] for the eight-vertex model when d is large, and by Liu [Liu18] for the six-vertex model. Fahrback and Randall [FR19] improved some of the results in [Liu18] and obtain some new results. We recall that on general 4-regular graphs, approximating the partition function of the six-vertex and eight-vertex models in the low temperature regime are NP-hard [CLL19, CLLY20].

Although MCMC methods fail at low temperatures, there is still hope to approximately count

and sample efficiently. In another unpublished work, we gave the first efficient approximate counting and sampling algorithms for the six-vertex model and the eight-vertex model on finite regions of the square lattice $\mathbb{Z}^2$ with free or periodic boundary conditions in all the ordered phases at sufficiently low temperatures. Our work uses cluster expansions, an old tool in statistical physics, combined with Barvinok’s method to approximate partition functions via Taylor expansions [Bar16, PR17], a recent approach pioneered by Helmuth, Perkins, and Regts [HPR19].

This dissertation is organized as follows. In the rest of Chapter 1, we introduce the even-coloring model (Section 1.2) and *holographic transformations* (Section 1.3), both playing an important role in multiple reductions and algorithms in later chapters. In Chapter 2, we present the approximate complexity of the six-vertex models on general 4-regular graphs. In Chapter 3, we present the approximate complexity of the eight-vertex models on general 4-regular graphs. In Chapter 4, we give approximation algorithms for planar graphs and for bipartite graphs. In Chapter 5, we prove torpid mixing of Markov chains and present fast approximation schemes for the square lattice six- and eight-vertex models in the low temperature regime. Results in Chapter 2 are from [CLL19], a joint work with Jin-Yi Cai and Pinyan Lu. Results in Chapter 3 are from [CLLY20, CL20]. [CLLY20] is joint with Jin-Yi Cai, Pinyan Lu, and Jing Yu. [CL20] is joint with Jin-Yi Cai. Results in Chapter 4 are from a submitted work joint with Jin-Yi Cai. Results in Chapter 5 are from [Liu18] and a work in submission joint with Jin-Yi Cai.

1.2 The even-coloring model

We introduce the following special edge-2-coloring model on 4-regular graphs which we call the *even-coloring model*: a valid configuration of this model assigns either *green* or *red* to every edge such that the number of green edges incident to any vertex is even (zero, two, or four). Similar to the eight-vertex model, there are also eight valid local configurations around a vertex (shown in Figure 1.3), and configurations 1 to 8 in Figure 1.3 can be associated with rational energy parameters $\epsilon'_1, \dots, \epsilon'_8$ respectively. Denote the set of these eight local configurations by S_{EC} .

Similarly, denote the set of eight local configurations of the eight-vertex model in Figure 1.1 by S_{8V} .

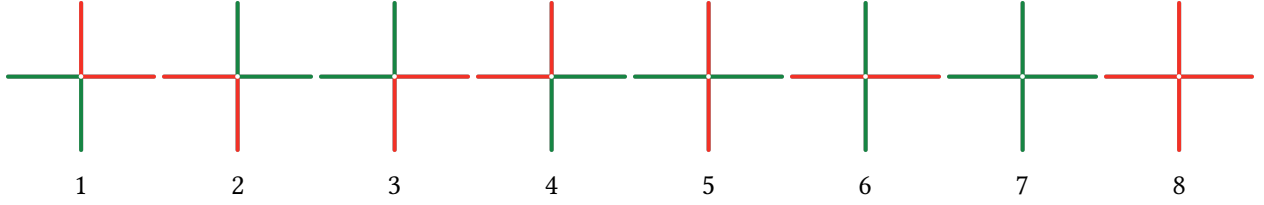


Figure 1.3: Valid configurations of the even-coloring model.

To set up a correspondence between the even-coloring model and the eight-vertex model satisfying arrow reversal symmetry, we consider the even-coloring model with weights that satisfy the *color reversal symmetry*. That is, the weight of a local configuration at a vertex remains unchanged if the color on every incident edge is changed. In this case we write

$$\begin{aligned} \exp(-\beta\epsilon'_1) &= \exp(-\beta\epsilon'_2) = \mu, & \exp(-\beta\epsilon'_3) &= \exp(-\beta\epsilon'_4) = \nu, \\ \exp(-\beta\epsilon'_5) &= \exp(-\beta\epsilon'_6) = \phi, & \exp(-\beta\epsilon'_7) &= \exp(-\beta\epsilon'_8) = \psi. \end{aligned} \quad (1.2.1)$$

In this dissertation, we make this assumption and further assume that $\mu, \nu, \phi, \psi \geq 0$. Given a 4-regular graph G , we label four incident edges of each vertex from 1 to 4. The *partition function* of the even-coloring model with parameters (μ, ν, ϕ, ψ) on G is defined as

$$Z_{\text{EC}}(G; \mu, \nu, \phi, \psi) = \sum_{\zeta \in \mathcal{C}_e(G)} \mu^{n_1+n_2} \nu^{n_3+n_4} \phi^{n_5+n_6} \psi^{n_7+n_8}, \quad (1.2.2)$$

where $\mathcal{C}_e(G)$ is the set of all even-colorings 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 1.3) under the even-coloring $\zeta \in \mathcal{C}_e(G)$.

1.3 Holographic transformations

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 ($w \rightarrow v$) 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 a *0-1 labeling* on G' , namely $\sigma : E' \rightarrow \{0, 1\}$, 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 DISQUALITY constraint, denoted by $(\neq_2)$. The values of a 4-ary *constraint function* 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 constraint function 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 locally index the left, down, right, and up edges incident to v by 1, 2, 3, and 4, respectively according to Figure 1.1. Thus computing the partition function $Z_{8V}(G; a, b, c, d)$ is equivalent to evaluating

$$\sum_{\sigma : E' \rightarrow \{0,1\}} \prod_{u \in U_E} (\neq_2)(\sigma|_{E'(u)}) \prod_{u \in U_V} f(\sigma|_{E'(u)}),$$

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:

$$Z_{8V}(G; a, b, c, d) = \text{Holant}(G'; \neq_2 \mid f)$$

where we use $\text{Holant}(H; g \mid f)$ to denote the Holant sum $\sum_{\sigma : E \rightarrow \{0,1\}} \prod_{u \in U} g(\sigma|_{E(u)}) \prod_{u \in V} f(\sigma|_{E(u)})$ on a bipartite graph $H = (U, V, E)$ for the Holant problem $\text{Holant}(g \mid f)$. Each vertex in U (or V) is assigned the constraint function g (or f , respectively). The constraint function g is written as a row vector, whereas the constraint function f is written as a column vector, both as truth tables. (See [CC17] for more on Holant problems.) The following theorem says that an invertible holographic transformation does not change the complexity of the Holant problem in the bipartite setting.

Theorem 1.3.1 ([Val08]). Suppose $T \in \mathbb{C}^2$ is an invertible matrix. Let $d_1 = \text{arity}(g)$ and $d_2 = \text{arity}(f)$. Define $g' = g (T^{-1})^{\otimes d_1}$ and $f' = T^{\otimes d_2} f$. Then for any bipartite graph H , $\text{Holant}(H; g \mid f) = \text{Holant}(H; g' \mid f')$.

We denote $\text{Holant}(G; f) = \text{Holant}(G'; =_2 \mid f)$. For the even-coloring model, if we view a green-red edge coloring by a 0-1 assignment to the edges such that an edge e is assigned 0 if it is colored green and assigned 1 if it is colored red, then the partition function of the even-coloring model $Z_{\text{EC}}(G; \mu, \nu, \phi, \psi)$ is exactly the value of the Holant problem $\text{Holant}\left(G; \begin{bmatrix} \psi & 0 & 0 & \mu \\ 0 & \nu & \phi & 0 \\ 0 & \phi & \nu & 0 \\ \mu & 0 & 0 & \psi \end{bmatrix}\right)$.

The following two lemmas show that the eight-vertex model and the even-coloring model are connected via suitable holographic transformations in unexpected ways as Holant problems.

Lemma 1.3.2. Let G be a 4-regular graph and let $M_Z = \frac{1}{2} \begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$. Then $Z_{8V}(G; a, b, c, d) = Z_{\text{EC}}(G; \mu, \nu, \phi, \psi)$ where $\begin{bmatrix} \mu \\ \nu \\ \phi \\ \psi \end{bmatrix} = M_Z \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$.

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' ,

$$Z_{8V}(G; a, b, c, d) = \text{Holant}(G'; \neq_2 \mid f),$$

where f is the 4-ary signature with $M(f) = \begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix}$. 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 $Z = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$ 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 Theorem 1.3.1, by the Z -transformation, we get

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

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 1.3.3. *Let G be a 4-regular graph and let $M_{HZ} = \frac{1}{2} \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$. Then $Z_{8V}(G; a, b, c, d) = Z_{EC}(G; \mu, \nu, \phi, \psi)$ where $\begin{bmatrix} \mu \\ \nu \\ \phi \\ \psi \end{bmatrix} = M_{HZ} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$.*

Proof. For the eight-vertex model as a Holant problem $\text{Holant}(G'; \neq_2 \mid f)$, we perform a holographic transformation by the matrix $\frac{1}{2} \begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}$. We note that this is the composition of a Z -transformation and an H -transformation where $Z = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix}$ and $H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$, namely $\frac{1}{2} \begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix} = HZ$. Then

$$\begin{aligned} Z_{8V}(G; a, b, c, d) &= \text{Holant}(G'; \neq_2 \mid f) \\ &= \text{Holant}(G'; (\neq_2) \cdot ((HZ)^{-1})^{\otimes 2} \mid (HZ)^{\otimes 4} \cdot f) \\ &= \text{Holant}(G'; =_2 \mid (HZ)^{\otimes 4} f) \\ &= \text{Holant}(G; (HZ)^{\otimes 4} f). \end{aligned}$$

Here $(\neq_2) \cdot ((HZ)^{-1})^{\otimes 2} = (\neq_2) \cdot (Z^{-1})^{\otimes 2} \cdot (H^{-1})^{\otimes 2} = (=_2) \cdot (H^{-1})^{\otimes 2} = (=_2)$, because H is orthogonal.

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

2 ON GENERAL 4-REGULAR GRAPHS (THE SIX-VERTEX MODEL)

In this section we initiate a study toward a classification of the approximate complexity of the six-vertex model in terms of the parameters. Our results conform to phase transitions in physics.

Here we briefly describe the phenomenon of phase transition of the zero-field six-vertex model (see Baxter's book [Bax82] for more details). On the square lattice in the thermodynamic limit:

- (1) When $a > b + c$ (FE: ferroelectric phase) any finite region tends to be frozen into one of the two configurations where either all arrows point up or to the right (Figure 1.1-1), or all point down or to the left (Figure 1.1-2).
- (2) Symmetrically when $b > a + c$ (also FE) all arrows point down or to the right (Figure 1.1-3), or all point up or to the left (Figure 1.1-4).
- (3) When $c > a + b$ (AFE: anti-ferroelectric phase) configurations in Figure 1.1-5 and Figure 1.1-6 alternate.
- (4) When $c < a + b$, $b < a + c$, and $a < b + c$, the system is disordered (DO: disordered phase) in the sense that all correlations decay to zero with increasing distance; in particular on the dashed curve $c^2 = a^2 + b^2$ the model can be solved by Pfaffians exactly [FW70], and the correlations decay inverse polynomially, rather than exponentially, in distance. See Figure 2.1a.

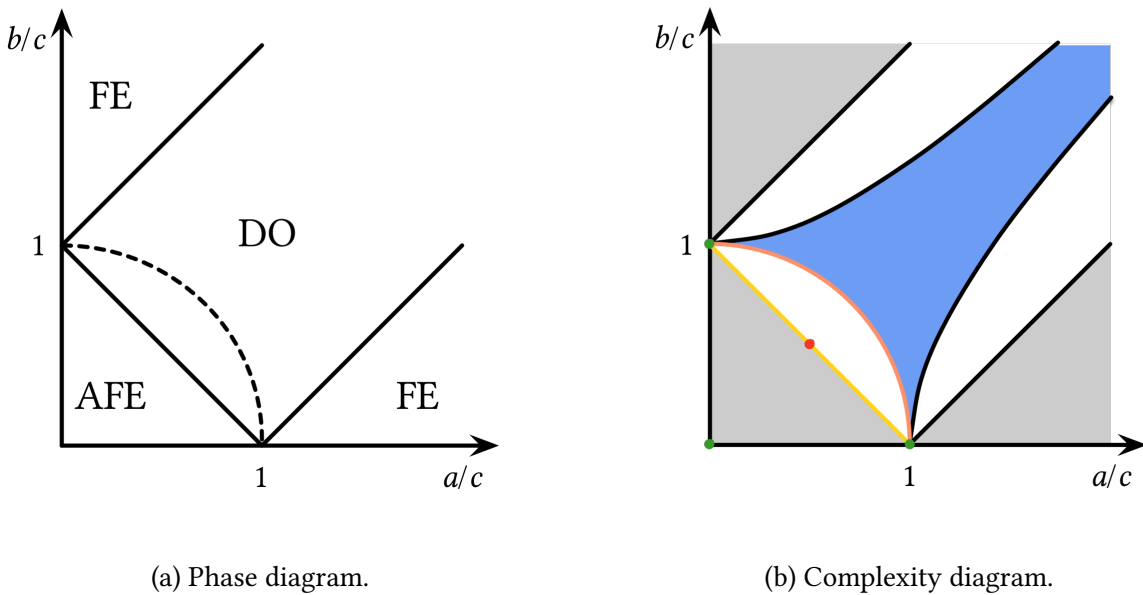


Figure 2.1: Phase diagram and complexity diagram of the zero-field six-vertex model.

In Figure 2.1b we have a corresponding complexity landscape, its main points are stated in:

Theorem 2.0.1. *There is an FPRAS for $Z_{6V}(G; a, b, c)$ if $a^2 \leq b^2 + c^2$, $b^2 \leq a^2 + c^2$, and $c^2 \leq a^2 + b^2$ (the blue region). There is no FPRAS for $Z_{6V}(G; a, b, c)$ if $a > b + c$ or $b > a + c$ or $c > a + b$ (the grey region), unless $RP = NP$.*

We obtain our FPRAS for $Z_{6V}(G; a, b, c)$ by designing a Markov chain and proving that it is rapidly mixing. The Markov chain is based on the *directed-loop algorithm*. This directed-loop algorithm is invented by Rahman and Stillinger [RS72] and is widely adopted in the literature (e.g., [YN79, BN98, SZ04]). The transitions of this algorithm are composed of creating, shifting, and merging of two “defects” on the edges. This is similar to *the worm process* introduced in [PS01] for the Ising model (see also [JS93]). Note that Markov chains that stay within the set of all Eulerian orientations, such as *Glauber dynamics*, are not irreducible for nonplanar 4-regular graphs (e.g. a torus). Therefore, introducing “defects” on the edges, i.e., considering “near-Eulerian” orientations, is indispensable. Moreover, the defects introduced in the directed-loop algorithm depict the *Bjerrum defects* that are observed to happen in real ice [BN98]. From this perspective, it would be rather artificial to insist that our Markov chain stay “inside” the set of Eulerian orientations, even in cases where such a chain is irreducible. There is no physical reason not to consider a wider variety, nor a mathematical one—as we shall show in this chapter that the first provable rapidly mixing Markov chain for a wider parameter space is based on a state space going beyond the set of Eulerian orientations. We give a formal description of the directed-loop algorithm in Section 2.2. The rapid mixing of this Markov chain is proved by bounding the *conductance* via a *flow argument* [JS89, DFK91, Sin92, Jer03]. The crucial ingredient in this proof is the closure properties we demonstrate in Section 2.1.

Our FPRAS result is actually stronger in that the FPRAS works even if different constraint functions from the blue region are assigned at different vertices. The blue region is a proper subset of the disordered phase. The point $(1, 1, 1)$ is contained in this region, which is the only previously known approximable case. The hardness part (the grey region) coincides with the FE/AFE phases. The three green points together with a point at infinity $((a, b, c) = (1, 1, 0))$ are exactly P-time

computable. All parameters belonging to the two axes ($a = 0$ and $b = 0$) and the orange curve $c^2 = a^2 + b^2$ are exactly P-time computable on planar graphs. Computing for the six-vertex model at $(1/2, 1/2, 1)$ (the red point) is equivalent to evaluating the Tutte polynomial $T(G; 3, 3)$ on planar graphs. Note that any 4-regular plane graph H is the medial graph of some plane graph G . The approximation complexity for the white region is unknown.

More structurally, 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. We use the term *a 4-ary construction* for a 4-regular graph Γ having 4 “external” edges, and consider all configurations on the edges of Γ where every vertex satisfies the ice rule and the arrow reversal symmetry. It turns out that this Γ defines a constraint function of arity 4 that also satisfies the ice rule and the arrow reversal symmetry. If we imagine the graph Γ , except the 4 external edges, is shrunk to a single point, then a 4-ary construction can be viewed as a virtual vertex with the six-vertex model parameter (a', b', c') for some a', b', c' .

In Theorem 2.1.1 of Section 2.1, we prove that the set of 4-ary constraint functions lying in the combined region of blue and white (this is the same as the DO region in Figure 2.1a) is closed under 4-ary constructions. In Theorem 2.1.2 we prove that the set of 4-ary constraint functions lying on the yellow line (phase transition threshold for AFE and DO) is closed under *planar* 4-ary constructions.

Theorem 2.1.1 is important not only for its revelation of a structural difference between parameter settings of the six-vertex model on the two sides of the phase transition threshold, but also for its crucial role in giving an FPRAS in Section 2.2. It is used to upper bound the ratio of near-Eulerian orientations over Eulerian orientations (valid six-vertex configurations). By Lemma 2.2.2, this shows that the directed-loop algorithm is rapidly mixing. Also, we use Theorem 2.1.1 to approximately sample the valid six-vertex configurations, which is necessary for our FPRAS to work.

Our FPRAS also has implications for counting weighted sum of directed Eulerian partitions (partition of edges of G into directed edge-disjoint circuits). A special case is an FPRAS for this

weighted sum when the weight of ---| is at least $\sqrt{2} - 1$ (more on the connection between directed Eulerian partitions and the three types of pairings $\text{---}\swarrow$, $\text{---}\nearrow$, and ---| can be found in Section 2.1).

The NP-hardness of approximation in the grey region (coincide with the ferroelectric/anti-ferroelectric phases) is given in Section 2.3 by an approximation-preserving reduction from computing the maximum independent set on 3-regular graphs. These are the first inapproximability results for the six-vertex model on general 4-regular graphs.

2.1 Closure properties

Theorem 2.1.1. *Consider a 4-ary construction using constraint functions from $\mathcal{X}^\ddagger$. Let f be the resulting constraint function of the 4-ary construction. Then $f \in \mathcal{X}$. In other words, the set of constraint functions in $\mathcal{X}$ is closed under 4-ary constructions.*

Theorem 2.1.2. *The set of constraint functions in $\{(a, b, c) \mid c = a + b\}$ is closed under 4-ary plane constructions.*

Before proving Theorem 2.1.1 and Theorem 2.1.2, we introduce another view of the six-vertex model. A valid configuration in the six-vertex model, i.e. a weighted Eulerian orientation, can also be viewed as a combination of weighted directed Eulerian partitions. An *Eulerian partition* of a graph G is a partition of the edges of G into edge-disjoint *circuits* (in which vertices may repeat whereas edges cannot). A *directed Eulerian partition* is an Eulerian partition where every edge-disjoint circuit takes one of the two cyclic orientations. Let $G = (V, E)$ be a 4-regular graph and v be a vertex of G . Let e_1, e_2, e_3, e_4 be the four edges incident to v . A *pairing* ϱ at v is a partition of $\{e_1, e_2, e_3, e_4\}$ into pairs. There are exactly three distinct pairings at v (Figure 2.2) which we denote by three special symbols: $\text{---}\swarrow$, $\text{---}\nearrow$, ---| , respectively. An Eulerian partition of G can be uniquely determined by a family of pairings $\varphi = \{\varrho_v\}_{v \in V}$, where $\varrho_v \in \{\text{---}\swarrow, \text{---}\nearrow, \text{---|}\}$ is a pairing

[‡]As a special case of the eight-vertex model where $d = 0$, the six-vertex model has $\mathcal{X} = \mathcal{Y} = \{(a, b, c) \mid a \leq b + c, \ b \leq a + c, \ c \leq b + a\}$. We use $\mathcal{X}$ to denote this region in this chapter.

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.

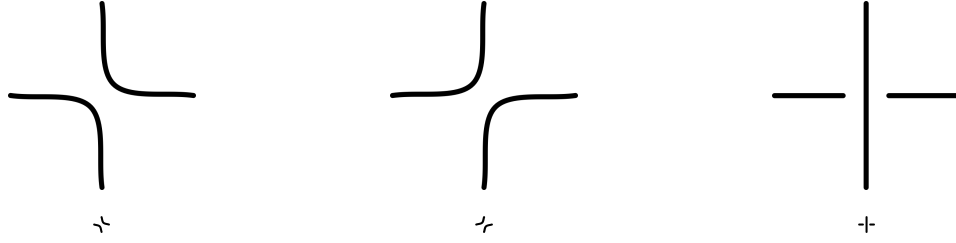


Figure 2.2: Pairings at a degree 4 vertex.

For any vertex v in a valid configuration τ of the six-vertex model (where ice rule is satisfied), incoming edges can be paired with outgoing edges in exactly two ways, corresponding to two of the three pairings at v . For example, the configuration in Figure 1.1-1 of the six-vertex model has two underlying pairings, $\nearrow$ and $\dashv$. Therefore, τ can be decomposed into $2^{|V|}$ distinct directed Eulerian partitions denoted by $\Phi(\tau)$. Since no two Eulerian orientations share one directed Eulerian partition and every directed Eulerian partition corresponds to a particular Eulerian orientation, the map from six-vertex configurations to directed Eulerian partitions is 1-to- $2^{|V|}$, non-overlapping, and surjective. Define w to be a function assigning a weight to every pairing at every vertex and let the weight $\tilde{w}(\varphi)$ of an Eulerian partition φ , undirected or directed, be the product of weights at each vertex. In particular, when w is defined such that $\begin{cases} w(\nearrow) = \frac{-a+b+c}{2} \\ w(\searrow) = \frac{a-b+c}{2} \\ w(\dashv) = \frac{a+b-c}{2} \end{cases}$, or equivalently $\begin{cases} a = w(\nearrow) + w(\dashv) \\ b = w(\nearrow) + w(\searrow) \\ c = w(\searrow) + w(\dashv) \end{cases}$, for every vertex with constraint matrix $\begin{bmatrix} & b & c & a \\ & c & b & \end{bmatrix}$, then the weight of a six-vertex model configuration τ is equal to $\sum_{\varphi \in \Phi(\tau)} \tilde{w}(\varphi)$, by expressing a product of sums as a sum of products.

The connection between Eulerian orientations and Eulerian partitions on 4-regular graphs has been explored in particular parameter settings for the six-vertex model. Las Vergnas [Ver88] demonstrated a special case for plane graphs: the number of directed *non-intersecting* Eulerian partitions is equal to the number of Eulerian orientations with weight 2 on every *saddle* configuration (Figure 1.1-5 and Figure 1.1-6), which is the six-vertex model at $(1, 1, 2)$. Jaeger [Jae90] proposed a graph polynomial called *transition polynomial* as a generalization of weighted Eulerian partitions, and related it with weighted Eulerian orientations. The idea of unweighted directed

Eulerian partitions was implicitly used in Mihail and Winkler’s paper [MW96] to approximate the number of unweighted Eulerian orientations, where they also adopted the notion of *pairings*. Here we give a general correspondence between weighted Eulerian orientations and weighted Eulerian partitions for the six-vertex model in the proof of Theorem 2.1.1 and Theorem 2.1.2. In particular we establish a weight-preserving 1-to- $2^{|V|}$, non-overlapping, and surjective mapping for Eulerian orientations in $\mathcal{X}$ and nonnegatively weighted Eulerian partitions. This “quantum graph” perspective is at the heart of the proof in this chapter.

Proof of Theorem 2.1.1. For the constraint function f of a 4-ary construction using vertices with constraint functions in $\mathcal{X}$ (Figure 2.3a), we first show that its constraint matrix must be of the form $\begin{bmatrix} & b' & c' & a' \\ & c' & b' & \\ a' & & & \end{bmatrix}$. This is to say that the ice rule and the arrow reversal symmetry are still satisfied. First, f still obeys the ice rule, i.e. it cannot take nonzero values on inputs with Hamming weight not 2. Including the dangling edges, every vertex has exactly two incoming edges and two outgoing edges. Thus if we sum the in-degrees over all vertices, it must equal to the sum of out-degrees over all vertices. Every internal edge contributes exactly 1 to each sum. Thus the number of incoming dangling edges is equal to the number of outgoing dangling edges, which must be 2 each since they sum to 4. Second, f still satisfies arrow reversal symmetry. For any valid orientation of edges in the 4-ary construction contributing a term to $f(x)$, reversing the orientations on all edges has the same contribution to $f(\bar{x})$, because the constraint function on each vertex of degree 4 satisfies the arrow reversal symmetry.

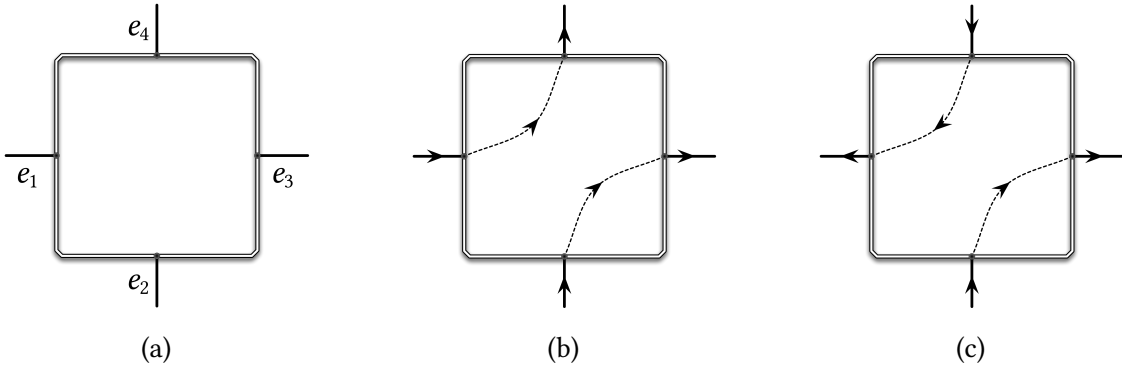


Figure 2.3: A 4-ary construction using vertices with constraint function in $\mathcal{X}$.

The notion of Eulerian partitions previously used for graphs (without external edges) can also be defined for 4-ary constructions. An *Eulerian partition* for a 4-ary construction g is a partition of the edges in g into edge-disjoint circuits and exactly two *trails* (walks with no repeated edges) whose ends are exactly the four dangling edges. The weight $\tilde{w}$ of such an Eulerian partition φ can be similarly defined. Set w such that
$$\begin{cases} w(\nearrow) = \frac{-a+b+c}{2} \\ w(\searrow) = \frac{a-b+c}{2} \\ w(\dashv) = \frac{a+b-c}{2} \end{cases}, \text{ or equivalently } \begin{cases} a = w(\searrow) + w(\dashv) \\ b = w(\nearrow) + w(\dashv) \\ c = w(\nearrow) + w(\searrow) \end{cases}.$$
 Observe that if a vertex has a constraint function $f \in \mathcal{X}$, then the weight of every pairing is *nonnegative*, and the weight of any directed Eulerian partition of a graph/4-ary construction comprised of such vertices is also *nonnegative*.

Consider $f(0011)$. Under the six-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 going in and e_3, e_4 going out. The contribution by τ is a weighted sum over a set $\Phi_{0011}(\tau)$ of directed Eulerian partitions. For every directed Eulerian partition $\varphi \in \Phi_{0011}(\tau)$, the two directed trails are either $\{e_1 \rightsquigarrow e_4, e_2 \rightsquigarrow e_3\}$ (Figure 2.3b) or $\{e_1 \rightsquigarrow e_3, e_2 \rightsquigarrow e_4\}$. Denote by $\Phi_{0011, \nearrow}$ the set of directed Eulerian partitions (distributed in potentially many different six-vertex configurations), each of which has directed trails $\{e_1 \rightsquigarrow e_4, e_2 \rightsquigarrow e_3\}$; denote by $\Phi_{0011, \dashv}$ the set of directed Eulerian partitions, each of which has directed trails $\{e_1 \rightsquigarrow e_3, e_2 \rightsquigarrow e_4\}$. In terms of directed Eulerian partitions of the 4-ary construction, $f(0011)$ can be seen as the weighted sum of elements from two disjoint sets $\Phi_{0011, \nearrow}$ and $\Phi_{0011, \dashv}$. Defining the weight of a set Φ of directed Eulerian partitions by $W(\Phi) = \sum_{\varphi \in \Phi} \tilde{w}(\varphi)$ yields $f(0011) = W(\Phi_{0011, \nearrow}) + W(\Phi_{0011, \dashv})$, and similarly $f(1100) = W(\Phi_{1100, \nearrow}) + W(\Phi_{1100, \dashv})$. Note that there is a bijective weight-preserving map between $\Phi_{0011, \nearrow}$ and $\Phi_{1100, \nearrow}$ by reversing the direction of every circuit and trail of an Eulerian partition. Thus, $W(\Phi_{0011, \nearrow}) = W(\Phi_{1100, \nearrow})$ and similarly $W(\Phi_{0011, \dashv}) = W(\Phi_{1100, \dashv})$. Consequently, $f(0011) = f(1100)$, $f(0110) = f(1001)$, and $f(0101) = f(1010)$.

An important observation is that for each Eulerian partition in $\Phi_{0011, \nearrow}$, if we only reverse the trail from $e_1 \rightsquigarrow e_4$ to $e_4 \rightsquigarrow e_1$ and keep the directions on all circuits and the other trail unchanged, this Eulerian partition has the same weight but now lies in $\Phi_{1010, \nearrow}$ (Figure 2.3c). This is because at every vertex v , reversing any orientation of a branch of the given pairing $\varrho_v \in \{\nearrow, \searrow, \dashv\}$ does

not change the value $w(\varrho_v)$. In this way, we set up a one-to-one weight-preserving map between $\Phi_{0011, \nearrow}$ and $\Phi_{1010, \nearrow}$, i.e. $W(\Phi_{0011, \nearrow}) = W(\Phi_{1010, \nearrow})$. Combining the result in the last paragraph, we can write

- $W(\nearrow) = W(\Phi_{0011, \nearrow}) = W(\Phi_{1100, \nearrow}) = W(\Phi_{0101, \nearrow}) = W(\Phi_{1010, \nearrow});$
- $W(\nwarrow) = W(\Phi_{0110, \nwarrow}) = W(\Phi_{1001, \nwarrow}) = W(\Phi_{0101, \nwarrow}) = W(\Phi_{1010, \nwarrow});$
- $W(-|-) = W(\Phi_{0011, -|-}) = W(\Phi_{1100, -|-}) = W(\Phi_{0110, -|-}) = W(\Phi_{1001, -|-}).$

Consequently, we have $\begin{cases} a' = W(\nearrow) + W(-|-) \\ b' = W(\nwarrow) + W(-|-) \\ c' = W(\nwarrow) + W(\nearrow) \end{cases}$. $W(\nearrow)$, $W(\nwarrow)$, and $W(-|-)$ are all nonnegative due to the fact that the weight of every directed Eulerian partition has a nonnegative weight. Therefore, $a' \leq b' + c'$, $b' \leq a' + c'$, and $c' \leq a' + b'$. This is to say, $f \in \mathcal{X}$. $\square$

Proof of Theorem 2.1.2. Inheriting the notations from the above proof, we have $w(-|-) = 0$ when $c = a + b$ for each vertex, which is to say no “crossing” can be made at any vertex in any Eulerian partition. Due to planarity, a trail $e_1 \rightsquigarrow e_3$ must cross a trail $e_2 \rightsquigarrow e_4$ at a vertex, thus $W(-|-) = W(\Phi_{0011, -|-}) = 0$. Therefore, $c' = a' + b'$. $\square$

Remark 2.1.1. Theorem 2.1.1 is important not only because it reveals a structural difference between parameter settings of the six-vertex model lying on the two sides of the phase transition threshold, but also for its crucial role in proving Corollary 2.2.3 which helps us upper bound the mixing time of the directed-loop algorithm and lower bound the proportion of valid six-vertex model configurations in the state space so that approximately counting via sampling [JVV86] leads to an FPRAS.

2.2 An FPRAS via Markov chains Monte Carlo

In this section we prove the following theorem.

Theorem 2.2.1. *There is an FPRAS for $Z_{6V}(a, b, c, d)$ if $(a, b, c) \in \mathcal{Z}$.*

For simplicity we prove Theorem 2.2.1 only for the case where all constraint functions of arity 4 are from a fixed finite subset $S \subset \mathcal{Z}$, i.e., we show that there is an FPRAS for computing $Z_{6V}(S)$. With some care the more general statement in Theorem 2.2.1 can also be proved.

We use the common approach to approximate counting via almost uniform sampling [JVV86] using a rapidly mixing Markov chain [JS89, DFK91, Sin92, Jer03]. The Markov chain is the widely-used directed-loop algorithm [RS72, YN79, BN98, SZ04] whose transitions are composed of creating, shifting, and merging of two “defects” on the edges. Some examples of the states in the directed-loop algorithm are shown in Figure 2.4 where the state in Figure 2.4a is an Eulerian orientation and the state in Figure 2.4b and the state in Figure 2.4c are “near-Eulerian” orientations with exactly two “defects”. Some typical moves in the directed-loop algorithm are as follows: the transition from the state in Figure 2.4a to the state in Figure 2.4b creates two defects; the transition from the state in Figure 2.4b to the state in Figure 2.4a merges two defects; the transitions between Figure 2.4b and Figure 2.4c shift one of the defects.

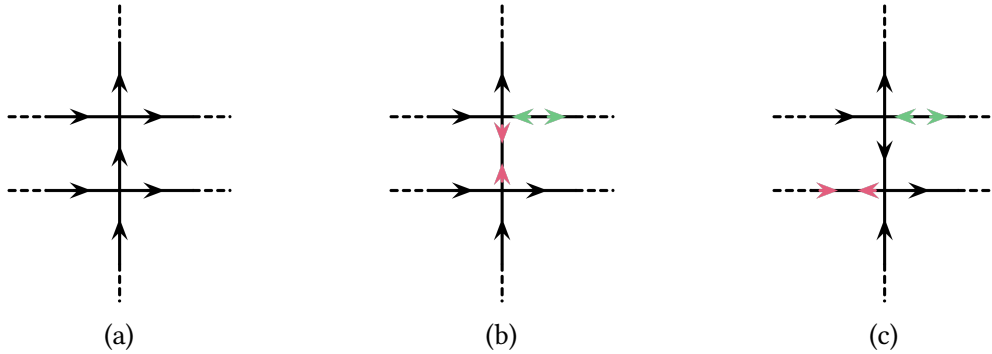


Figure 2.4: Examples of the states in the directed-loop algorithm.

Although the directed-loop algorithm $\mathcal{MC}$ runs on the Eulerian orientations of a 4-regular graph G , it is formally defined and analyzed using the edge-vertex incidence graph G' of G , introduced in Section 1.3. Let $G' = (V, U, E)$ be the edge-vertex incidence graph of G , an instance of $Z_{6V}(\mathcal{Z})$. Each vertex in V is assigned $(\neq_2)$; each vertex $u \in U$ is assigned a constraint function $f_u \in \mathcal{Z}$. An *assignment* σ assigns a value in $\{0, 1\}$ to each edge $e \in E$. The state space of $\mathcal{MC}$ is $\Omega = \Omega_0 \cup \Omega_2$, which consists of “perfect” or “near-perfect” assignments to E , defined as follows: All assignments satisfy the “two-0 two-1” ice rule at every vertex $u \in U$ of degree 4. We also

insist that all assignments satisfy the “one-0 one-1” at every $v \in V$ with *possibly exactly two* exceptions. Assignments in Ω_0 have no exceptions, and are “perfect” (corresponding to the Eulerian orientations in G). Assignments in Ω_2 have exactly two exceptions, and are “near-perfect” (corresponding to the near-Eulerian orientations in G). Thus any $\sigma \in \Omega_0$ satisfies all $(\neq_2)$ on V , and any $\sigma \in \Omega_2$ satisfies all $(\neq_2)$ on $V - \{v', v''\}$ for some two vertices $v', v'' \in V$ where it satisfies $(=_2)$ (which outputs 1 on inputs 00, 11 and outputs 0 on 01, 10).

For any assignment $\sigma \in \Omega$ and any subset $S \subseteq \Omega$, define the *weight* function $\mathcal{W}$ by $\mathcal{W}(\sigma) = \prod_{u \in U} f_u(\sigma|_{E(u)})$ and $Z(S) = \sum_{\sigma \in S} \mathcal{W}(\sigma)$. Then the *Gibbs measure* for Ω is defined by $\pi(\sigma) = \frac{\mathcal{W}(\sigma)}{Z(\Omega)}$, assuming $Z(\Omega) > 0$. Observe that if a state $\sigma \in \Omega_2$ assigns 00 to both edges incident to $v' \in V$ (satisfying $(=_2)$ at v') then it must assign 11 to both edges incident to $v'' \in V$, and vice versa. Indeed, having 00 at v' models the fact that v' has two arrows going out (to degree-4 vertices in U). To maintain the property that the number of incoming arrows is equal to the number of outgoing arrows everywhere else, v'' must have two arrows coming in, which is equivalent to having 11 at v'' in the edge 2-coloring model of the edge-vertex incidence graph. An example state is shown in Figure 2.6a.

Transitions in $\mathcal{MC}$ are comprised of three types of moves. Suppose $\sigma \in \Omega_0$. An Ω_0 -to- Ω_2 move from σ takes a degree 4 vertex $u \in U$ and two incident edges $e' = (v', u), e'' = (v'', u) \in V \times U$ satisfying $\{\sigma(e'), \sigma(e'')\} = \{0, 1\}$, and changes it to $\sigma_2 \in \Omega_2$ which flips both $\sigma(e')$ and $\sigma(e'')$. The effect is that we still have $\{\sigma_2(e'), \sigma_2(e'')\} = \{0, 1\}$, but at v' and v'' , σ_2 satisfies $(=_2)$ instead. An Ω_2 -to- Ω_0 move is the opposite. An Ω_2 -to- Ω_2 move is, intuitively, to *shift* one $(=_2)$ from one vertex $v' \in V$ to another $v^* \in V$, where for some $u \in U$, v' and v^* are both incident to u and the “two-0 two-1” rule at u is preserved. Formally, let $\sigma \in \Omega_2$ be a near-perfect assignment with $v', v'' \in V$ being the two exceptional vertices (i.e., σ satisfies $(=_2)$ at v' and v''). Let $v^* \in V - \{v', v''\}$ be such that for some $u \in U$, both $e' = (v', u), e^* = (v^*, u) \in E$, and $\{\sigma(e'), \sigma(e^*)\} = \{0, 1\}$. Then an Ω_2 -to- Ω_2 move changes σ to σ^* by flipping both $\sigma(e')$ and $\sigma(e^*)$. The effect is that we still have $\{\sigma^*(e'), \sigma^*(e^*)\} = \{0, 1\}$, but σ^* satisfies $(\neq_2)$ at v' and $(=_2)$ at v^* . Note that σ^* continues to satisfy $(=_2)$ at v'' .

The above describes a symmetric binary relation *neighbor* ($\sim$) on Ω . No two states in Ω_0 are neighbors. Set $n = |U|$. The number of neighbors of a Ω_0 -state is at most $4n$ (by first picking a vertex and then picking an “in-out” pair of edges incident to this vertex) and the number of neighbors of a Ω_2 -state is at most a constant. The transition probabilities $P(\cdot, \cdot)$ of $\mathcal{MC}$ are *Metropolis* moves between neighboring states:

$$P(\sigma_1, \sigma_2) = \begin{cases} \frac{1}{8n} \min\left(1, \frac{\pi(\sigma_2)}{\pi(\sigma_1)}\right) & \text{if } \sigma_2 \sim \sigma_1; \\ 1 - \frac{1}{8n} \sum_{\sigma' \sim \sigma_1} \min\left(1, \frac{\pi(\sigma')}{\pi(\sigma_1)}\right) & \text{if } \sigma_1 = \sigma_2; \\ 0 & \text{otherwise.} \end{cases}$$

$\mathcal{MC}$ is aperiodic due to the “lazy” movement; one can verify that $\mathcal{MC}$ is irreducible by creating, shifting, and merging of a pair of ($=_2$)’s; as the transitions are Metropolis moves, detailed balance conditions are satisfied with regard to π . By results from [JS89, Sin92], such a Markov chain is rapidly mixing if there is a *flow* whose congestion can be bounded by a polynomial in n .

Lemma 2.2.2. *Assume $Z(\Omega_0) > 0$. There is a flow on Ω with congestion at most $O\left(n^2 \left(\frac{Z(\Omega)}{Z(\Omega_0)}\right)^2\right)$, using paths of length $O(n)$.*

Proof. The idea is to design a flow $\mathfrak{F} : \mathcal{P} \rightarrow \mathbb{R}^+$ from Ω_2 to Ω_0 which satisfies

$$\sum_{p \in \mathcal{P}_{\sigma_2 \sigma_0}} \mathfrak{F}(p) = \pi(\sigma_2)\pi(\sigma_0), \quad \text{for all } \sigma_2 \in \Omega_2, \sigma_0 \in \Omega_0,$$

where $\mathcal{P}_{\sigma_2 \sigma_0}$ is defined to be a set of simple directed paths from σ_2 to σ_0 in $\mathcal{MC}$ and $\mathcal{P} = \bigcup_{\sigma_2 \in \Omega_2, \sigma_0 \in \Omega_0} \mathcal{P}_{\sigma_2 \sigma_0}$. Once the congestion of $\mathfrak{F}$ from Ω_2 to Ω_0 is polynomially bounded, so is the flow from Ω_0 to Ω_2 by symmetric construction. Moreover, there is a flow from Ω_2 to Ω_2 (or from Ω_0 to Ω_0) whose congestion can also be polynomially bounded by randomly picking an intermediate state in Ω_0 (or Ω_2 , respectively). Thus we have a flow on Ω with polynomially bounded congestion. This technique has been used in [JSV04, McQ13]. In the following we show that the congestion of $\mathfrak{F}$ from Ω_2 to Ω_0 is bounded by $O(n^2 \frac{Z(\Omega_2)}{Z(\Omega_0)})$. Then the bound in the lemma for a flow on Ω follows.

To describe the flow $\mathfrak{F}$, we first specify the sets of paths that are going to take the flow. In line with the definition of Ω_0 and Ω_2 , we define Ω_4 to be the set of assignments where there are exactly four violations of $(\neq_2)$ in V . Let $\Omega' = \Omega_0 \cup \Omega_2 \cup \Omega_4$. For $\sigma, \sigma' \in \Omega'$, let $\sigma \oplus \sigma'$ denote the *symmetric difference* (or bitwise XOR), where we view σ and σ' as two bit strings in $\{0, 1\}^{|E|}$. This is a 0-1 assignment to the edge set of the edge-vertex incidence graph $G' = (V, U, E)$ of G . We also treat $\sigma \oplus \sigma'$ as an edge subset of E (corresponding to bit positions having bit 1, where σ and σ' assign opposite values), and this defines an induced subgraph of G' . Since at every $u \in U$ of degree 4, the “two-0 two-1” rule is satisfied by both σ and σ' , this induced subgraph has even degree (0, 2, or 4) at every $u \in U$.

Denote by $U_4 \subseteq U$ the degree-4 vertices in $\sigma \oplus \sigma'$. Then there are exactly $2^{|U_4|}$ Eulerian partitions for $\sigma \oplus \sigma'$. Recall that an Eulerian partition of $\sigma \oplus \sigma'$ is uniquely determined by a family of pairings on U_4 . This is a 1-1 correspondence and we will identify the two sets. For any pairing in $\{\nearrow, \searrow, \dashv\}$ on a vertex u with constraint matrix $M(f_u) = \begin{bmatrix} & b & c \\ a & c & b \end{bmatrix}$, define the weight function $\mathfrak{w}$ for pairings as follows, $\begin{cases} \mathfrak{w}(\nearrow) = \frac{-a^2+b^2+c^2}{2} \\ \mathfrak{w}(\searrow) = \frac{a^2-b^2+c^2}{2} \\ \mathfrak{w}(\dashv) = \frac{a^2+b^2-c^2}{2} \end{cases}$, or equivalently $\begin{cases} a^2 = \mathfrak{w}(\searrow) + \mathfrak{w}(\dashv) \\ b^2 = \mathfrak{w}(\nearrow) + \mathfrak{w}(\dashv) \\ c^2 = \mathfrak{w}(\nearrow) + \mathfrak{w}(\searrow) \end{cases}$. Note that when $f_u \in \mathcal{Z}$, $\mathfrak{w}$ takes *nonnegative* values. Let $\Phi_{\sigma \oplus \sigma'}$ be the set of Eulerian partitions for $\sigma \oplus \sigma'$. For $\varphi \in \Phi_{\sigma \oplus \sigma'}$, define

$$\mathfrak{W}(\sigma, \sigma', \varphi) := \left(\prod_{u \in U \setminus U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \right) \left(\prod_{u \in U_4} \mathfrak{w}(\varphi(u)) \right).$$

Then for all distinct $\sigma, \sigma' \in \Omega'$, we have

$$\begin{aligned} \sum_{\varphi \in \Phi_{\sigma \oplus \sigma'}} \mathfrak{W}(\sigma, \sigma', \varphi) &= \sum_{\varphi \in \Phi_{\sigma \oplus \sigma'}} \left(\prod_{u \in U \setminus U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \right) \left(\prod_{u \in U_4} \mathfrak{w}(\varphi(u)) \right) \\ &= \left(\prod_{u \in U \setminus U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \right) \left(\sum_{\varphi \in \Phi_{\sigma \oplus \sigma'}} \prod_{u \in U_4} \mathfrak{w}(\varphi(u)) \right) \\ &= \left(\prod_{u \in U \setminus U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \right) \left(\prod_{u \in U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \right) \\ &= \prod_{u \in U} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \\ &= \mathcal{W}(\sigma) \mathcal{W}(\sigma'). \end{aligned}$$

The equality from line 2 to line 3 is due to the following: when the degree (in the induced subgraph $\sigma \oplus \sigma'$) of a vertex $u \in U$ is 4, σ and σ' must take the same value at u , since one represents a total reversal of all arrows of another; thus $f_u(\sigma|_{E(u)})f_u(\sigma'|_{E(u)})$ is in $\{a^2, b^2, c^2\}$. Then

$$\prod_{u \in U_4} f_u(\sigma|_{E(u)})f_u(\sigma'|_{E(u)}) = \sum_{\varphi \in \Phi_{\sigma \oplus \sigma'}} \prod_{u \in U_4} w(\varphi(u))$$

is obtained by using the sum expressions for a^2 , b^2 and c^2 in terms of $w(\curvearrowright)$, $w(\curvearrowleft)$, and $w(\dashv)$, and then expressing the product-of-sums as a sum-of-products.

Now we are ready to specify the “paths” which take nonzero flow from $\sigma_2 \in \Omega_2$ to $\sigma_0 \in \Omega_0$. In order to transit from σ_2 to σ_0 , paths in $\mathcal{P}_{\sigma_2 \sigma_0}$ go through states in Ω that gradually decrease the number of conflicting assignments along trails and circuits in $\sigma_2 \oplus \sigma_0$. We first specify a total order on E , the set of edges of G' . This induces a total order on circuits by lexicographic order. In the induced subgraph $\sigma_2 \oplus \sigma_0$, exactly two vertices in V have degree 1 (called *endpoints*) and all other vertices have degree 2 or degree 4. The set of paths in $\mathcal{P}_{\sigma_2 \sigma_0}$ are designed to be in 1-to-1 correspondence with elements in $\Phi_{\sigma_2 \oplus \sigma_0}$. Given any family of pairings $\varphi \in \Phi_{\sigma_2 \oplus \sigma_0}$, we have a unique decomposition of the induced subgraph $\sigma_2 \oplus \sigma_0$ as an edge disjoint union of one trail $[e_1](v_1, e'_1, u_1, e_2, v_2, e'_2, u_2, \dots, e_k, v_k)[e'_k]$ (where e_1 and e'_k are not part of the trail), and zero or more edge disjoint circuits, which are ordered lexicographically. Here $v_i \in V$ and $u_i \in U$, and we may assume $\sigma_2(e_1) = \sigma_2(e'_1) = 0$, $\sigma_2(e_2) = 1$, $\sigma_2(e'_2) = 0$, $\dots$, $\sigma_2(e_k) = \sigma_2(e'_k) = 1$. So the two exceptional vertices are v_1 and v_k , where σ_2 satisfies $(=)_2$. The unique path p_φ first “pushes” the $(=)_2$ from v_1 , to v_2 , then to $v_3, \dots, v_{k-1}$, and then “merge” at v_k , arriving at a configuration in Ω_0 . Then p_φ reverses all arrows on each circuit in lexicographic order, and within each circuit C it starts at the least edge e (according to the edge order) and reverses all arrows on C in the direction defined by the starting cyclic orientation of σ_2 . (Technically it flips a pair of incident edges to vertices in U in each step.) Such paths p_φ are well-defined and are valid paths in $\mathcal{MC}$ since along any path every state is in $\Omega = \Omega_0 \cup \Omega_2$ and every move is a valid transition defined in $\mathcal{MC}$. With regard to the flow distribution, the flow value put on p_φ is $\frac{\mathfrak{W}(\sigma_2, \sigma_0, \varphi)}{(Z(\Omega))^2}$, making the following hold for all

$\sigma_2 \in \Omega_2, \sigma_0 \in \Omega_0$:

$$\begin{aligned} \sum_{p_\varphi \in \mathcal{P}_{\sigma_2 \sigma_0}} \mathfrak{F}(p_\varphi) &= \sum_{\varphi \in \Phi_{\sigma_2 \oplus \sigma_0}} \frac{\mathfrak{W}(\sigma_2, \sigma_0, \varphi)}{(Z(\Omega))^2} \\ &= \frac{\mathcal{W}(\sigma_2) \mathcal{W}(\sigma_0)}{(Z(\Omega))^2} \\ &= \pi(\sigma_2) \pi(\sigma_0). \end{aligned}$$

Note that in each path, no edge is flipped more than once, so the length is $O(n)$. For any transition (σ', σ'') where $\sigma' \neq \sigma''$, we have $P(\sigma', \sigma'') = \frac{1}{8n} \min \left(1, \frac{\pi(\sigma'')}{\pi(\sigma')} \right) = \Omega \left(\frac{1}{n} \right)$, as $\frac{\pi(\sigma'')}{\pi(\sigma')}$ is a constant. (This is a constant because we have restricted the constraint function f_u to be from a fixed finite set S .) Let $H_{\sigma'} = \{\sigma_2 \oplus \sigma_0 \mid \sigma_2 \in \Omega_2, \sigma_0 \in \Omega_0, \exists \varphi \in \Phi_{\sigma_2 \oplus \sigma_0} \text{ s.t. } \sigma' \in p_\varphi\}$. The congestion of $\mathfrak{F}$ is

$$\begin{aligned} &\max_{\text{transition } (\sigma', \sigma'')} \frac{1}{\pi(\sigma') P(\sigma', \sigma'')} \sum_{\substack{\sigma_2 \in \Omega_2 \\ \sigma_0 \in \Omega_0}} \sum_{\substack{p_\varphi \in \mathcal{P}_{\sigma_2 \sigma_0} \\ p_\varphi \ni (\sigma', \sigma'')}} \frac{\mathfrak{W}(\sigma_2, \sigma_0, \varphi)}{(Z(\Omega))^2} \\ &\leq \max_{\sigma' \in \Omega} \frac{O(n)}{\mathcal{W}(\sigma') Z(\Omega)} \sum_{\substack{\sigma_2 \in \Omega_2 \\ \sigma_0 \in \Omega_0}} \sum_{\substack{\varphi \in \Phi_{\sigma_2 \oplus \sigma_0} \\ p_\varphi \ni \sigma'}} \mathfrak{W}(\sigma_2, \sigma_0, \varphi) \\ &= \max_{\sigma' \in \Omega} \frac{O(n)}{\mathcal{W}(\sigma') Z(\Omega)} \sum_{\sigma_2 \in \Omega_2} \sum_{\eta \in H_{\sigma'}} \sum_{\varphi \in \Phi_\eta} \mathfrak{W}(\sigma_2, \sigma_2 \oplus \eta, \varphi) \\ &= \max_{\sigma' \in \Omega} \frac{O(n)}{\mathcal{W}(\sigma') Z(\Omega)} \sum_{\eta \in H_{\sigma'}} \sum_{\varphi \in \Phi_\eta} \sum_{\sigma_2 \in \tilde{\Omega}_2} \mathfrak{W}(\sigma_2, \sigma_2 \oplus \eta, \varphi). \end{aligned}$$

On the last line above we exchange the order of summations where $\tilde{\Omega}_2$ is the set of Ω_2 -states that are compatible with the symmetric difference η and its Eulerian partition φ . The number of states in $\tilde{\Omega}_2$ is bounded by the length of the longest path $O(n)$ because σ' is an intermediate state on a path. Fix any $\sigma' \in \Omega$. For any $\sigma_2 \in \Omega_2$, and $\eta \in H_{\sigma'}$ consisting of exactly one connected component with two endpoints of degree 1 and all other vertices having even degree (and zero or more connected components of even degree vertices), observe that $\sigma' \oplus \eta \in \Omega'$. Indeed, if $\sigma' \in \Omega_0$ then $\sigma' \oplus \eta \in \Omega_2$; if $\sigma' \in \Omega_2$ then depending on whether σ'

(1) is σ_2 , or

(2) appears in the process of reversing arrows on the trail with two endpoints, or

(3) appears after reversing arrows on the trail with endpoints,

$\sigma' \oplus \eta$ lies in Ω_0 , Ω_2 , or Ω_4 , respectively. For the edges not in η , σ' agrees with σ_2 and $\sigma_2 \oplus \eta$ as the path p_φ never “touches” them, and so does $\sigma' \oplus \eta$. Recall that

$$\mathfrak{W}(\sigma_2, \sigma_2 \oplus \eta, \varphi) = \left(\prod_{u \in U \setminus U_4} f_u(\sigma_2|_{E(u)}) f_u((\sigma_2 \oplus \eta)|_{E(u)}) \right) \left(\prod_{u \in U_4} \mathfrak{w}(\varphi(u)) \right).$$

For every vertex $u \in U$ that is not in η , f_u takes the same value in all σ_2 , $\sigma_2 \oplus \eta$, σ' , and $\sigma' \oplus \eta$. For every vertex $u \in U$ that is degree-2 in η , assuming $M(f_u) = \begin{bmatrix} & b & c & a \\ a & & & \end{bmatrix}$, $f_u(\sigma_2|_{E(u)})$ and $f_u((\sigma_2 \oplus \eta)|_{E(u)})$ take two different elements in $\{a, b, c\}$. Meanwhile, $f_u(\sigma'|_{E(u)})$ and $f_u(\sigma' \oplus \eta|_{E(u)})$ also take these two elements (possibly in the opposite order). For example, at the vertex u shown in Figure 2.5, $f_u(\sigma_2|_{E(u)}) = a$ and $f_u((\sigma_2 \oplus \eta)|_{E(u)}) = c$. The two solid edges are in η and assignments on the two dotted edges are shared by σ_2 and $\sigma_2 \oplus \eta$, as well as σ' and $\sigma' \oplus \eta$. On the path p_φ from σ_2 to $\sigma_2 \oplus \eta$ decided by φ : if σ' appears before reversing the two solid edges, then σ' agrees with σ_2 on them ($f_u(\sigma'|_{E(u)}) = a$) and $\sigma' \oplus \eta$ agrees with $\sigma_2 \oplus \eta$ on them ($f_u(\sigma' \oplus \eta|_{E(u)}) = c$); if σ' appears after reversing the two solid edges, then σ' agrees with $\sigma_2 \oplus \eta$ on them ($f_u(\sigma'|_{E(u)}) = c$) and $\sigma' \oplus \eta$ agrees with σ_2 on them ($f_u(\sigma' \oplus \eta|_{E(u)}) = a$). For every vertex $u \in U$ that is degree-4 in η , $\mathfrak{w}(\varphi(u))$ takes the same value in $\mathfrak{W}(\sigma_2, \sigma_2 \oplus \eta, \varphi)$ and $\mathfrak{W}(\sigma', \sigma' \oplus \eta, \varphi)$ as the weight only depends on $\varphi(u)$, the pairing at u .



Figure 2.5: A vertex of degree-2 in η .

By the above argument, we established that $\mathfrak{W}(\sigma_2, \sigma_2 \oplus \eta, \varphi) = \mathfrak{W}(\sigma', \sigma' \oplus \eta, \varphi)$. Therefore, the

congestion of $\mathfrak{F}$ can be bounded by

$$\begin{aligned}
& \max_{\sigma' \in \Omega} \frac{O(n)}{\mathcal{W}(\sigma')Z(\Omega)} \sum_{\eta \in H_{\sigma'}} \sum_{\varphi \in \Phi_\eta} \sum_{\sigma_2 \in \Omega_2} \mathfrak{W}(\sigma', \sigma' \oplus \eta, \varphi) \\
& \leq \max_{\sigma' \in \Omega} \frac{O(n^2)}{\mathcal{W}(\sigma')Z(\Omega)} \sum_{\eta \in H_{\sigma'}} \sum_{\varphi \in \Phi_\eta} \mathfrak{W}(\sigma', \sigma' \oplus \eta, \varphi) \\
& \leq \max_{\sigma' \in \Omega} \frac{O(n^2)}{\mathcal{W}(\sigma')Z(\Omega)} \sum_{\eta \in H_{\sigma'}} \mathcal{W}(\sigma')\mathcal{W}(\sigma' \oplus \eta) \\
& = \max_{\sigma' \in \Omega} \frac{O(n^2)}{Z(\Omega)} \sum_{\eta \in H_{\sigma'}} \mathcal{W}(\sigma' \oplus \eta) \\
& \leq \frac{O(n^2)}{Z(\Omega)} \sum_{\sigma \in \Omega'} \mathcal{W}(\sigma) \\
& = O(n^2) \frac{Z(\Omega')}{Z(\Omega)}.
\end{aligned}$$

By a standard argument as in [JS89, MW96, McQ13], $\frac{Z(\Omega_4)}{Z(\Omega_2)} \leq \frac{Z(\Omega_2)}{Z(\Omega_0)}$. Therefore, the congestion is bounded by $O(n^2) \frac{Z(\Omega_2)}{Z(\Omega_0)}$. $\square$

Remark 2.2.1. We have an alternative derivation of Lemma 2.2.2 using the notion of “windability” [McQ13]. However, this alternative derivation does not yield a proof of Theorem 2.2.1; we still require the results from Section 2.1 to show that.

In order to show $\mathcal{MC}$ is rapidly mixing, we need to show $\frac{Z(\Omega_2)}{Z(\Omega_0)}$ is polynomially bounded. This bound is also needed to get an FPRAS from a rapidly mixing Markov chain in Ω , since ultimately we are only interested in Ω_0 . Such a bound is a corollary of Theorem 2.1.1.

Corollary 2.2.3. $\frac{Z(\Omega_2)}{Z(\Omega_0)} = O(n^2)$.

Proof. For each $\sigma \in \Omega_2$, there are exactly two vertices in V satisfying $(=_2)$. Let $\Omega_2^{\{v', v''\}} \subseteq \Omega_2$ be the set of states in which v', v'' are these two vertices. We have $\frac{Z(\Omega_2)}{Z(\Omega_0)} = \sum_{\{v', v''\} \in \binom{V}{2}} \frac{Z(\Omega_2^{\{v', v''\}})}{Z(\Omega_0)}$. For any $\sigma \in \Omega_2^{\{v', v''\}}$, the local assignments around v' and v'' must be 00 on one and 11 on the other. An example is in Figure 2.6a. If we “delete” v' and v'' as shown in Figure 2.6b, we get a 4-ary construction g using degree 4 vertices with constraint functions in $\mathcal{Z} \subset \mathcal{Z}$. Denote the constraint matrix of g by $M(g) = \begin{bmatrix} & b' & c' & a' \\ & c' & b' & \\ a' & & & \end{bmatrix}$, with the input order being counter-clockwise starting from the

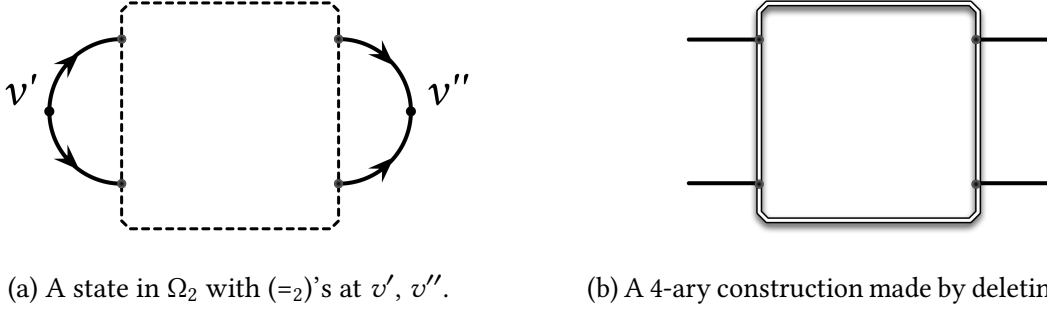


Figure 2.6: The correspondence between configurations before and after deleting two vertices with $(=_2)$.

upper-left edge. For this 4-ary construction g we observe that: the states in $\Omega_2^{\{v', v''\}}$ where edges incident to v' (also v'') take the same value contribute a total weight $(a' + a')$, i.e. $Z(\Omega_2^{\{v', v''\}}) = 2a'$; the states in Ω_0 where v', v'' satisfy $(\neq_2)$ have a total weight $Z(\Omega_0) = 2b' + 2c'$. Note that $\mathcal{Z} \subset \mathcal{Z}$. By Theorem 2.1.1 we know that for 4-ary construction g , $a' \leq b' + c'$. Therefore, $\frac{Z(\Omega_2^{\{v', v''\}})}{Z(\Omega_0)} \leq 1$. In total, $\frac{Z(\Omega_2)}{Z(\Omega_0)} \leq \binom{|V|}{2}$. $\square$

Combining Lemma 2.2.2 and Corollary 2.2.3, we conclude that $\mathcal{MC}$ is rapidly mixing, and Ω_0 , the set of valid six-vertex configurations, in total takes a non-negligible proportion in the stationary distribution. As a consequence, we are able to efficiently sample six-vertex configurations according to the Gibbs measure on Ω_0 , and in the following algorithm we only work with states in Ω_0 . We design the following algorithm to approximately compute $Z_{6V}(\mathcal{Z})$ via sampling with the directed-loop algorithm $\mathcal{MC}$. As we have argued in Section 2.1, the partition function of the six-vertex models can be viewed as the weighted sum of Eulerian partitions. For a vertex $v \in U$, the ratios among different pairings ($\curvearrowright$, $\curvearrowleft$, and $\dashv$) in weighted Eulerian partitions can be uniquely determined by the ratios among different orientations (represented by a , b , and c) at v . As long as the partition function is not zero (this can be easily tested in P), there must be a pairing ϱ showing up at v with probability at least $\frac{1}{3}$ among all three pairings. Therefore, running $\mathcal{MC}$ on G , we can approximate, with a sufficient $1/\text{poly}(n)$ precision, the probability of having ϱ 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 , each assigned a $(\neq_2)$ and the edges reconnected according to ϱ . Write the partition function of $G_{v, \varrho}$ as $Z_{6V}(G_{v, \varrho})$, we

have $\Pr_v(\varrho) = w(\varrho)Z_{6V}(G_{v,\varrho})/Z_{6V}(G)$ which means $Z_{6V}(G) = w(\varrho)Z_{6V}(G_{v,\varrho})/\Pr_v(\varrho)$. To approximate $Z_{6V}(G)$ it suffices to approximate $Z_{6V}(G_{v,\varrho})$, which can be done by running $\mathcal{MC}$ on $G_{v,\varrho}$ and recursing. Repeating this process for $|U|$ steps we decompose the graph G into the base case, a set of disjoint cycles with even number of vertices, each assigned a $(\neq_2)$. The partition function of this cycle graph is just 2^C where C is the number of cycles. By this self-reduction, the partition function for G can be approximated.

Therefore, Theorem 2.2.1 is proved. Note that for the special case $(1, 1, 1)$, the FPRAS by Mihail and Winkler is a reduction [MW96] to computing the number of perfect matchings in a bipartite graph. We give a direct algorithm using Markov chain Monte Carlo.

2.3 NP-hardness

Theorem 2.3.1. *If $(a, b, c) \in \overline{\mathcal{X}}$ and $a, b, c > 0$, then $Z_{6V}(a, b, c)$ does not have an FPRAS unless $\text{RP} = \text{NP}$.*

Proof. Let 3-MIS denote the NP-hard problem of computing the cardinality of a maximum independent set in a 3-regular graph [GJS76]. We reduce 3-MIS to approximating $Z_{6V}(a, b, c)$. Since $(a, b, c) \in \overline{\mathcal{X}}$, all $a, b, c > 0$. Since the proof of NP-hardness for $Z_{6V}(a, b, c)$ is for general graphs (i.e., not necessarily planar), we can permute the parameters so that $c > a + b$, and normalize $c > b \geq a = 1$. Let $\gamma = \frac{c}{a+b}$. Then $\gamma > 1$.

Before proving this theorem we briefly state our idea. Denote an instance of 3-MIS by $G = (V, E)$. For any independent set, no two adjacent vertices $u, v \in V$ can both appear. The only possible configurations for adjacent $\{u, v\}$ in any independent set S are $(u \in S, v \notin S)$, $(u \notin S, v \in S)$, and $(u \notin S, v \notin S)$. We want to encode this local constraint by a local fragment of a graph G' in terms of configurations in the six-vertex model.

In Figure 2.7a we show how to implement a toy example—a single edge $\{u, v\}$ —by a construction in the six-vertex model with parameters $(a = 1, b \geq 1, c > a + b)$. We create two vertices, the left one for u and the right one for v , both given the constraint function f , and connect them as is

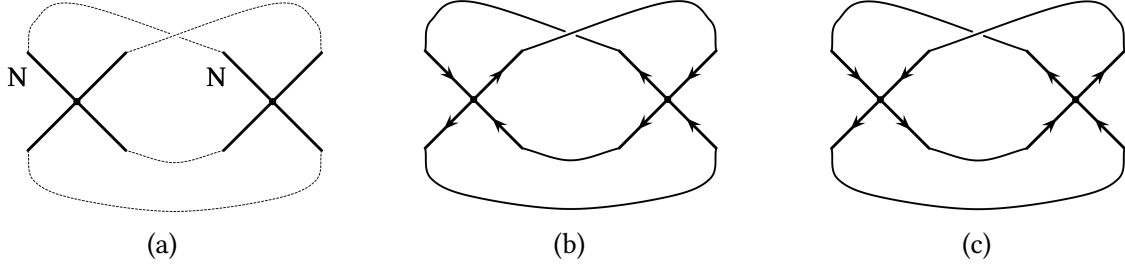


Figure 2.7: A construction implementing a single edge in an independent set.

shown in Figure 2.7a. The order of the 4 edges at each vertex is aligned to Figure 1.1 by a rotation so that the northwest edge marked by “N” corresponds to the north edge in Figure 1.1. There are a total of 4 edges in Figure 2.7a. Every 2-in 2-out configuration on the left vertex uniquely extends to a 2-in 2-out configuration on the right, and vice versa. Hence there are a total of 6 valid configurations. When the left vertex has a saddle configuration (in-out-in-out, or its reversal) which has weight c , the right must have a non-saddle configuration of weight b . Figure 2.7b depicts one such configuration; reversing all arrows gives another one having the same weight. Similarly if the right vertex has a saddle configuration (or its reversal) having weight c then the left must be a non-saddle having weight b . There are two more configurations with two non-saddles (Figure 2.7c and its reversal); these both have weight $a^2 = 1$. This models how two adjacent vertices interact in 3-MIS. We will call the connection pattern described in Figure 2.7a between two sets of 4 dangling edges the *four-way connection*.

However, when a vertex in G has more than one neighbor, simply duplicating this elementary implementation will not work, because we cannot make sure that the duplicate copies corresponding to the same vertex v behave consistently. To handle this difficulty, we design a locking device (Figure 2.8) for every $v \in V$ such that the property whether v belongs to an independent set in G is consistently reflected in G' in terms of being in a saddle configuration or not. This locking mechanism is enforced in the sense of approximation.

In Figure 2.8, we identify the leftmost node J with the rightmost node J —there are three “circles” in total. Most nodes depicted in Figure 2.8 will be replaced by some construction implementing the locking mechanism. Each circle has 4 dangling edges. The “left circle” has two dangling edges

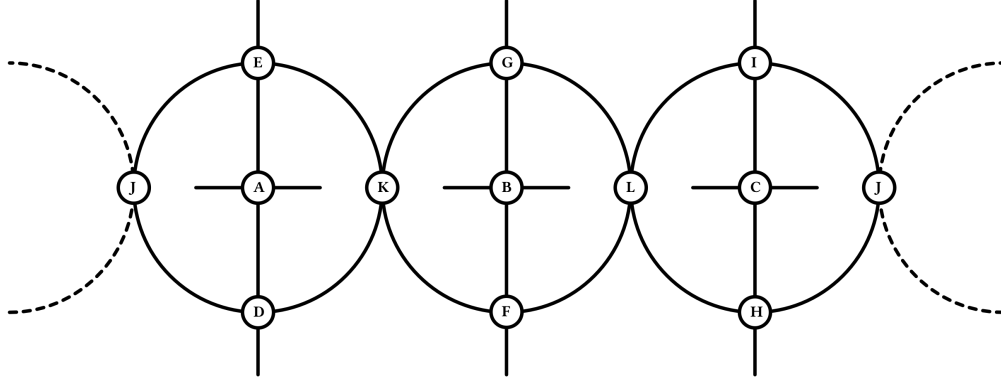


Figure 2.8: A locking device implementing a degree three vertex and its incident edges.

incident to A , one incident to D , and one incident to E . Similarly for the “middle circle” and the “right circle”. Each edge $\{u, v\}$ in G is modeled by a four-way connection of the 4 dangling edges between one circle of the construction for u and another circle of the construction for v . The 4 dangling edges of any circle is said to be in a saddle (or non-saddle) configuration if it is so when viewed externally in the cyclic order depicted in Figure 2.8.

The locking mechanism is to realize the following: when the four dangling edges of one of the 3 circles take a saddle configuration, (either in-out-in-out, or out-in-out-in), the other two circles must also take the *identical* saddle configuration (in-out-in-out, or out-in-out-in, respectively); when one circle takes any non-saddle configuration, the other two circles can take independently any non-saddle configurations, with no linkage (aside being a non-saddle). This is made possible by *chaining*, and the guarantee is enforced by approximate counting.

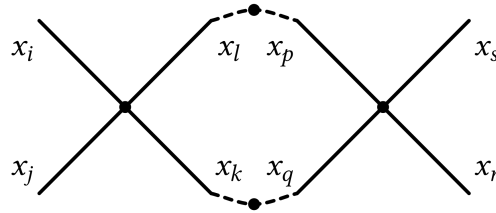


Figure 2.9: 2-chain.

Figure 2.9 depicts a *2-chain*. Given a 4-ary signature f with $M(f) = \begin{bmatrix} b & c & a \\ a & c & b \end{bmatrix}$, we place the constraint function ($\neq_2$) on the two degree 2 vertices connecting the two degree 4 vertices, each assigned a copy of f , and the edges are ordered so that $M = M_{x_i x_j, x_l x_k}(f) = M_{x_p x_q, x_s x_r}(f) = \begin{bmatrix} a & b & c \\ c & b & a \end{bmatrix}$.

Then the constraint function f_2 of this 2-chain construction is obtained by matrix multiplication $M(f_2) = M_{x_i x_j, x_s x_r}(f_2) = MNM$, where $N = \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix}$. Thus $M(f_2) = \begin{bmatrix} a_2 & b_2 & c_2 \\ & b_2 & a_2 \\ c_2 & & \end{bmatrix} = \begin{bmatrix} 2ab & a^2+b^2 & c^2 \\ a^2+b^2 & 2ab & \\ c^2 & & \end{bmatrix}$, where $c_2 = c^2$, $a_2 + b_2 = (a + b)^2$. Thus $c_2 > a_2 + b_2$, and $\gamma_2 := \frac{c_2}{a_2+b_2} = \left(\frac{c}{a+b}\right)^2 = \gamma^2$.

This can be generalized to a k -chain, which connects k vertices with constraint function f by $k - 1$ copies of N , such that $c_k = c^k$, $a_k + b_k = (a + b)^k$, $\gamma_k = \gamma^k$. Notice that when $c > a + b$, the ratio $\frac{c}{a+b}$ can be amplified exponentially in k in a k -chain. Therefore, by a chain of polynomially bounded size we can ensure the undesirable configurations are negligible — the 4-ary construction is locked into the only two complementary configurations which c represents. It can be verified that $b_k = (\mu^k + \nu^k)/2 \geq a_k = (\mu^k - \nu^k)/2 \geq 1$, where $\mu = a + b$ and $\nu = b - a$. We can “normalize” a k -chain by dividing a_k , so that its parameters are $\tilde{c}_k = c^k/a_k > \tilde{b}_k = b_k/a_k \geq \tilde{a}_k = 1$, and the ratio $\frac{\tilde{c}_k}{\tilde{a}_k + \tilde{b}_k} = \gamma^k$.

To reduce the problem 3-MIS to approximating $Z_{6V}(a, b, c)$, let $\kappa > \lambda \geq 1$ be two constants that will be fixed later. For each 3-MIS instance $G = (V, E)$ with $|V| = n$, we construct a graph G' where a device in Figure 2.8 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 between two circles corresponding to $\{u, v\}$ as in Figure 2.7a. For each device in Figure 2.8, each of the nodes A, B, C is replaced by a *normalized* λn -chain to boost the ratio of the saddle configuration over other configurations; each of the nodes D, F, H is replaced by a κn^2 -chain to lock in the configuration “all arrows pointing up and right” and its reversal; each of the nodes E, G, I is also replaced by a κn^2 -chain to lock in the configuration “all arrows pointing down and right” and its reversal (these configurations at D, F, H , and at E, G, I respectively, will be called locking configurations); at each of J, K, L , we just put f , of which the maximum weight of a configuration over the minimum is a constant $\frac{c}{\min\{a, b\}} = c$. Note that the constraint function in Figure 2.9 has the dominating entry at 0011 and 1100. Since our graph G' does not need to be planar, we can reorder the 4 external edges arbitrarily. In particular, for A, B, C the dominating entry $\tilde{c}_{\lambda n}$ is in the saddle 0101 and 1010 positions, as depicted in Figure 2.8. Similarly the 4 external edges of D, F, H and E, G, I are also properly reordered from the order given in Figure 2.9, as a κn^2 -chain to achieve the proper locking configurations.

Next we argue that the maximum size s of independent sets in G can be recovered from an approximate solution to $Z_{6V}(G'; a, b, c)$.

Given an independent set $S \subset V$ of size s , we show there is a valid configuration (at the granularity of nodes and edges shown in Figure 2.8) of weight $\geq c^{6\kappa n^3} (\tilde{c}_{\lambda n} \tilde{b}_{\lambda n})^{3s}$. For any vertex $v \in S$ we set the following configuration for its locking device: set each of 3 nodes A, B, C to the same saddle configuration in-out-in-out cyclically starting from the upper edge—each has weight $\tilde{c}_{\lambda n}$; set each of 3 nodes D, F, H to the same out-out-in-in locking configuration (clockwise) cyclically starting from the upper edge—each has weight $c^{\kappa n^2}$; set each of 3 nodes E, G, I to the same in-out-out-in locking configuration (clockwise) cyclically starting from the upper edge—each also has weight $c^{\kappa n^2}$; set each of 3 nodes J, K, L to the same configuration “two in from the left and two out to the right”, which has a non-zero weight ≥ 1 . For any vertex $v \notin S$ we set the following configuration for its locking device: All D, F, H, E, G, I will be in some locking configurations. Consider any of the 3 circles in the device, for example the circle formed by A, D, E, J, K . The node A is involved in a four-way connection to another circle belonging to a device for some vertex u . If $u \in S$, the assigned configuration just defined at u forces a non-saddle configuration here; more specifically the horizontal two dangling edges at A must either both point right (or both point left), and the upper edge $\hat{e}$ of E and the lower edge $\underline{e}$ of D must either both point down (or both point up, respectively). Regardless of which of the two assignments for the pair $(\hat{e}, \underline{e})$, either both down or both up, we can assign a locking configuration for E and D so that the upper and lower edges of A are either both point down or both point up respectively, inheriting the orientation at $(\hat{e}, \underline{e})$. Note that in either case, the left two edges of K are one-in-one-out; similarly the right two edges of J are also one-in-one-out (this allows “freedom” between the 3 circles where each of J, K, L can take a nonzero weight ≥ 1). Continuing our description at the circle A, D, E, J, K , if $u \notin S$, then we will pick an arbitrary non-saddle to non-saddle configuration in the four-way connection for $\{u, v\}$. These can all be extended to a valid configuration at A, D, E such that the configuration at A is non-saddle having weight ≥ 1 , the configurations at D and E are locking, and the right two edges of J and the left two edges of K are both one-in-one-out. The weight at D and E are still

$c^{\kappa n^2}$. Because J and K each has one-in-one-out from within the side of the circle, the 3 circles can be assigned independently from each other. This allows us to handle the situation where, for the same $v \notin S$, some edge $\{v, u\}$ connects to $u \in S$ and some edge $\{v, u'\}$ connects to $u' \notin S$.

We have defined a valid configuration, and it has weight $\geq \prod_{v \in V} c^{6\kappa n^2} \prod_{v \in S} (\tilde{c}_{\lambda n} \tilde{b}_{\lambda n})^3 = c^{6\kappa n^3} (\tilde{c}_{\lambda n} \tilde{b}_{\lambda n})^{3s}$, where 6 comes from the 6 locking nodes D, E, F, G, H, I in each locking device. (Omitted factors are all ≥ 1 .)

Next we show that the weighted sum of all configurations is smaller than $\frac{1}{2} c^{6\kappa n^3} (\tilde{c}_{\lambda n} \tilde{b}_{\lambda n})^{3(s+1)}$, where s is the maximum size of independent sets in G . First we bound W_{lock} , the sum of weights for configurations where all nodes labeled D, E, F, G, H, I are locked. Consider any circle such as the one labeled A, D, E, J, K for any $v \in V$. It is involved in a four-way connection with another such circle for a vertex $u \in V$, say A', D', E', J', K' , where $\{u, v\} \in E(G)$. The fact that D is locked forces that the upper and lower edges of D are to be consistently oriented, i.e., both up or both down. Similarly, consistency holds at E, D' and E' . Thus the four-way connection forces that there can be at most one of A and A' is in a saddle configuration. Furthermore, if A is in a particular saddle configuration, say in-out-in-out starting from the upper edge, both upper and lower edges of D must point up, and both upper and lower edges of E must point down, and then both right edges of D and E must point right, causing the left two edges of K point in, and thus the right two edges of K point out. This forces both F and G to take *exactly* the same locked configurations of D and E respectively, which forces B to be in *exactly* the same saddle configuration as A . Similarly so is C . We conclude that when all nodes labeled D, E, F, G, H, I are locked, for any $v \in V$, if any of its A, B, C is in a saddle configuration, then all 3 are in exactly the same saddle configuration, and none of A', B', C' for $u \in V$ is a saddle, if $\{u, v\} \in E$. In particular, there can be at most $3s$ many saddles among A, B, C 's in G' . If $0 \leq i \leq s$ is the number of $\{A, B, C\}$'s being in saddle, their weight is $(\tilde{c}_{\lambda n})^{3i}$, and their corresponding non-saddles in respective four-way connections must take weight $(\tilde{b}_{\lambda n})^{3i}$. Those $(3n - 6i)/2$ pairwise four-way connections (here $(3n - 6i)/2$ is an integer, as $n = |V|$ is even for a 3-regular graph G) between two non-saddles have weight $\tilde{a}_{\lambda n} = 1$. Note that, if any of those non-saddles were to take weight $\tilde{b}_{\lambda n}$, then the corresponding paired node in

its four-way connection must be in saddle, a contradiction.

It follows that

$$W_{\text{lock}} \leq 2^{6n} \left(c^{\kappa n^2} \right)^{6n} \sum_{i=0}^s \binom{n}{i} (\tilde{c}_{\lambda n} \tilde{b}_{\lambda n})^{3i} c^{3n} \leq 2^{7n} c^{6\kappa n^3 + 3n} (\tilde{c}_{\lambda n} \tilde{b}_{\lambda n})^{3s},$$

where each locked node D, E, F, G, H, I has 2 possible locking configurations each with weight $c^{\kappa n^2}$, and given a particular assignment of $6n$ locking configurations, there can be at most s batches of A, B, C 's in saddle configurations (same for each batch and determined by the locks) with weight $\tilde{c}_{\lambda n}$. Hence $W_{\text{lock}} < \frac{1}{4} c^{6\kappa n^3} (\tilde{c}_{\lambda n} \tilde{b}_{\lambda n})^{3(s+1)}$, when $\lambda \geq 1$ is large.

It remains to upper-bound the weighted sum of configurations where there is at least one device with some lock broken. This quantity is bounded by

$$\begin{aligned} & \sum_{i=0}^{6n-1} \binom{6n}{i} c^{i\kappa n^2} (a+b)^{(6n-i)\kappa n^2} 6^{6n} [2(\tilde{c}_{\lambda n} + \tilde{b}_{\lambda n} + 1)]^{3n} [2(a+b+c)]^{3n} \\ & \leq 2^{30n} (a+b)^{6\kappa n^3} \left(\frac{c}{a+b} \right)^{(6n-1)\kappa n^2} [\tilde{c}_{\lambda n} + \tilde{b}_{\lambda n} + 1]^{3n} [a+b+c]^{3n} \\ & \leq 2^{30n} [\Theta(1)]^{\lambda n^2} c^{6\kappa n^3} \frac{1}{c^{\kappa n^2}} (a+b)^{\kappa n^2} \\ & = 2^{30n} [\Theta(1)]^{\lambda n^2} c^{6\kappa n^3} \frac{1}{\gamma^{\kappa n^2}}, \end{aligned}$$

which is $< \frac{1}{4} c^{6\kappa n^3}$ when $\kappa \gg \lambda \geq 1$ is large. Here $\Theta(1)$ is a constant depending only on a, b, c , and 6^{6n} comes from six possible valid configurations for at most $6n$ vertices of the type D, E, F, G, H, I , not necessarily in locking configurations. $\square$

3 ON GENERAL 4-REGULAR GRAPHS (THE EIGHT-VERTEX MODEL)

In this chapter, we study the approximate complexity of the eight-vertex model on general 4-regular graphs in terms of the parameters (see Figure 1.2).

Notation. In this chapter and Chapter 4, we adopt the following notations assuming $a, b, c, d \geq 0$.

- $\mathcal{A} = \{(a, b, c, d) \mid a \leq b + c + d\}$, $\mathcal{B} = \{(a, b, c, d) \mid b \leq a + c + d\}$, $\mathcal{C} = \{(a, b, c, d) \mid c \leq a + b + d\}$,
 $\mathcal{D} = \{(a, b, c, d) \mid d \leq a + b + c\}$;
- $\mathcal{AD} = \{(a, b, c, d) \mid a + d \leq b + c\}$, $\mathcal{BD} = \{(a, b, c, d) \mid b + d \leq a + c\}$, $\mathcal{CD} = \{(a, b, c, d) \mid c + d \leq a + b\}$.

Remark 3.0.1. $\mathcal{AD} \subset \mathcal{A} \cap \mathcal{D}$, $\mathcal{BD} \subset \mathcal{B} \cap \mathcal{D}$, $\mathcal{CD} \subset \mathcal{C} \cap \mathcal{D}$. $\mathcal{X} = \mathcal{A} \cap \mathcal{B} \cap \mathcal{C} \cap \mathcal{D}$. $\mathcal{Y} = \mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD}$.

In addition, we abuse the notation and use $\overline{\mathcal{C}} = \{(a, b, c, d) \mid c \geq a + b + d\}$, $\overline{\mathcal{D}} = \{(a, b, c, d) \mid d \geq a + b + c\}$, $\overline{\mathcal{AD}} = \{(a, b, c, d) \mid a + d \geq b + c\}$, $\overline{\mathcal{BD}} = \{(a, b, c, d) \mid b + d \geq a + c\}$, and $\overline{\mathcal{CD}} = \{(a, b, c, d) \mid c + d \geq a + b\}$.

To prove the FPRAS result in Section 3.2, a most important building block is a set of *closure properties*. We prove these closure properties for the eight-vertex model in Section 3.1. We then use these closure properties to show that a Markov chain designed originally only for the six-vertex model can be adapted to provide our FPRAS. The Markov chain we adapt is the *directed-loop algorithm* which was invented by Rahman and Stillinger [RS72] and is widely used for the six-vertex model (e.g., [YN79, BN98, SZ04]). The state space of 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 (MCMC)* 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* [JS89, DFK91, Sin92, Jer03], (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) [JVV86]. All three parts depend on the closure properties. Specifically, we show that when $(a, b, c, d) \in \mathcal{Z}$, 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 $(a, b, c, d) \in \mathcal{Y}$, this ratio is indeed polynomially bounded according to the closure properties. Finally a self-reduction whose success in $\mathcal{Y}$ requires an additional closure property. Therefore, there is an FPRAS in the intersection of $\mathcal{Y}$ and $\mathcal{Z}$.

The closure properties are keys to our FPRAS. We use the term *a 4-ary construction* to denote a 4-regular graph Γ having four “dangling” edges, and consider all configurations on the edges of Γ where every vertex satisfies the even orientation rule and has arrow reversal symmetry. We can prove that this Γ defines a *constraint function* of arity 4 that also satisfies the even orientation rule and has arrow reversal symmetry. If we imagine the graph Γ is shrunken to a single point except the 4 dangling 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 $a', b', c', d' \geq 0$.

In Theorem 3.1.2 we show that the set of 4-ary constraint functions in $\mathcal{Y}$ is closed under 4-ary constructions. This is achieved by inventing a “*quantum decomposition*” of even-orientations. Theorem 2.1.1, a special case of Theorem 3.1.2 when $d = 0$, is proved for the six-vertex model using a decomposition of Eulerian orientations. Given $G = (V, E)$, every Eulerian orientation defines a set of $2^{|V|}$ directed *Eulerian partitions* by pairing up the four edges around a vertex in one of two ways such that each pair of edges satisfies “1-in-1-out”. However, such a decomposition does not exist when sinks and sources appear in the eight-vertex model.

In order to overcome this difficulty, we introduce a quantum decomposition where each vertex has a “signed” pairing. 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 gives rise to a quantum decomposition of $3^{|V|}$ “*annotated circuit partitions*”. (Details are in Section 3.1.) Although the idea of “pairings” and decompositions of

Eulerian orientations has been used before [Ver88, Jae90, MW96], the idea of a signed pairing and the associated quantum decomposition of even orientations into annotated circuit partitions is new. Just as statistical physicists introduce the eight-vertex model on the square lattice for certain desirable properties and better universality over the six-vertex model, in approximate complexity on 4-regular graphs our technique that gives FPRAS for the eight-vertex model extends significantly beyond those for the six-vertex model.

Not only more sophisticated techniques are needed, the landscape of approximate complexity for the eight-vertex model is also richer. In the six-vertex model we have $d = 0$. Then it follows that $\mathcal{Z} \subset \mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD}$ which means whenever the conductance of the directed-loop algorithm can be bounded by the ratio of near-even orientations over even orientations, there is an FPRAS. In the eight-vertex model, however, there are parameter settings in $\mathcal{Z}$ where the ratio can be exponentially large. This indicates that the current MCMC method is unable to give FPRAS for the whole region $\mathcal{Z}$, even though we can prove a nice upper bound for the conductance of this Markov chain.

Moreover, in the eight-vertex model we can give more positive results for planar graphs than for general graphs, unlike in the six-vertex model whenever we have an FPRAS for planar graphs we also have one for general graphs for the same parameters. For planar graphs, in Theorem 3.1.3 and Corollary 3.1.4 we show that the extra regions $\mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}} \cap \mathcal{X}$ and $\mathcal{AD} \cap \mathcal{BD} \cap \{(a, b, c, d) \mid c + d = a + b\}$ also enjoy closure properties. It turns out that (only) on planar graphs we have an FPRAS when the parameter setting is in the intersection of $\mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}} \cap \mathcal{X}$ and $\mathcal{Z}$. And since $\mathcal{Z} \subset \mathcal{X}$, combined with the FPRAS on general graphs, we get an FPRAS for $\mathcal{Z} \cap \mathcal{AD} \cap \mathcal{BD}$ for all planar graphs. This tolerance of dropping off the requirement $c + d \leq a + b$ is in perfect accordance with the special role that “saddle” configurations (Figure 1.1-5 and Figure 1.1-6) play on planar graphs.

In addition to the complexity result, in Section 3.1 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 3.1.1, we show that the set of 4-ary constraint functions lying in the

disordered phase $\mathcal{X}$ is closed under 4-ary constructions. We prove in Section 2.3 and Section 3.5 that approximation is NP-hard on $\overline{\mathcal{X}}$ by reductions from the problem of computing the maximum cut on a 3-regular graph. In Section 3.3 and Section 3.4, the complexity of approximating the partition function of the eight-vertex model is related to that of approximately counting perfect matchings.

3.1 Closure properties

Theorem 3.1.1. *The set of constraint functions in $\mathcal{X}$ is closed under 4-ary constructions, i.e., the constraint function of any 4-ary construction using constraint functions from the set $\mathcal{X}$ also belongs to the same set.*

Theorem 3.1.2. *The set of constraint functions in $\mathcal{Y}$ is closed under 4-ary constructions.*

Theorem 3.1.3. *The set of constraint functions in $AD \cap BD \cap \overline{CD} \cap \mathcal{X}$ is closed under 4-ary plane constructions.*

Corollary 3.1.4. *The set of constraint functions in $AD \cap BD \cap \{(a, b, c, d) \mid c + d = a + b\}$ is closed under 4-ary plane constructions.*

In order to prove the above closure properties, we introduce a quantum decomposition for the eight-vertex model, in which every even orientation of a 4-regular graph $G = (V, E)$ is a “superposition” of $3^{|V|}$ annotated circuit partitions (to be defined shortly).

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 3.1) which we denote by three special symbols: $\curvearrowright$, $\curvearrowleft$, $-|$, respectively. A *circuit partition* of a graph 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 $\varrho_v \in \{\curvearrowright, \curvearrowleft, -|\}$ 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.

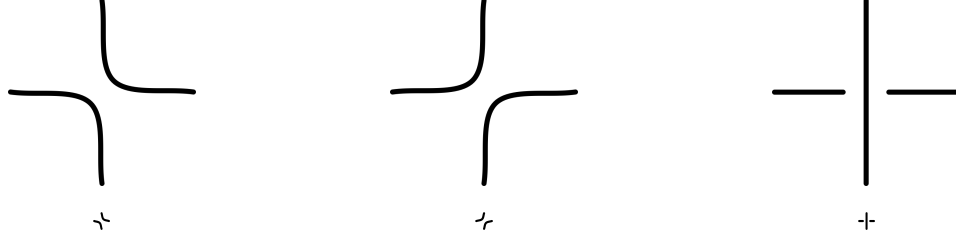


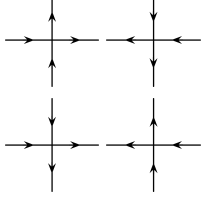
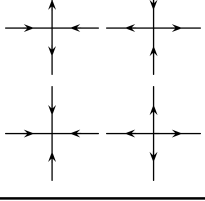
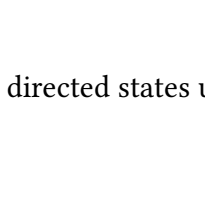
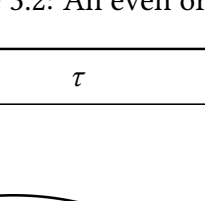
Figure 3.1: (Unsigned) pairings at a degree 4 vertex.

A *signed pairing* ϱ_v at v is a pairing with a sign, either *plus* (+) or *minus* (−). In other words, it is an element in $\{\curvearrowright, \curvearrowleft, \perp\} \times \{+, -\}$. 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 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 3.1. 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 .

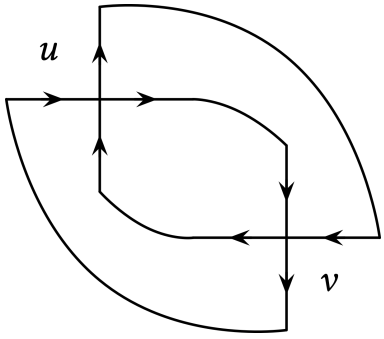
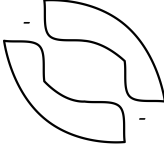
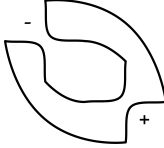
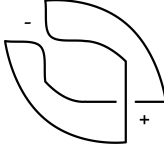
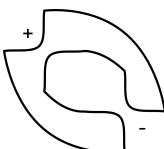
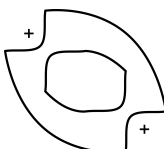
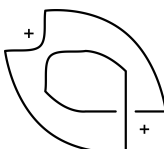
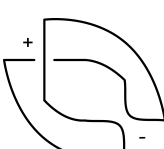
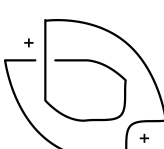
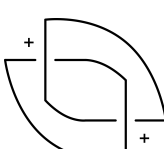
In this way, every even orientation τ defines $3^{|V|}$ *acp*’s, denoted by $\Phi(\tau)$. See Table 3.2 and Table 3.3 for two examples. Moreover, for any *acp* $\varphi \in \Phi(\tau)$, every circuit in φ is in one of the

Table 3.1: Map from eight local configurations to signed pairings.

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

two well-defined directed states under the orientation τ . Thus each even orientation τ defines $3^{|V|}$ $dacp$'s.

Table 3.2: An even orientation and its quantum decomposition into acp 's.

τ	$\Phi(\tau)$		
			
			
			

Conversely, for any $dacp$, if we ignore the signs at all vertices we get a valid even orientation

Table 3.3: Another even orientation and its quantum decomposition into *acp*'s.

τ	$\Phi(\tau)$		

(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. Define w to be a function assigning a weight to every signed pairing at every vertex and let 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. For every vertex in the eight-vertex model with the parameter setting (a, b, c, d) , we define w such that

$$\begin{cases} a = w(\swarrow_-) + w(\nearrow_+) + w(\dashv_+) \\ b = w(\swarrow_+) + w(\nearrow_-) + w(\dashv_+) \\ c = w(\swarrow_+) + w(\nearrow_+) + w(\dashv_-) \\ d = w(\swarrow_-) + w(\nearrow_-) + w(\dashv_-) \end{cases} \quad (3.1.1)$$

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 3.1.7 discusses this freedom). 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 (1.1.3), which is a product of sums by (3.1.1), 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 3.2 and Table 3.3. 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. $Z_{8V}(G) = 2[a^2 + b^2 + c^2 + d^2]$. When we expand $Z_{8V}(G)$ using (3.1.1) 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 3.2, 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 3.2 and Table 3.3.

Remark 3.1.1. While a weight function w satisfying (3.1.1) 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 (3.1.1), and the specification is independent of the choice of the weight function. E.g., the region $\mathcal{X}$ is specified by $\begin{cases} w(\mathcal{J}_+) + w(\mathcal{J}_-) + w(\mathcal{I}_-) \geq 0 \\ w(\mathcal{J}_-) + w(\mathcal{J}_+) + w(\mathcal{I}_-) \geq 0 \\ w(\mathcal{J}_-) + w(\mathcal{J}_-) + w(\mathcal{I}_+) \geq 0 \\ w(\mathcal{J}_+) + w(\mathcal{J}_+) + w(\mathcal{I}_+) \geq 0 \end{cases}$. Also $\mathcal{AD}$ is specified by $w(\mathcal{J}_-) \leq w(\mathcal{J}_+)$, $\mathcal{BD}$ by $w(\mathcal{J}_-) \leq w(\mathcal{J}_+)$, and $\mathcal{CD}$ by $w(\mathcal{I}_-) \leq w(\mathcal{I}_+)$. In Lemma 3.1.7, we will show that a nonnegative weight function w satisfying (3.1.1) exists iff $(a, b, c, d) \in \mathcal{X}$.

Remark 3.1.2. Although the association from even orientations to *dacp*'s is 1-to- $3^{|V|}$, non-overlapping, and surjective, the association from even orientations to *acp*'s is overlapping. If an *acp* has k circuits, it will be associated with 2^k even orientations. It is this many-to-many association, with corresponding weights, between even orientations and *acp*'s, that we call a *quantum decomposition*.

of eight-vertex model configurations, and each is expressed as a “superposition” (weighted sum) of acp 's[§].

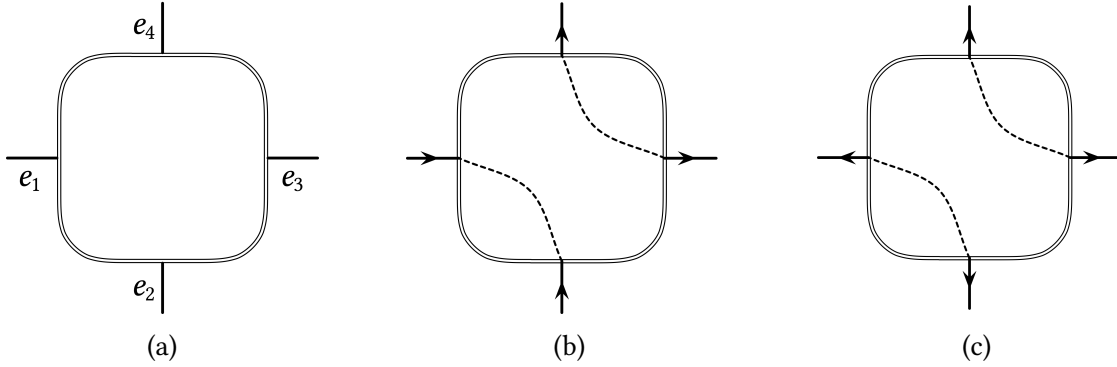


Figure 3.2: A 4-ary construction in the eight-vertex model.

A *4-ary construction* is a 4-regular graph Γ having four “dangling” edges (Figure 3.2a), and a constraint function on each node. It defines a 4-ary constraint function with these four dangling edges as input variables, when we sum the product of constraint function values on all vertices, over all configurations on the internal edges of Γ (i.e., the partial sum in the partition function $Z_{8V}(\Gamma)$ with a given boundary assignment on the dangling edges). If we imagine the graph Γ is shrunk to a single point except the 4 dangling 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 $a', b', c', d' \geq 0$. This is proved in the following lemma. 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 .

Lemma 3.1.5. *If constraint functions in Γ satisfy the even orientation rule and have arrow reversal symmetry, then the constraint function f defined by Γ also satisfies the even orientation rule and has arrow reversal symmetry.*

Proof. Consider any even orientation on Γ , and let Δ be the (sum of all in-degrees) – (the sum of all out-degrees). Each internal edge contrutes 0 to Δ . By the even orientation rule, at every vertex

[§]There is a superficial similarity between the quantum decomposition and skein relations in knot theory [Lic97]. This remains to be explored further.

this difference is 0 (mod 4). Thus $\Delta \equiv 0 \pmod{4}$. Thus among the dangling edges, it also satisfies the even orientation rule.

If we reverse all directions of an even orientation, which is an involution, each vertex contributes the same weight by arrow reversal symmetry. Hence f also has the arrow reversal symmetry. $\square$

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 dangling 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* ($datcp$'s). 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 (3.1.1).

Denote the constraint function of Γ by f and use the notations introduced in Section 1.3. 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 $datcp$'s. 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) $\{\xrightarrow{e_1} \square \xleftarrow{e_2}, \xleftarrow{e_3} \square \xrightarrow{e_4}\}$ and on both trails the numbers of minus pairings are odd; or
- (2) $\{\xrightarrow{e_1} \square \xrightarrow{e_4}, \xrightarrow{e_2} \square \xrightarrow{e_3}\}$ and on both trails the numbers of minus pairings are even (Figure 3.2b);

or

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

Let $\Phi_{0011, \curvearrowright_-}$, $\Phi_{0011, \curvearrowleft_+}$ and $\Phi_{0011, \dashv_+}$ be the subsets of $datcp$'s 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*'s by $W(\Phi) = \sum_{\varphi \in \Phi} \tilde{w}(\varphi)$ yields $f(0011) = W(\Phi_{0011, \nearrow_-}) + W(\Phi_{0011, \searrow_+}) + W(\Phi_{0011, \dashv_+})$. Similarly we can define $\Phi_{1100, \nearrow_-}$, $\Phi_{1100, \searrow_+}$ and $\Phi_{1100, \dashv_+}$, and get $f(1100) = W(\Phi_{1100, \nearrow_-}) + W(\Phi_{1100, \searrow_+}) + W(\Phi_{1100, \dashv_+})$. Note that there is a bijective weight-preserving map between $\Phi_{0011, \nearrow_-}$ and $\Phi_{1100, \nearrow_-}$ by reversing the direction of every circuit and trail of a *datcp*. Thus, $W(\Phi_{0011, \nearrow_-}) = W(\Phi_{1100, \nearrow_-})$, $W(\Phi_{0011, \searrow_+}) = W(\Phi_{1100, \searrow_+})$, and $W(\Phi_{0011, \dashv_+}) = W(\Phi_{1100, \dashv_+})$. 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 $b_i = b_j$. Then a further important observation is that for each *datcp* in $\Phi_{0011, \nearrow_-}$, if we only reverse every edge in the trail between $\xrightarrow{e_1} \square \xleftarrow{e_2}$ and keep the states of all circuits and the other trail unchanged, this *datcp* has the same weight but now lies in $\Phi_{1111, \nearrow_-}$. This is because at every vertex v , reversing the orientation of any one branch of the given (annotated) pairing $\varrho_v \in \{\nearrow, \searrow, \dashv\} \times \{+, -\}$ does not change the value $w(\varrho_v)$. In this way, we set up a one-to-one weight-preserving map between $\Phi_{0011, \nearrow_-}$ and $\Phi_{1111, \nearrow_-}$, hence $W(\Phi_{0011, \nearrow_-}) = W(\Phi_{1111, \nearrow_-})$. Combining the result in the last paragraph we have proved the first item below, and we name its common value $W(\nearrow_-)$. The other items are proved similarly.

- $W(\nearrow_-) = W(\Phi_{0011, \nearrow_-}) = W(\Phi_{1100, \nearrow_-}) = W(\Phi_{0000, \nearrow_-}) = W(\Phi_{1111, \nearrow_-})$;
- $W(\searrow_-) = W(\Phi_{0110, \searrow_-}) = W(\Phi_{1001, \searrow_-}) = W(\Phi_{0000, \searrow_-}) = W(\Phi_{1111, \searrow_-})$;
- $W(\dashv_-) = W(\Phi_{0101, \dashv_-}) = W(\Phi_{1010, \dashv_-}) = W(\Phi_{0000, \dashv_-}) = W(\Phi_{1111, \dashv_-})$;
- $W(\nearrow_+) = W(\Phi_{0110, \nearrow_+}) = W(\Phi_{1001, \nearrow_+}) = W(\Phi_{0101, \nearrow_+}) = W(\Phi_{1010, \nearrow_+})$;
- $W(\searrow_+) = W(\Phi_{0011, \searrow_+}) = W(\Phi_{1100, \searrow_+}) = W(\Phi_{0101, \searrow_+}) = W(\Phi_{1010, \searrow_+})$;
- $W(\dashv_+) = W(\Phi_{0011, \dashv_+}) = W(\Phi_{1100, \dashv_+}) = W(\Phi_{0110, \dashv_+}) = W(\Phi_{1001, \dashv_+})$.

Consequently, f has parameters $\begin{cases} a' = W(\nearrow_-) + W(\searrow_+) + W(\dashv_+) \\ b' = W(\nearrow_+) + W(\searrow_-) + W(\dashv_+) \\ c' = W(\nearrow_+) + W(\searrow_+) + W(\dashv_-) \\ d' = W(\nearrow_-) + W(\searrow_-) + W(\dashv_-) \end{cases}$.

Proof of Theorem 3.1.1. For any weight function satisfying (3.1.1), one can easily verify that $(a, b, c, d) \in \mathcal{X}$ iff the following inequalities hold:
$$\begin{cases} w(\mathfrak{L}_+) + w(\mathcal{J}_-) + w(\mathfrak{L}_-) \geq 0 \\ w(\mathfrak{L}_-) + w(\mathcal{J}_+) + w(\mathfrak{L}_+) \geq 0 \\ w(\mathfrak{L}_-) + w(\mathcal{J}_-) + w(\mathfrak{L}_+) \geq 0 \\ w(\mathfrak{L}_+) + w(\mathcal{J}_+) + w(\mathfrak{L}_-) \geq 0 \end{cases}.$$

By Lemma 3.1.7 (stated later), we can assume w is a nonnegative weight function. By definition, each of the six quantities $W(\mathfrak{L}_+)$, $W(\mathfrak{L}_-)$, $W(\mathcal{J}_+)$, $W(\mathcal{J}_-)$, $W(\mathfrak{L}_+)$ and $W(\mathfrak{L}_-)$ is a sum over a set of *datcp*'s of products of values of w , and thus they are all nonnegative. Hence, the constraint function f defined by Γ satisfies
$$\begin{cases} W(\mathfrak{L}_+) + W(\mathcal{J}_-) + W(\mathfrak{L}_-) \geq 0 \\ W(\mathfrak{L}_-) + W(\mathcal{J}_+) + W(\mathfrak{L}_+) \geq 0 \\ W(\mathfrak{L}_-) + W(\mathcal{J}_-) + W(\mathfrak{L}_+) \geq 0 \\ W(\mathfrak{L}_+) + W(\mathcal{J}_+) + W(\mathfrak{L}_-) \geq 0 \end{cases}.$$
 This is equivalent to the assertion that the parameters (a', b', c', d') of f belong to the region $\mathcal{X}$. $\square$

Proof of Theorem 3.1.2. By definition $(a, b, c, d) \in \mathcal{Y}$ 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 (3.1.1) this is equivalent to
$$\begin{cases} w(\mathfrak{L}_+) \geq w(\mathfrak{L}_-) \\ w(\mathcal{J}_+) \geq w(\mathcal{J}_-) \\ w(\mathfrak{L}_+) \geq w(\mathfrak{L}_-) \end{cases}.$$
 Since $\mathcal{Y} \subset \mathcal{X}$, by Lemma 3.1.7 we can assume w is nonnegative. To prove Theorem 3.1.2 we only need to establish
$$\begin{cases} W(\mathfrak{L}_+) \geq W(\mathfrak{L}_-) \\ W(\mathcal{J}_+) \geq W(\mathcal{J}_-) \\ W(\mathfrak{L}_+) \geq W(\mathfrak{L}_-) \end{cases}.$$
 We prove $W(\mathfrak{L}_+) \geq W(\mathfrak{L}_-)$. Proof for the other two inequalities is symmetric.

An *atcp* is a *tcp* together with a valid annotation. Consider the set Ψ of *tcp*'s 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 $W(\mathfrak{L}_+) \geq W(\mathfrak{L}_-)$, 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 $N = 0$. Let us first also 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 $w(\varrho_v)$ 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 $w_+(v) = w(\varrho_+)$, and $w_-(v) = w(\varrho_-)$, 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 $w(\varrho_v)$ 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

$$S_+(C) - S_-(C) = \prod_{i=1}^s (w_+(x_i) - w_-(x_i)) \geq 0, \quad T_+(C) - T_-(C) = \prod_{j=1}^t (w_+(y_j) - w_-(y_j)) \geq 0.$$

Both differences are nonnegative by the hypothesis of Theorem 3.1.2.

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 ($1 \leq l \leq q$). 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 $p' + q' \equiv 0 \pmod{2}$.

Now let us deal with the case when there are self-intersecting trails or circuits. Suppose v is a self-intersecting vertex. Let its four incident edges be $\{e, f, g, h\}$. Without loss of generality we assume the pairing ϱ_v in ψ is $\{e, f\}$ and $\{g, h\}$ (Figure 3.3a). Define Γ' to be the 4-ary construction obtained from Γ by deleting v , and merging e with f , and g with h (Figure 3.3b). Define W'_+ and W'_- similarly for Γ' with ψ' being $\psi \setminus \{\varrho_v\}$.

Since v contributes either zero or two - to the trail or circuit it belongs to, an annotation is valid for ψ iff its restriction on Γ' is valid for ψ' . Moreover, for every valid annotation of vertices in ψ' contributing a factor to W'_+ (or W'_-), if we impose an arbitrary sign on ϱ_v , we get a valid annotation for ψ contributing a factor to W_+ (or W_- , respectively). If the sign of the annotation at v is + (or - respectively) then each product term in W'_+ or W'_- gains the same extra factor $w_+(v)$. If the annotation at v is -, then they gain the factor $w_-(v)$. Therefore, we have $W_+ - W_- = (w_+(v) + w_-(v))(W'_+ - W'_-)$. Hence $W_+ \geq W_-$ if $W'_+ \geq W'_-$. Thus we have reduced from Γ to Γ' which has one fewer self-intersections. Repeating this finitely many times we end up with no self-intersections.

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 3.3a). Define Γ' to be the 4-ary construction obtained from Γ by deleting v and merging e with f , and g with h (Figure 3.3b). Define Γ'' to be the 4-ary construction obtained from Γ by deleting v and merging e

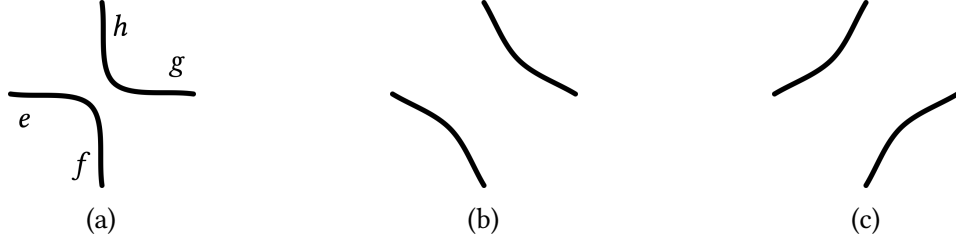


Figure 3.3: Possible ways of deleting a vertex. The vertex (not explicitly shown) at the center of part (a) is removed in part (b) and (c).

with h , and f with g (Figure 3.3c). 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

$$W_+ - W_- = w_+(v)[W_+ - W_-]_e + w_-(v)[W_+ - W_-]_o, \quad (3.1.2)$$

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

$$W'_+ - W'_- = [W_+ - W_-]_e, \quad (3.1.3)$$

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

$$W''_+ - W''_- = [W_+ - W_-]_e + [W_+ - W_-]_o, \quad (3.1.4)$$

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 (3.1.2,3.1.3,3.1.4) we have

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

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

Proof of Theorem 3.1.3. By Lemma 3.1.7 (stated later), for $(a, b, c, d) \in \mathcal{X}$ we can choose a nonnegative function w to satisfy (3.1.1). It is easily verified that for any weight function w satisfying (3.1.1), $w(\curvearrowright_+) \geq w(\curvearrowright_-)$ iff $a + d \leq b + c$ (in $\mathcal{AD}$), $w(\curvearrowleft_+) \geq w(\curvearrowleft_-)$ iff $b + d \leq a + c$ (in $\mathcal{BD}$), $w(\dashv_+) \leq w(\dashv_-)$ iff $c + d \geq a + b$ (in $\overline{\mathcal{CD}}$). Since $(a, b, c, d) \in \mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}} \cap \mathcal{X}$ by hypothesis, we have a nonnegative function w satisfying (3.1.1) and
$$\begin{cases} w(\curvearrowright_+) \geq w(\curvearrowright_-) \\ w(\curvearrowleft_+) \geq w(\curvearrowleft_-) \\ w(\dashv_+) \leq w(\dashv_-) \end{cases}.$$

We say a *tcp* ψ of a 4-ary construction has type- $\curvearrowright$ if its two trails connect dangling edges e_1 with e_2 and e_3 with e_4 , type- $\curvearrowleft$ if they connect e_1 with e_4 and e_2 with e_3 , and type- $\dashv$ if they connect e_1 with e_3 and e_2 with e_4 . Sometimes we also say a pairing $\varrho \in \{\curvearrowright, \curvearrowleft, \dashv\}$ (without a sign) has type- ϱ .

We prove this theorem not only for 4-ary plane constructions, but for any 4-ary construction Γ that satisfies the following property $\mathcal{P}$.

For any *tcp* ψ of Γ the number of vertices that have type- $\dashv$ pairings shared: (1) by any two distinct circuits is even; (2) by a trail and a circuit is even; (3) by two trails is even, if ψ has type- $\curvearrowright$ or type- $\curvearrowleft$; and (4) by two trails is odd, if ψ has type- $\dashv$. ($\mathcal{P}$)

Observe that every 4-ary plane construction satisfies property $\mathcal{P}$ by *Jordan Curve Theorem*.

The structure of this proof is similar to that of the proof of Theorem 3.1.2, but the details are more delicate because of the reversed inequality $w(\dashv_+) \leq w(\dashv_-)$, which we need to use property

$\mathcal{P}$ and a parity argument to finesse.

Inheriting notations from the proof of Theorem 3.1.2, we prove that for any $tcp \psi \in \Psi$, $W_+ \geq W_-$ if ψ has type- $\curvearrowright$ or type- $\curvearrowleft$; and $W_+ \leq W_-$ if ψ has type- $-|-$. We prove this statement still by induction on the number N of vertices shared by any two distinct circuits in ψ .

Base case: The base case is $N = 0$. Let us first assume that no trail or circuit is self-intersecting. Consider the case ψ has type- $\curvearrowright$. 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.

For a circuit $C \in \psi$, by property $\mathcal{P}$ the number of vertices it shares with a trail that have a type- $-|-$ pairing is even. Denote the number of vertices it shares with χ_{12} that have a type- $\curvearrowright$ or type- $\curvearrowleft$ pairing by s and those that have a type- $-|-$ pairing by s' , and let the vertices be x_i ($1 \leq i \leq s$) and x'_i ($1 \leq i \leq s'$) respectively; similarly denote the number of vertices it shares with χ_{34} that have a type- $\curvearrowright$ or type- $\curvearrowleft$ pairing by t and those that have a type- $-|-$ pairing by t' , and let the vertices be y_j ($1 \leq j \leq t$) and y'_j ($1 \leq j \leq t'$) respectively (the following statement is still true if there is any zero among s, s', t, t'). Define the quantities $S_+(C), S_-(C), T_+(C)$ and $T_-(C)$ as in the proof of Theorem 3.1.2, then we have

$$S_+(C) - S_-(C) = \prod_{i=1}^{s'} (w_+(x'_i) - w_-(x'_i)) \prod_{i=1}^s (w_+(x_i) - w_-(x_i)),$$

$$T_+(C) - T_-(C) = \prod_{j=1}^{t'} (w_+(y'_j) - w_-(y'_j)) \prod_{j=1}^t (w_+(y_j) - w_-(y_j)).$$

We have $S_+(C) \geq S_-(C)$ because each $w_+(x_i) - w_-(x_i) \geq 0$ and each $w_+(x'_i) - w_-(x'_i) \leq 0$ but s' is even. By the same argument, $T_+(C) \geq T_-(C)$.

Now we account for the shared vertices between the two trails. According to property $\mathcal{P}$, the number of vertices shared by χ_{12} and χ_{34} that have a type- $-|-$ pairing must also be even. Denote the number of vertices in ψ shared by χ_{12} and χ_{34} that have a type- $\curvearrowright$ or type- $\curvearrowleft$ pairing by p and those that have a type- $-|-$ pairing by p' , and denote these vertices by z_k ($1 \leq k \leq p$) and z'_k ($1 \leq k \leq p'$)

(again the following is still true if p or p' is 0). Then, by the same proof,

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

Since $S_+(C_l)T_+(C_l) \geq S_-(C_l)T_-(C_l)$ ($1 \leq l \leq q$), $w_+(z_k) - w_-(z_k) \geq 0$ ($1 \leq k \leq p$), and $w_+(z'_k) - w_-(z'_k) \leq 0$ ($1 \leq k \leq p'$ and p' is even), we get $W_+ \geq W_-$.

The same proof applies for ψ of type- $\mathcal{J}$. For type- $\mathcal{I}$ we have the corresponding p' odd, and thus $W_+ \leq W_-$.

The way to deal with self-intersections is exactly the same as in the proof of Theorem 3.1.2. We will not repeat here.

Induction step: When the pairings at intersections between distinct circuits are all of type- $\mathcal{K}$ or type- $\mathcal{J}$ only, our proof is the same as the induction step of the proof of Theorem 3.1.2. We only note that the constructions of Γ' and Γ'' in that proof preserve property $\mathcal{P}$. When there are type- $\mathcal{I}$ intersections between distinct circuits, we show how to reduce to the previous case by getting rid of all type- $\mathcal{I}$ intersections while preserving property $\mathcal{P}$.

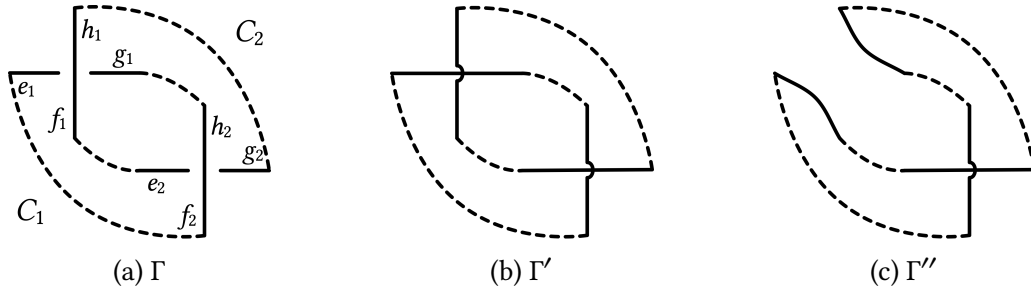


Figure 3.4: Possible ways of deleting two vertices with intersections of type- $\mathcal{I}$.

For any two circuits C_1 and C_2 in ψ , the number of intersections of type- $\mathcal{I}$ between these two circuits must be even (according to property $\mathcal{P}$). Suppose this number is not zero, let u and v be two vertices with intersections of type- $\mathcal{I}$ between C_1 and C_2 . For $b \in \{1, 2\}$, let $\{e_b, f_b, g_b, h_b\}$ be the edges incident to u and v in Γ respectively. We may assume the pairings at u and v are $\{e_b, g_b\}$ and $\{f_b, h_b\}$ and thus e_b, g_b are in one circuit, while f_b, h_b are in another circuit. Furthermore, we may name the edges so that e_1, g_1 and f_2, h_2 are in the same circuit, say C_1 , and e_2, g_2 and f_1, h_1 are

in another circuit (in this case C_2) (Figure 3.4a). Define Γ' to be the 4-ary construction obtained from Γ by deleting u, v and merging e_b with g_b , and f_b with h_b for $b \in \{1, 2\}$ (Figure 3.4b). Define Γ'' to be the 4-ary construction obtained from Γ by deleting u, v and merging e_1 with f_1 , g_1 with h_1 , e_2 with g_2 , and f_2 with h_2 (Figure 3.4c). Note that in Γ' , we have two circuits C'_1 and C'_2 (each has two fewer vertices u and 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 $\psi' = \psi \setminus \{\varrho_u, \varrho_v\}$.

We can decompose $W_+ - W_-$ according to whether the signs on ϱ_u and ϱ_v are $+$ or $-$. Recall that for any valid annotation of ψ , one encounters an even number of $-$ along C_1 and C_2 . If the signs on ϱ_u and ϱ_v are both $+$ or both $-$, the number of $-$ along C_1 (and C_2) at all vertices other than u and v in any valid annotation is always even; if the signs on ϱ_u and ϱ_v are different (one $+$ and one $-$), this number (for both C_1 and C_2) is always odd. $W_+ - W_-$ can be decomposed into four parts, corresponding to terms with the signs on ϱ_u and ϱ_v being $\pm$. So we can write

$$W_+ - W_- = w_+(u)w_+(v)[W_+ - W_-]_{++} + w_-(u)w_-(v)[W_+ - W_-]_{--} \quad (3.1.5)$$

$$+ w_+(u)w_-(v)[W_+ - W_-]_{+-} + w_-(u)w_+(v)[W_+ - W_-]_{-+}, \quad (3.1.6)$$

where $[W_+ - W_-]_{\pm\pm}$ collect terms in $W_+ - W_-$ in the respective parts, but without the factors at u and v .

Let X (respectively X') be the set of vertices of C_1 (excluding u, v) between e_1 and f_2 (respectively between g_1 and h_2). Let Y (respectively Y') be the set of vertices of C_2 (excluding u, v) between h_1 and g_2 (respectively between f_1 and e_2). If we write $\sigma(x) = 1$ if the annotation on x is $-$, and $\sigma(x) = 0$ otherwise, then the requirement for an annotation on C'_1 and C'_2 to be valid is $\sum_{x \in X \cup X'} \sigma(x) \equiv \sum_{x \in Y \cup Y'} \sigma(x) \equiv 0 \pmod{2}$. This is equivalent to requiring an extension that assigns the same sign to both u and v (either $(++)$ or $(--)$) to be a valid annotation on C_1 and C_2 . The latter

is just $\sum_{x \in X \cup X' \cup \{u, v\}} \sigma(x) \equiv \sum_{x \in Y \cup Y' \cup \{u, v\}} \sigma(x) \equiv 0 \pmod{2}$, conditioned on $\sigma(u) = \sigma(v)$. Hence

$$W'_+ - W'_- = [W_+ - W_-]_{++} = [W_+ - W_-]_{--} \quad (3.1.7)$$

The requirement for an annotation to be valid on C^* is $\sum_{x \in X \cup X' \cup Y \cup Y'} \sigma(x) \equiv 0 \pmod{2}$, which is equivalent to either $\sum_{x \in X \cup X'} \sigma(x) \equiv \sum_{x \in Y \cup Y'} \sigma(x) \equiv 0 \pmod{2}$, or $\sum_{x \in X \cup X'} \sigma(x) \equiv \sum_{x \in Y \cup Y'} \sigma(x) \equiv 1 \pmod{2}$. This is equivalent to combining two types of extensions to a valid annotation on C_1 and C_2 , where type (1) assigns the same sign to both u and v (either $(++)$ or $(--)$), or type (2) assigns different signs to u and v (either $(+-)$ or $(-+)$). Hence, in addition to (3.1.7) we have

$$[W_+ - W_-]_{+-} = [W_+ - W_-]_{-+}, \quad (3.1.8)$$

and also

$$W''_+ - W''_- = [W_+ - W_-]_{++} + [W_+ - W_-]_{+-}. \quad (3.1.9)$$

It follows from (3.1.5, 3.1.8, 3.1.7, 3.1.9) that

$$\begin{aligned} & W_+ - W_- \\ &= (w_+(u)w_+(v) + w_-(u)w_-(v))[W_+ - W_-]_{++} + (w_+(u)w_-(v) + w_-(u)w_+(v))[W_+ - W_-]_{+-} \\ &= [(w_+(u) - w_-(u))(w_+(v) - w_-(v))](W'_+ - W'_-) + [w_+(u)w_-(v) + w_-(u)w_+(v)](W''_+ - W''_-). \end{aligned}$$

Note that if Γ satisfies property $\mathcal{P}$, Γ' and Γ'' also satisfy property $\mathcal{P}$, but with fewer intersections of type $--|$ between distinct circuits. By induction, both $W'_+ - W'_- \geq 0$ and $W''_+ - W''_- \geq 0$. Since $w_+(u) \leq w_-(u)$ and $w_+(v) \leq w_-(v)$ are given by hypothesis, the product $(w_+(u) - w_-(u))(w_+(v) - w_-(v)) \geq 0$. Also $w_+(u)w_-(v) + w_-(u)w_+(v) \geq 0$ as w is nonnegative. Therefore, $W_+ \geq W_-$.

We have finished the proof for $W(\mathcal{N}_+) \geq W(\mathcal{N}_-)$. The proof for $W(\mathcal{J}_+) \geq W(\mathcal{J}_-)$ is the same. The proof for $W(-|_+) \leq W(-|_-)$ can be adapted. We only need to note that the two trails in Γ' and Γ'' are unchanged from Γ , thus both are still of type $--|$. Thus inductively we have $W'_+ - W'_- \leq 0$ and

$W_+'' - W_-'' \leq 0$, and thus $W_+ - W_- \leq 0$ as a nonnegative combination of these two quantities. $\square$

Corollary 3.1.4 follows immediately from Theorem 3.1.2 and Theorem 3.1.3 since $\{(a, b, c, d) \mid c + d = a + b\} = CD \cap \overline{CD}$ and $\mathcal{Y} \subset \mathcal{X}$, and therefore the intersection of the two regions $\mathcal{Y}$ and $AD \cap BD \cap \overline{CD} \cap \mathcal{X}$ is precisely $AD \cap BD \cap \{(a, b, c, d) \mid c + d = a + b\}$.

Lemma 3.1.6. *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 $\tilde{x}, \tilde{x}', \tilde{y}, \tilde{y}', \tilde{z}, \tilde{z}'$.*

Proof. The condition is obviously symmetric so that there is a symmetry group $S_2 \times S_2 \times S_2$ acting on $\{x, x', y, y', z, z'\}$. Thus, we may assume without loss of generality that $x \leq x', y \leq y', z \leq z'$. Let α, β be two distinct symbols among x, y, z . For any $c \in \mathbb{R}$, if we add c to α and α' , and subtract c from β and β' , the eight sums $X + Y + Z$ are unchanged, because in each $X + Y + Z$ exactly one of α and α' appears once and also exactly one of β and β' appears once.

Note that $x + y + z \geq 0$. In two steps we can replace $\{x, x', y, y', z, z'\}$ by

$$\begin{aligned} x &\rightarrow x + y + z, & x' &\rightarrow x' + y + z \\ y &\rightarrow y - y = 0, & y' &\rightarrow y' - y \\ z &\rightarrow z - z = 0, & z' &\rightarrow z' - z \end{aligned}$$

This completes the proof. $\square$

Lemma 3.1.7. *The parameter setting (a, b, c, d) belongs to $\mathcal{X}$ iff there exists a nonnegative weight function w satisfying (3.1.1).*

Proof. The assignment of the weight function w satisfying (3.1.1) can be viewed as a linear system on six variables $(x, x', y, y', z, z') = (w(\swarrow_+), w(\swarrow_-), w(\searrow_+), w(\searrow_-), w(\dashv_+), w(\dashv_-))$. This linear system has rank 4 and therefore there is a nonempty solution space of dimension 2.

Pick any solution to (3.1.1). Recall that membership $(a, b, c, d) \in \mathcal{X}$ is characterized by

$$\begin{cases} w(\neg_+) + w(\neg_-) + w(\neg_+) \geq 0 \\ w(\neg_+) + w(\neg_-) + w(\neg_+) \geq 0 \\ w(\neg_+) + w(\neg_-) + w(\neg_+) \geq 0 \\ w(\neg_+) + w(\neg_-) + w(\neg_+) \geq 0 \end{cases}.$$

We also have $a, b, c, d \geq 0$. Hence we can apply Lemma 3.1.6 and get a nonnegative valued w satisfying (3.1.1).

The reverse direction is obvious, once we realize that membership in $\mathcal{X}$ is characterized by the four inequalities above, for any solution to (3.1.1). $\square$

Notation. Fix for each vertex v in a 4-regular graph G a weight function w on signed pairings (satisfying (3.1.1) at v). Let $Z_v(\varrho)$ be the weighted sum of the set of all $dacp$ having the signed pairing ϱ at v .

Corollary 3.1.8. *If at each vertex in a 4-regular graph G we have a nonnegative weight function w such that $w(\neg_+) \geq w(\neg_-)$, $w(\neg_+) \geq w(\neg_-)$, and $w(\neg_+) \geq w(\neg_-)$, then $Z_v(\neg_+) \geq Z_v(\neg_-)$, $Z_v(\neg_+) \geq Z_v(\neg_-)$, and $Z_v(\neg_+) \geq Z_v(\neg_-)$ at each vertex v in G .*

Proof. Let (a, b, c, d) be the parameters of the constraint function at a vertex v . We first collect terms in the partition function $Z_{8V}(G)$ according to which of the 8 local configurations in Figure 1.1 v is in. If we remove the vertex v from G , the rest of G forms a 4-ary construction whose dangling edges are those incident to v . Using notations for 4-ary constructions, we can write $Z_{8V}(G)$ as

$$\begin{aligned} & 2a \left[W(\Phi_{0011}, \neg_+) + W(\Phi_{0011}, \neg_-) + W(\Phi_{0011}, \neg_+) \right] + 2b \left[W(\Phi_{0110}, \neg_+) + W(\Phi_{0110}, \neg_-) + W(\Phi_{0110}, \neg_+) \right] + \\ & 2c \left[W(\Phi_{0101}, \neg_+) + W(\Phi_{0101}, \neg_-) + W(\Phi_{0101}, \neg_+) \right] + 2d \left[W(\Phi_{0000}, \neg_+) + W(\Phi_{0000}, \neg_-) + W(\Phi_{0000}, \neg_+) \right] \\ &= 2(w(\neg_+) + w(\neg_-) + w(\neg_+)) [W(\neg_+) + W(\neg_-) + W(\neg_+)] + 2(w(\neg_+) + w(\neg_-) + w(\neg_+)) [W(\neg_+) + W(\neg_-) + W(\neg_+)] + \\ & 2(w(\neg_+) + w(\neg_-) + w(\neg_+)) [W(\neg_+) + W(\neg_-) + W(\neg_+)] + 2(w(\neg_+) + w(\neg_-) + w(\neg_+)) [W(\neg_+) + W(\neg_-) + W(\neg_+)]. \end{aligned}$$

Now we collect terms according to the 6 signed pairings ϱ at v . These are precisely $Z_v(\neg_+)$, $Z_v(\neg_-)$, $Z_v(\neg_+)$, $Z_v(\neg_-)$, $Z_v(\neg_+)$, and $Z_v(\neg_-)$ respectively, and $Z_{8V}(G)$ is the sum of these 6 terms

$$\begin{aligned} & 2w(\neg_+) [2W(\neg_+) + W(\neg_+) + W(\neg_-) + W(\neg_+) + W(\neg_-)] + 2w(\neg_-) [2W(\neg_-) + W(\neg_+) + W(\neg_-) + W(\neg_+) + W(\neg_-)] + \\ & 2w(\neg_+) [2W(\neg_+) + W(\neg_+) + W(\neg_-) + W(\neg_+) + W(\neg_-)] + 2w(\neg_-) [2W(\neg_-) + W(\neg_+) + W(\neg_-) + W(\neg_+) + W(\neg_-)] + \\ & 2w(\neg_+) [2W(\neg_+) + W(\neg_+) + W(\neg_-) + W(\neg_+) + W(\neg_-)] + 2w(\neg_-) [2W(\neg_-) + W(\neg_+) + W(\neg_-) + W(\neg_+) + W(\neg_-)]. \end{aligned}$$

Notice the common multipliers when comparing the three pairs $Z_v(\nearrow_+)$ vs. $Z_v(\nwarrow_-)$, $Z_v(\searrow_+)$ vs. $Z_v(\swarrow_-)$, and $Z_v(-|_+)$ vs. $Z_v(-|-)$. The corollary follows because $w(\nearrow_+) \geq w(\nwarrow_-)$, $w(\searrow_+) \geq w(\swarrow_-)$, and $w(-|_+) \geq w(-|-)$ by the assumption, and $W(\nearrow_+) \geq W(\nwarrow_-)$, $W(\searrow_+) \geq W(\swarrow_-)$, and $W(-|_+) \geq W(-|-)$ by Theorem 3.1.2. $\square$

Corollary 3.1.9. *If at each vertex in a 4-regular plane graph G we have a nonnegative weight function w such that $w(\nearrow_+) \geq w(\nwarrow_-)$, $w(\searrow_+) \geq w(\swarrow_-)$, and $w(-|_+) \leq w(-|-)$, then $Z_v(\nearrow_+) \geq Z_v(\nwarrow_-)$, $Z_v(\searrow_+) \geq Z_v(\swarrow_-)$, and $Z_v(-|_+) \leq Z_v(-|-)$ at each vertex v in G .*

Proof. The proof is similar to that of Corollary 3.1.8, with the only difference that we are given $w(-|_+) \leq w(-|-)$, and we have $W(-|_+) \leq W(-|-)$ by Theorem 3.1.3. $\square$

3.2 An FPRAS via Markov chains Monte Carlo

Theorem 3.2.1. *There is an FPRAS for $Z_{8V}(a, b, c, d)$ if $(a, b, c, d) \in \mathcal{Y} \cap \mathcal{Z}$.*

Theorem 3.2.2. *There is an FPRAS for $Z_{8V}(a, b, c, d)$ on planar graphs if $(a, b, c, d) \in \mathcal{Z} \cap \mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}}$.*

Remark 3.2.1. Our FPRAS result is actually stronger. The FPRAS in Theorem 3.2.1 for general graphs (including planar graphs) works even if different constraint functions from $\mathcal{Y} \cap \mathcal{Z}$ are assigned at different vertices. Similarly the FPRAS in Theorem 3.2.2 for (only) planar graphs works even if different constraint functions from $\mathcal{Z} \cap \mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}}$ are assigned at different vertices. (For logical reasons concerning models of computation, the functions should take values in algebraic numbers, and if these functions are not chosen from a fixed finite set then the description of each constraint function used must be included in the input. In this section for simplicity, we assume all constraint functions are from a fixed finite subset.)

We design our FPRAS using the common approach of approximately counting via almost uniformly sampling [JVV86, JS89, DFK91, Sin92, Jer03] by showing that a Markov chain designed for the six-vertex model can be adapted for the eight-vertex model. The Markov chain we adapt is the *directed-loop algorithm* which was invented by Rahman and Stillinger [RS72] and is widely

used for the six-vertex model (e.g., [YN79, BN98, SZ04]). The state space of our Markov chain $\mathcal{MC}$ 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 the two defects on edges. Some examples of the states in the directed-loop algorithm are shown in Figure 3.5 where the state in Figure 3.5a is an even orientation and the state in Figure 3.5b and the state in Figure 3.5c are near-even orientations with exactly two defects. Some typical moves in the directed-loop algorithm are as follows: the transition from the state in Figure 3.5a to the state in Figure 3.5b creates two defects; the transition from the state in Figure 3.5b to the state in Figure 3.5a merges two defects; the transitions between Figure 3.5b and Figure 3.5c shift one of the defects. (Formal description of this Markov chain will be given shortly.)

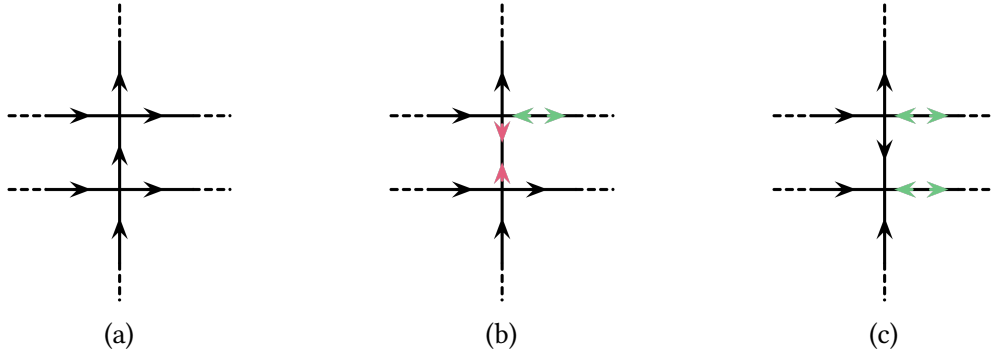


Figure 3.5: Examples of the states in the directed-loop algorithm.

Notation. For a 4-regular graph, denote the set of even orientations by Ω_0 and the set of near-even orientations by Ω_2 . The state space of $\mathcal{MC}$ is $\Omega = \Omega_0 \cup \Omega_2$. Let $Z(S)$ be the weighted sum of states in the set S .

Proof of Theorem 3.2.1. We will show (later) 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. In this proof we show that for a 4-regular graph, if all constraint functions used in an instance belong to $\mathcal{Y} \cap \mathcal{Z}$, then

- (1) the $\mathcal{MC}$ is rapidly mixing via a *conductance argument* [JS89, DFK91, Sin92, Jer03];
- (2) even orientations take a non-negligible proportion in the state space;
- (3) there exists a self-reduction (to reduce the computation of the partition function of a graph to that of a “smaller” graph) [JVV86].

We remark that all three parts (1)(2)(3) depend on the idea of quantum decomposition and the closure properties shown in Section 3.1.

According to Lemma 3.2.5, when $(a, b, c, d) \in \mathcal{Z}$, the conductance of this $\mathcal{MC}$ is polynomially bounded if $\frac{Z(\Omega_2)}{Z(\Omega_0)}$ is polynomially bounded. According to Corollary 3.2.3, when $(a, b, c, d) \in \mathcal{Y}$, $\frac{Z(\Omega_2)}{Z(\Omega_0)}$ is polynomially bounded, which proves part (2) above. Combining Lemma 3.2.5 and Corollary 3.2.3, we can also conclude part (1). As a consequence of (1) and (2), we are able to efficiently sample valid eight-vertex configurations according to the Gibbs measure on Ω_0 (almost uniformly), and in the following algorithm we only work with states in Ω_0 , the set of even orientations.

Before we state the algorithm, we need to extend the type of vertices a graph can have in the eight-vertex model. Previously, a graph can only have degree 4 vertices, on each of which a constraint function satisfies the even orientation rule and 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 3.2.5 and Corollary 3.2.3 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 compute the partition function $Z_{8V}(G)$ via sampling with the directed-loop algorithm $\mathcal{MC}$. As we have argued in Section 3.1, the partition function of the eight-vertex models can be viewed as the weighted sum over a 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 3.1.7 for each vertex we can choose a nonnegative weight function w on signed pairings at v . For a vertex $v \in V$, the ratios among different signed pairings $\{\curvearrowright, \curvearrowleft, \dashv\} \times \{+, -\}$ in weighted *dacp*’s can be uniquely determined by the ratios among different orientations (represented by $a, b,$

c , and d) at v . For example, if we express $Z_{8V}(G)$ as $2aA + 2bB + 2cC + 2dD$ according to the local orientation configuration at v , as in the proof of Corollary 3.1.8, we see that indeed $2w(\swarrow_+)(A + D)$ is the weight for finding the signed pairing $\swarrow_+$ at v . As long as the partition function is not zero (this can be easily tested in polynomial time), there is a signed pairing ρ showing up at v with probability at least $\frac{1}{6}$ among all six signed pairings. Moreover, according to Corollary 3.1.8, one of the pairings in $\{\swarrow_+, \nearrow_+, \dashv_+\}$ shows up at v with probability at least $\frac{1}{6}$. Therefore, running $\mathcal{MC}$ on G , we can approximate, with a sufficient $1/\text{poly}(n)$ precision, the probability of having $\rho \in \{\swarrow_+, \nearrow_+, \dashv_+\}$ at v , denoted by $\Pr_v(\rho)$. Denote by $G_{v,\rho}$ the graph with v being split into v_1 and v_2 and the edges reconnected according to ρ . Recall that the degree 2 vertices v_1 and v_2 must satisfy the “1-in-1-out” rule in any valid configuration. Write the partition function of $G_{v,\rho}$ as $Z_{8V}(G_{v,\rho})$, we have $\Pr_v(\rho) = w(\rho)Z_{8V}(G_{v,\rho})/Z_{8V}(G)$ which means $Z_{8V}(G) = w(\rho)Z_{8V}(G_{v,\rho})/\Pr_v(\rho)$. To approximate $Z_{8V}(G)$ it suffices to approximate $Z_{8V}(G_{v,\rho})$, which can be done by running $\mathcal{MC}$ on $G_{v,\rho}$ and recursing. Repeating this process for $|V|$ steps we decompose the graph G into the base case, a set of disjoint cycles. The partition function of this cycle graph is just 2^C where C is the number of cycles. By this self-reduction, the partition function $Z_{8V}(G)$ can be approximated. $\square$

Proof of Theorem 3.2.2. The proof is similar to that of Theorem 3.2.1, with the help of Corollary 3.2.4 and Corollary 3.1.9, two corollaries of the closure property Theorem 3.1.3 which holds on planar graphs.

Given a plane graph G with a constraint function on every vertex from $\mathcal{Z} \cap \mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}}$, we can still efficiently sample even orientations according to the Gibbs measure. However, in order to do self-reduction, we have to prove something more.

To make our algorithm work, we need to extend the type of vertices in the eight-vertex model again. Previously in the proof of Theorem 3.2.1, a graph can have degree 4 vertices, on each of which the constraint function satisfies the even orientation rule and arrow reversal symmetry, and 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. Now, a graph can also have degree 2 vertices, on each of which the constraint function satisfies the “2-in/2-out” rule and both valid local configurations

have weight 1. One can check that Lemma 3.2.5 still holds even with this extension.

The self-reduction still processes one vertex v at a time. As long as the partition function is not zero, there is a signed pairing ρ showing up at v with probability at least $\frac{1}{6}$ among all six signed pairings. Moreover, according to Corollary 3.1.9, one of the pairings $\{\nearrow_+, \searrow_+, \dashv\vdash_-\}$ shows up at v with probability at least $\frac{1}{6}$. If ρ is $\nearrow_+$ or $\searrow_+$, let $G_{v,\rho}$ be the graph with v being split into v_1 and v_2 and the edges reconnected according to ρ . The degree 2 vertices v_1 and v_2 must satisfy the “1-in-1-out” rule in any valid configuration, just as in the proof of Theorem 3.2.1.

If ρ is $\dashv\vdash_-$, let $G_{v,\dashv\vdash_-}$ be the graph with v being split into v_1 and v_2 and the edges reconnected according to $\dashv\vdash_-$. This time, the degree 2 vertices v_1 and v_2 must satisfy the “2-in/2-out” rule in any valid configuration. Observe that Theorem 3.1.3 holds for $G_{v,\dashv\vdash_-}$ if and only if it holds for $G'_{v,\dashv\vdash_-}$, which is obtained from G by replacing v by a virtual vertex v' with parameter setting $(a, b, c, d) = (0, 0, 1, 1)$ (this is equivalent to choosing $w(\dashv\vdash_-) = 1$ and w being 0 on the other five signed pairings, for a nonnegative w at v'). Since $(0, 0, 1, 1) \in \mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}} \cap \overline{\mathcal{F}_>}$, Theorem 3.1.3 and consequently Corollary 3.2.4 still hold for $G'_{v,\dashv\vdash_-}$ thus also for $G_{v,\dashv\vdash_-}$. (Note that this $G'_{v,\dashv\vdash_-}$ is not involved algorithmically in subsequent steps; its only purpose is to show that Theorem 3.1.3 holds for $G_{v,\dashv\vdash_-}$, on which the algorithm continues.)

The subsequent steps in the self-reduction step for v are the same as in the proof of Theorem 3.2.1. The base case is a decomposition of G into a set of disjoint cycles with an even number of degree 2 vertices that satisfy the “2-in/2-out” rule. This is proved by using the *Jordan Curve Theorem*: The graph is initially planar. Any step replacing v with v_1 and v_2 for $\nearrow_+$ or $\searrow_+$ in $G_{v,\rho}$ does not create any non-planar crossings nor vertices satisfying the “2-in/2-out” rule. Only the third type of steps replacing v with v_1 and v_2 for $\dashv\vdash_-$ in $G_{v,\dashv\vdash_-}$ create a non-planar crossing and also a vertex satisfying the “2-in/2-out” rule at each crossing locally at each branch of the crossing. Thus at the end we are left with a set of disjoint cycles where along each cycle degree 2 vertices satisfying the “2-in/2-out” rule are in 1-1 correspondence with non-planar crossings. By the *Jordan Curve Theorem* this number is even, for every cycle. The partition function of this cycle graph is just 2^C where C is the number of cycles. Again, the partition function $Z_{8V}(G)$ can

be approximated. □

Corollary 3.2.3. *Given a 4-regular graph $G = (V, E)$, if the constraint function on every vertex is from $\mathcal{Y}$, then $\frac{Z(\Omega_2)}{Z(\Omega_0)} \leq \binom{|E|}{2}$.*

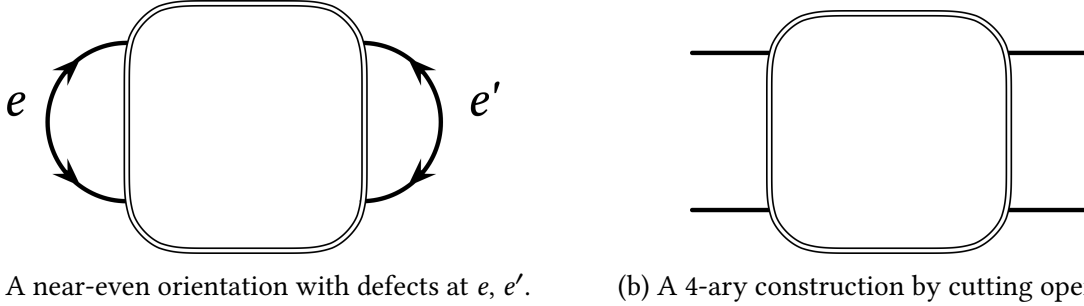


Figure 3.6: The correspondence between configurations before and after cutting open two edges.

Proof. For each near-even orientation, there are exactly two defective edges. Let $\Omega_2^{\{e, e'\}} \subseteq \Omega_2$ be the set of near-even orientations in which e, e' are these 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 coming in or going out, with 4 possibilities. An example is in Figure 3.6a where both e and e' have their half-edges going out. If we “cut open” e and e' as shown in Figure 3.6b, we get a 4-ary construction Γ using degree 4 vertices with constraint functions in $\mathcal{Y}$. 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 Γ we observe that: the set of near-even orientations in $\Omega_2^{\{e, e'\}}$ contributes a total weight $(a' + a' + d' + d')$, i.e. $Z(\Omega_2^{\{e, e'\}}) = 2(a' + d')$; the set of even orientations in Ω_0 has a total weight $Z(\Omega_0) = 2(b' + c')$. By Theorem 3.1.2 we know that for the 4-ary construction Γ , $a' + d' \leq b' + c'$. Therefore, $\frac{Z(\Omega_2^{\{e, e'\}})}{Z(\Omega_0)} \leq 1$. In total, $\frac{Z(\Omega_2)}{Z(\Omega_0)} \leq \binom{|E|}{2}$. □

Corollary 3.2.4. *Given a 4-regular plane graph $G = (V, E)$, if the constraint function on every vertex is from $AD \cap BD \cap \overline{CD} \cap \overline{F}_>$, then $\frac{Z(\Omega_2)}{Z(\Omega_0)} \leq \binom{|E|}{2}$.*

Proof. For any 4-regular plane graph $G = (V, E)$, if we cut the two defective edges of $\tau \in \Omega_2$, we obtain a planar Γ with 4 dangling edges using constraint functions from $AD \cap BD \cap \overline{CD} \cap \overline{F}_>$.

We name e_1 and e_2 the two dangling edges cut from one edge in G , and e_3 and e_4 cut from the other. Both e_1 and e_2 now reside in a single face of Γ , and so do e_3 and e_4 . We can modify the proof of Theorem 3.1.3 to establish that for Γ , we still have $a' + d' \leq b' + c'$. $\square$

Although $\mathcal{MC}$ runs on the even orientations and near-even orientations of a 4-regular graph G , it is formally defined and analyzed using the edge-vertex incidence graph G' of G introduced in Section 1.3.

Let $G' = (V, U, E)$ be the edge-vertex incidence graph of G , an instance of $Z_{8V}(a, b, c, d)$. Each vertex in V is assigned $(\neq_2)$; each vertex $u \in U$ is assigned a constraint function $f_u \in \mathcal{Z}$. An assignment σ assigns a value in $\{0, 1\}$ to each edge $e \in E$. The state space of $\mathcal{MC}$ is $\Omega = \Omega_0 \cup \Omega_2$, which consists of “perfect” or “near-perfect” assignments to E , defined as follows: all assignments satisfy the “two-0-two-1/four-0/four-1” rule at every vertex $u \in U$ of degree 4; all assignments satisfy the “one-0-one-1” at every $v \in V$ with *possibly exactly two* exceptions. Assignments in Ω_0 have no exceptions, and are “perfect” (corresponding to the even orientations in G). Assignments in Ω_2 have exactly two exceptions, and are “near-perfect” (corresponding to the near-even orientations in G). Thus any $\sigma \in \Omega_0$ satisfies all $(\neq_2)$ on V , and any $\sigma \in \Omega_2$ satisfies all $(\neq_2)$ on $V - \{v', v''\}$ for some two vertices $v', v'' \in V$ where it satisfies $(=_2)$ (which outputs 1 on inputs 00, 11 and outputs 0 on 01, 10). For any assignment $\sigma \in \Omega$ and any subset $S \subseteq \Omega$, define the *weight* function $\mathcal{W}$ by $\mathcal{W}(\sigma) = \prod_{u \in U} f_u(\sigma|_{E(u)})$ and $Z(S) = \sum_{\sigma \in S} \mathcal{W}(\sigma)$. Then the *Gibbs measure* for Ω is defined by $\pi(\sigma) = \frac{\mathcal{W}(\sigma)}{Z(\Omega)}$, assuming $Z(\Omega) > 0$.

Transitions in $\mathcal{MC}$ are comprised of three types of moves. Suppose $\sigma \in \Omega_0$. An Ω_0 -to- Ω_2 move from σ takes a degree 4 vertex $u \in U$ and two incident edges $e' = (v', u), e'' = (v'', u) \in V \times U$, and changes it to $\sigma_2 \in \Omega_2$ which flips both $\sigma(e')$ and $\sigma(e'')$. The effect is that at v' and v'' , σ_2 satisfies $(=_2)$ instead of $(\neq_2)$. An Ω_2 -to- Ω_0 move is the opposite. An Ω_2 -to- Ω_2 move is, intuitively, to *shift* one $(=_2)$ from one vertex $v' \in V$ to another $v^* \in V$, where for some $u \in U$, v' and v^* are both incident to u and the “two-0-two-1/four-0/four-1” rule at u is preserved. Formally, let $\sigma \in \Omega_2$ be a near-perfect assignment with $v', v'' \in V$ being the two exceptional vertices (i.e., σ satisfies $(=_2)$ at v' and v''). Let $v^* \in V - \{v', v''\}$ be such that for some $u \in U$, both $e' = (v', u), e^* = (v^*, u) \in E$.

Then an Ω_2 -to- Ω_2 move changes σ to σ^* by flipping both $\sigma(e')$ and $\sigma(e^*)$. The effect is that σ^* satisfies $(\neq_2)$ at v' and $(=_2)$ at v^* . Note that σ^* continues to satisfy $(=_2)$ at v'' .

The above describes a symmetric binary relation *neighbor* ($\sim$) on Ω . No two states in Ω_0 are neighbors. Set $n = |U|$. The number of neighbors of a Ω_0 -state is at most $6n$ (by first picking a vertex and then picking a pair of edges incident to this vertex) and the number of neighbors of a Ω_2 -state is at most a constant. The transition probabilities $P(\cdot, \cdot)$ of $\mathcal{MC}$ are *Metropolis* moves between neighboring states:

$$P(\sigma_1, \sigma_2) = \begin{cases} \frac{1}{12n} \min\left(1, \frac{\pi(\sigma_2)}{\pi(\sigma_1)}\right) & \text{if } \sigma_2 \sim \sigma_1; \\ 1 - \frac{1}{12n} \sum_{\sigma' \sim \sigma_1} \min\left(1, \frac{\pi(\sigma')}{\pi(\sigma_1)}\right) & \text{if } \sigma_1 = \sigma_2; \\ 0 & \text{otherwise.} \end{cases}$$

$\mathcal{MC}$ is aperiodic due to the “lazy” movement; one can verify that $\mathcal{MC}$ is irreducible by creating, shifting, and merging two $(=_2)$ ’s; as the transitions are Metropolis moves, detailed balance conditions are satisfied with regard to π . By results from [JS89, Sin92], such a Markov chain is rapidly mixing if there is a *flow* whose congestion can be bounded by a polynomial in n .

Lemma 3.2.5. *Assume $Z(\Omega_0) > 0$. Given $f_u \in \mathcal{Z}$ for every vertex $u \in U$, there is a flow on Ω with congestion at most $O\left(n^2 \left(\frac{Z(\Omega)}{Z(\Omega_0)}\right)^2\right)$, using paths of length $O(n)$.*

Proof. The idea is to design a flow $\mathfrak{F} : \mathcal{P} \rightarrow \mathbb{R}^+$ from Ω_2 to Ω_0 which satisfies

$$\sum_{p \in \mathcal{P}_{\sigma_2 \sigma_0}} \mathfrak{F}(p) = \pi(\sigma_2)\pi(\sigma_0), \quad \text{for all } \sigma_2 \in \Omega_2, \sigma_0 \in \Omega_0,$$

where $\mathcal{P}_{\sigma_2 \sigma_0}$ is defined to be a set of simple directed paths from σ_2 to σ_0 in $\mathcal{MC}$ and $\mathcal{P} = \bigcup_{\sigma_2 \in \Omega_2, \sigma_0 \in \Omega_0} \mathcal{P}_{\sigma_2 \sigma_0}$. Once the congestion of $\mathfrak{F}$ from Ω_2 to Ω_0 is polynomially bounded, so is the flow from Ω_0 to Ω_2 by symmetric construction. Moreover, there is a flow from Ω_2 to Ω_2 (or from Ω_0 to Ω_0) whose congestion can also be polynomially bounded by randomly picking an intermediate state in Ω_0 (or Ω_2 , respectively). Thus we have a flow on Ω with polynomially bounded congestion.

This technique has been used in [JSV04, McQ13]. In the following we show that the congestion of $\mathfrak{F}$ from Ω_2 to Ω_0 is bounded by $O(n^2 \frac{Z(\Omega_2)}{Z(\Omega_0)})$. Then the bound in the lemma for a flow on Ω follows.

To describe the flow $\mathfrak{F}$, we first specify the sets of paths that are going to take the flow. In line with the definition of Ω_0 and Ω_2 , we define Ω_4 to be the set of assignments where there are exactly four violations of $(\neq_2)$ in V . Let $\Omega' = \Omega_0 \cup \Omega_2 \cup \Omega_4$. For $\sigma, \sigma' \in \Omega'$, let $\sigma \oplus \sigma'$ denote the *symmetric difference* (or bitwise XOR), where we view σ and σ' as two bit strings in $\{0, 1\}^{|E|}$. This is a 0-1 assignment to the edge set of the edge-vertex incidence graph $G' = (V, U, E)$ of G . We also treat $\sigma \oplus \sigma'$ as an edge subset of E (corresponding to bit positions having bit 1, where σ and σ' assign opposite values), and this defines an edge-induced subgraph of G' , which we will just call it $\sigma \oplus \sigma'$. Since at every $u \in U$ of degree 4, the “two-0-two-1/four-0/four-1” rule is satisfied by both σ and σ' , this edge-induced subgraph has even degree (0, 2, or 4) at every $u \in U$.

Let us introduce the set of *atcp*’s (annotated trail & circuit partitions) for the symmetric difference $\sigma \oplus \sigma'$. It is similar to the notions of *acp* for 4-regular graphs and *atcp* for 4-ary constructions defined in Section 3.1. Let us assume $\sigma \in \Omega_0$ and $\sigma' \in \Omega_2$, and the set of *atcp*’s for $\sigma \oplus \sigma'$ in general cases when $\sigma, \sigma' \in \Omega'$ can be similarly defined. If $\sigma \in \Omega_0$ and $\sigma' \in \Omega_2$, on the edge where σ' is defective (but $\sigma \in \Omega_0$ is not), $\sigma \oplus \sigma'$ has a degree 1 vertex. First we assign a pairing (that groups four incident edges into two unordered pairs) at every vertex of degree 4 in $\sigma \oplus \sigma'$. This partitions the edges of $\sigma \oplus \sigma'$ into a set of edge-disjoint circuits and exactly one *trail* which ends in the two vertices in V of degree 1. Then we affix a $\pm$ at every vertex $u \in U$ of degree 2 or degree 4 in $\sigma \oplus \sigma'$ as follows: If $u \in U$ has degree 4 in $\sigma \oplus \sigma'$ then σ and σ' represent total reversal orientations of each other at u , and thus the pairing at u has the same sign according to Table 3.1 for σ and σ' . We affix this sign at u . If $u \in U$ has degree 2 in $\sigma \oplus \sigma'$, then σ and σ' disagree on exactly two edges. On these two edges, if one assigns 01 the other assigns 10 (and vice versa), and if one assigns 00 the other assigns 11 (and vice versa). We affix $+$ at u in the first case, and $-$ in the second case. One can check that for any *atcp* φ of $\sigma \oplus \sigma'$, one encounters an even number of $-$ along any circuit of φ .

Denote by $U_4 \subseteq U$ the degree-4 vertices in $\sigma \oplus \sigma'$. Then there are exactly $3^{|U_4|}$ *atcp*’s for $\sigma \oplus \sigma'$.

Note that an *atcp* of $\sigma \oplus \sigma'$ is uniquely determined by a family of signed pairings on U_4 . This is a 1-1 correspondence and we will identify the two sets. For any signed pairing in $\{\curvearrowright, \curvearrowleft, \dashv, \vdash\} \times \{+, -\}$ on a vertex u with constraint matrix $M(f_u) = \begin{bmatrix} d & b & c & a \\ a & b & c & d \\ c & b & a & d \\ d & a & c & b \end{bmatrix}$, define the weight function $\mathfrak{w}$ for signed pairings as follows, $\begin{cases} a^2 = \mathfrak{w}(\curvearrowright_+) + \mathfrak{w}(\curvearrowleft_+) + \mathfrak{w}(\vdash_+) \\ b^2 = \mathfrak{w}(\curvearrowright_+) + \mathfrak{w}(\curvearrowleft_-) + \mathfrak{w}(\vdash_+) \\ c^2 = \mathfrak{w}(\curvearrowright_+) + \mathfrak{w}(\curvearrowleft_+) + \mathfrak{w}(\vdash_-) \\ d^2 = \mathfrak{w}(\curvearrowright_-) + \mathfrak{w}(\curvearrowleft_-) + \mathfrak{w}(\vdash_-) \end{cases}$. Note that $\mathfrak{w}$ has a *nonnegative* solution if and only if $f_u \in \mathcal{Z}$ by a proof similar to that of Lemma 3.1.7. Let $\Phi_{\sigma \oplus \sigma'}$ be the set of *atcp*'s for $\sigma \oplus \sigma'$. For $\varphi \in \Phi_{\sigma \oplus \sigma'}$, define

$$\mathfrak{W}(\sigma, \sigma', \varphi) := \left(\prod_{u \in U \setminus U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \right) \left(\prod_{u \in U_4} \mathfrak{w}(\varphi(u)) \right),$$

where $\varphi(u)$ is the signed pairing given by φ at u . Then for all distinct $\sigma, \sigma' \in \Omega'$, we have

$$\begin{aligned} \sum_{\varphi \in \Phi_{\sigma \oplus \sigma'}} \mathfrak{W}(\sigma, \sigma', \varphi) &= \sum_{\varphi \in \Phi_{\sigma \oplus \sigma'}} \left(\prod_{u \in U \setminus U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \right) \left(\prod_{u \in U_4} \mathfrak{w}(\varphi(u)) \right) \\ &= \left(\prod_{u \in U \setminus U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \right) \left(\sum_{\varphi \in \Phi_{\sigma \oplus \sigma'}} \prod_{u \in U_4} \mathfrak{w}(\varphi(u)) \right) \\ &= \left(\prod_{u \in U \setminus U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \right) \left(\prod_{u \in U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \right) \\ &= \prod_{u \in U} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) \\ &= \mathcal{W}(\sigma) \mathcal{W}(\sigma'). \end{aligned}$$

The equality from line 2 to line 3 is due to the following: when the degree (in the induced subgraph $\sigma \oplus \sigma'$) of a vertex $u \in U$ is 4, σ and σ' must take the same value at u , since one represents a total reversal of all arrows of another; thus $f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)})$ is in $\{a^2, b^2, c^2, d^2\}$. Then

$$\prod_{u \in U_4} f_u(\sigma|_{E(u)}) f_u(\sigma'|_{E(u)}) = \sum_{\varphi \in \Phi_{\sigma \oplus \sigma'}} \prod_{u \in U_4} \mathfrak{w}(\varphi(u))$$

is obtained by using the sum expressions for a^2 , b^2 , c^2 , and d^2 in terms of $\mathfrak{w}(\curvearrowright_+)$, $\mathfrak{w}(\curvearrowright_-)$, $\mathfrak{w}(\curvearrowleft_+)$, $\mathfrak{w}(\curvearrowleft_-)$, $\mathfrak{w}(\vdash_+)$, and $\mathfrak{w}(\vdash_-)$, and then expressing the product-of-sums as a sum-of-products.

Now we are ready to specify the “paths” which take nonzero flow from $\sigma_2 \in \Omega_2$ to $\sigma_0 \in \Omega_0$.

In order to transit from σ_2 to σ_0 , paths in $\mathcal{P}_{\sigma_2\sigma_0}$ go through states in Ω that gradually decrease the number of conflicting assignments along trails and circuits in $\sigma_2 \oplus \sigma_0$. We first specify a total order on E , the set of edges of G' . This induces a total order on circuits by lexicographic order. In the induced subgraph $\sigma_2 \oplus \sigma_0$, exactly two vertices in V have degree 1 (called *endpoints*) and all other vertices have degree 2 or degree 4. The set of paths in $\mathcal{P}_{\sigma_2\sigma_0}$ are designed to be in 1-to-1 correspondence with elements in $\Phi_{\sigma_2\oplus\sigma_0}$. Given any family of signed pairings $\varphi \in \Phi_{\sigma_2\oplus\sigma_0}$, we have a unique decomposition of the induced subgraph $\sigma_2 \oplus \sigma_0$ as an edge disjoint union of one trail $[e_1](v_1, e'_1, u_1, e_2, v_2, e'_2, u_2, \dots, e_k, v_k)[e'_k]$ (where e_1 and e'_k are not part of the trail), and zero or more edge disjoint circuits, which are ordered lexicographically. Here $v_i \in V$ and $u_i \in U$, and the two exceptional vertices are v_1 and v_k where σ_2 satisfies $(=_2)$. The unique path p_φ first reverses all arrows along the trail, starting from the smaller of e'_1 and e_k . If we assume, without loss of generality, e'_1 is the smaller one, then p_φ “pushes” the $(=_2)$ from v_1 , to v_2 , then to $v_3, \dots, v_{k-1}$, and then “merge” at v_k , arriving at a configuration in Ω_0 . Next p_φ reverses all arrows on each circuit in lexicographic order, and within each circuit C it starts at the least edge e (according to the edge order) and reverses all arrows on C in a cyclic order starting in the direction indicated by σ_2 on e . (Technically it flips a pair of incident edges to vertices in U in each step.) Such paths p_φ are well-defined and are valid paths in $\mathcal{MC}$ since along any path every state is in $\Omega = \Omega_0 \cup \Omega_2$ and every move is a valid transition defined in $\mathcal{MC}$. With regard to the flow distribution, the flow value put on p_φ is $\frac{\mathfrak{W}(\sigma_2, \sigma_0, \varphi)}{(Z(\Omega))^2}$, making the following hold for all $\sigma_2 \in \Omega_2, \sigma_0 \in \Omega_0$:

$$\begin{aligned} \sum_{p_\varphi \in \mathcal{P}_{\sigma_2\sigma_0}} \mathfrak{F}(p_\varphi) &= \sum_{\varphi \in \Phi_{\sigma_2\oplus\sigma_0}} \frac{\mathfrak{W}(\sigma_2, \sigma_0, \varphi)}{(Z(\Omega))^2} \\ &= \frac{\mathcal{W}(\sigma_2)\mathcal{W}(\sigma_0)}{(Z(\Omega))^2} \\ &= \pi(\sigma_2)\pi(\sigma_0). \end{aligned}$$

Note that in each path, no edge is flipped more than once, so the length is $O(n)$. For any transition (σ', σ'') where $\sigma' \neq \sigma''$, we have $P(\sigma', \sigma'') = \frac{1}{12n} \min\left(1, \frac{\pi(\sigma'')}{\pi(\sigma')}\right) = \Omega\left(\frac{1}{n}\right)$, as $\frac{\pi(\sigma'')}{\pi(\sigma')}$ is a constant. (This is a constant because we have restricted the constraint function f_u to be from a

fixed finite set $\mathcal{F}$.) Let $H_{\sigma'} = \{\sigma_2 \oplus \sigma_0 \mid \sigma_2 \in \Omega_2, \sigma_0 \in \Omega_0, \exists \varphi \in \Phi_{\sigma_2 \oplus \sigma_0} \text{ s.t. } \sigma' \in p_\varphi\}$. The congestion of $\mathfrak{F}$ is

$$\begin{aligned}
& \max_{\text{transition } (\sigma', \sigma'')} \frac{1}{\pi(\sigma')P(\sigma', \sigma'')} \sum_{\substack{\sigma_2 \in \Omega_2 \\ \sigma_0 \in \Omega_0}} \sum_{\substack{p_\varphi \in \mathcal{P}_{\sigma_2 \oplus \sigma_0} \\ p_\varphi \ni (\sigma', \sigma'')}} \frac{\mathfrak{W}(\sigma_2, \sigma_0, \varphi)}{(Z(\Omega))^2} \\
& \leq \max_{\sigma' \in \Omega} \frac{O(n)}{\mathcal{W}(\sigma')Z(\Omega)} \sum_{\substack{\sigma_2 \in \Omega_2 \\ \sigma_0 \in \Omega_0}} \sum_{\substack{\varphi \in \Phi_{\sigma_2 \oplus \sigma_0} \\ p_\varphi \ni \sigma'}} \mathfrak{W}(\sigma_2, \sigma_0, \varphi) \\
& \leq \max_{\sigma' \in \Omega} \frac{O(n)}{\mathcal{W}(\sigma')Z(\Omega)} \sum_{\sigma_2 \in \Omega_2} \sum_{\eta \in H_{\sigma'}} \sum_{\varphi \in \Phi_\eta} \mathfrak{W}(\sigma_2, \sigma_2 \oplus \eta, \varphi) \\
& = \max_{\sigma' \in \Omega} \frac{O(n)}{\mathcal{W}(\sigma')Z(\Omega)} \sum_{\eta \in H_{\sigma'}} \sum_{\varphi \in \Phi_\eta} \sum_{\sigma_2 \in \tilde{\Omega}_2} \mathfrak{W}(\sigma_2, \sigma_2 \oplus \eta, \varphi).
\end{aligned}$$

On the last line above we exchange the order of summations where $\tilde{\Omega}_2$ is the subset of Ω_2 -states of the form $\sigma_2 = \eta \oplus \sigma_0$, for some $\sigma_0 \in \Omega_0$ such that p_φ (which passes through σ') goes from σ_2 to σ_0 . These are Ω_2 -states “compatible” with the symmetric difference η and its *atcp* φ . The number of states in $\tilde{\Omega}_2$ is bounded by the length of the longest path $O(n)$ because σ' is an intermediate state on a path. Fix any $\sigma' \in \Omega$. For any $\sigma_2 \in \Omega_2$, and $\eta \in H_{\sigma'}$ consisting of exactly one connected component with two endpoints of degree 1 and all other vertices having even degree (and zero or more connected components of even degree vertices), observe that $\sigma' \oplus \eta \in \Omega'$. Indeed, if $\sigma' \in \Omega_0$ then $\sigma' \oplus \eta \in \Omega_2$; if $\sigma' \in \Omega_2$ then depending on whether σ'

(1) is σ_2 , or

(2) appears in the process of reversing arrows on the trail with two endpoints, or

(3) appears after reversing arrows on the trail with endpoints,

$\sigma' \oplus \eta$ lies in Ω_0 , Ω_2 , or Ω_4 , respectively. For the edges not in η , σ' agrees with σ_2 and $\sigma_2 \oplus \eta$ as the path p_φ never “touches” them, and so does $\sigma' \oplus \eta$. Recall that

$$\mathfrak{W}(\sigma_2, \sigma_2 \oplus \eta, \varphi) = \left(\prod_{u \in U \setminus U_4} f_u(\sigma_2|_{E(u)}) f_u((\sigma_2 \oplus \eta)|_{E(u)}) \right) \left(\prod_{u \in U_4} \mathfrak{w}(\varphi(u)) \right).$$

For every vertex $u \in U$ that is not in η , f_u takes the same value in all σ_2 , $\sigma_2 \oplus \eta$, σ' , and $\sigma' \oplus \eta$. For every vertex $u \in U$ that is degree-2 in η , assuming $M(f_u) = \begin{bmatrix} d & b & c & a \\ a & c & b & d \end{bmatrix}$, $f_u(\sigma_2|_{E(u)})$ and $f_u((\sigma_2 \oplus \eta)|_{E(u)})$ take two different elements in $\{a, b, c, d\}$. Meanwhile, $f_u(\sigma'|_{E(u)})$ and $f_u(\sigma' \oplus \eta|_{E(u)})$ also take these two elements (possibly in the opposite order). For example, at the vertex u shown in Figure 3.7, $f_u(\sigma_2|_{E(u)}) = a$ and $f_u((\sigma_2 \oplus \eta)|_{E(u)}) = c$. The two solid edges are in η and assignments on the two dotted edges are shared by σ_2 and $\sigma_2 \oplus \eta$, as well as σ' and $\sigma' \oplus \eta$. On the path p_φ from σ_2 to $\sigma_2 \oplus \eta$ decided by φ : if σ' appears before reversing the two solid edges, then σ' agrees with σ_2 on them ($f_u(\sigma'|_{E(u)}) = a$) and $\sigma' \oplus \eta$ agrees with $\sigma_2 \oplus \eta$ on them ($f_u(\sigma' \oplus \eta|_{E(u)}) = c$); if σ' appears after reversing the two solid edges, then σ' agrees with $\sigma_2 \oplus \eta$ on them ($f_u(\sigma'|_{E(u)}) = c$) and $\sigma' \oplus \eta$ agrees with σ_2 on them ($f_u(\sigma' \oplus \eta|_{E(u)}) = a$). For every vertex $u \in U$ that is degree-4 in η , $\mathfrak{w}(\varphi(u))$ takes the same value in $\mathfrak{W}(\sigma_2, \sigma_2 \oplus \eta, \varphi)$ and $\mathfrak{W}(\sigma', \sigma' \oplus \eta, \varphi)$ as the weight only depends on $\varphi(u)$, the signed pairing at u .



Figure 3.7: A vertex of degree-2 in η .

By the above argument, we established that $\mathfrak{W}(\sigma_2, \sigma_2 \oplus \eta, \varphi) = \mathfrak{W}(\sigma', \sigma' \oplus \eta, \varphi)$. Therefore, the

congestion of $\mathfrak{F}$ can be bounded by

$$\begin{aligned}
& \max_{\sigma' \in \Omega} \frac{O(n)}{\mathcal{W}(\sigma')Z(\Omega)} \sum_{\eta \in H_{\sigma'}} \sum_{\varphi \in \Phi_{\eta}} \sum_{\sigma_2 \in \Omega_2} \mathfrak{W}(\sigma', \sigma' \oplus \eta, \varphi) \\
& \leq \max_{\sigma' \in \Omega} \frac{O(n^2)}{\mathcal{W}(\sigma')Z(\Omega)} \sum_{\eta \in H_{\sigma'}} \sum_{\varphi \in \Phi_{\eta}} \mathfrak{W}(\sigma', \sigma' \oplus \eta, \varphi) \\
& \leq \max_{\sigma' \in \Omega} \frac{O(n^2)}{\mathcal{W}(\sigma')Z(\Omega)} \sum_{\eta \in H_{\sigma'}} \mathcal{W}(\sigma')\mathcal{W}(\sigma' \oplus \eta) \\
& = \max_{\sigma' \in \Omega} \frac{O(n^2)}{Z(\Omega)} \sum_{\eta \in H_{\sigma'}} \mathcal{W}(\sigma' \oplus \eta) \\
& \leq \frac{O(n^2)}{Z(\Omega)} \sum_{\sigma \in \Omega'} \mathcal{W}(\sigma) \\
& = O(n^2) \frac{Z(\Omega')}{Z(\Omega)}.
\end{aligned}$$

By a standard argument as in [JS89, MW96, McQ13], $\frac{Z(\Omega_4)}{Z(\Omega_2)} \leq \frac{Z(\Omega_2)}{Z(\Omega_0)}$. Therefore, the congestion is bounded by $O(n^2) \frac{Z(\Omega_2)}{Z(\Omega_0)}$. $\square$

Remark 3.2.2. Lemma 3.2.5 can be alternatively derived using the notion of “windability” [McQ13].

3.3 #PERFECTMATCHINGS-easiness

In this section and Section 3.4, the complexity of approximating the partition function of the eight-vertex model ($\#8V(a, b, c, d)$) is related to that of approximately counting perfect matchings ($\#PM$).

Theorem 3.3.1. *For any $(a, b, c, d) \in \mathcal{Z}$,*

$$\#8V(a, b, c, d) \leq_{AP} \#PM.$$

To prove the $\#PM$ -easiness result, we express the eight-vertex model in the Holant framework (see Section 1.3) 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 3.3.1).

We know that

$$Z_{8V}(G; a, b, c, d) = \text{Holant} \left(G'; \#_2 \mid f \right),$$

where f is the 4-ary signature with $M(f) = \begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix}$. Considering the fact that $(\#_2)$ can be easily realized by a matchgate (a vertex with two dangling edges), Theorem 3.3.1 is a direct consequence of Lemma 3.3.3 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). We note that allowing nonnegative edge-weights does not add more computational power to the unweighted #PM [McQ13, Proposition 5]. The crucial ingredient of our proof is a *geometric* lemma (Lemma 3.3.4) 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 3.3.5). In other words, we give a complete characterization of signatures of arity 4 that can be realized by matchgates with nonnegative real edges. Moreover, the general version of our result also works for the eight-vertex model without the arrow reversal symmetry.

Although the set of signatures that can be realized by planar matchgates with complex edge weights have been completely characterized [CC17], 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. As part of this section, we give a complete characterization of the constraint functions that can be expressed by 4-ary matchgates in Theorem 3.3.2. This solves an important question that has been open for several years [McQ13, BGJ⁺17]. We believe it is of independent interest.

Definition 3.3.1. We use the term a *k*-ary *matchgate* to denote a graph Γ having k “dangling” edges, labelled $i_1, \dots, i_k$. 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 $f(b_1, \dots, b_k)$ for $(b_1, \dots, b_k) \in \{0, 1\}^k$ 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 3.3.1. Contrary to Definition 3.3.1 which does not require planarity, planar matchgates with complex edge weights has been completely characterized [Val02, CC17]. As computing the weighted sum of perfect matchings is in polynomial time over planar graphs by the *FKT algorithm* [TF61, Kas61, Kas67], problems that can be locally expressed by planar matchgates are tractable over planar graphs.

$$\text{Notation. } S_{\leq^2}^E = \{f \mid M(f) = \begin{bmatrix} d_1 & 0 & 0 & a_1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ a_2 & 0 & 0 & d_2 \end{bmatrix} \text{ satisfying } \begin{cases} a_1 a_2 \leq b_1 b_2 + c_1 c_2 + d_1 d_2 \\ b_1 b_2 \leq a_1 a_2 + c_1 c_2 + d_1 d_2 \\ c_1 c_2 \leq a_1 a_2 + b_1 b_2 + d_1 d_2 \\ d_1 d_2 \leq a_1 a_2 + b_1 b_2 + c_1 c_2 \end{cases}, a_1, \dots, d_2 \geq 0.\}, S_{\leq^2}^O = \\ \{f \mid M(f) = \begin{bmatrix} 0 & d_1 & a_1 & 0 \\ b_1 & 0 & 0 & c_1 \\ c_2 & 0 & 0 & b_2 \\ 0 & a_2 & d_2 & 0 \end{bmatrix} \text{ satisfying } \begin{cases} a_1 a_2 \leq b_1 b_2 + c_1 c_2 + d_1 d_2 \\ b_1 b_2 \leq a_1 a_2 + c_1 c_2 + d_1 d_2 \\ c_1 c_2 \leq a_1 a_2 + b_1 b_2 + d_1 d_2 \\ d_1 d_2 \leq a_1 a_2 + b_1 b_2 + c_1 c_2 \end{cases}, a_1, \dots, d_2 \geq 0.\}.$$

Theorem 3.3.2. Denote by $\mathcal{M}$ the set of signatures that can be realized by 4-ary matchgates. Then $\mathcal{M} = S_{\leq^2}^E \cup S_{\leq^2}^O$.

Remark 3.3.2. 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 3.3.2 for the even parity part ($S_{\leq^2}^E$) is a combination of Lemma 3.3.3 and Lemma 3.3.5. The odd parity part can be proved similarly.

Lemma 3.3.3. Suppose $f \in S_{\leq^2}^E$. Then there is a 4-ary matchgate of constant size whose signature is f .

Proof. We first note that if any of the four inequalities in the definition of $S_{\leq^2}$ is an equality, then the remaining three inequalities automatically hold, since the 8 values $a_1, \dots, d_2$ 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 $a_1 a_2 = b_1 b_2 = c_1 c_2 = d_1 d_2 = 0$, 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 $\begin{bmatrix} a_1 a_2 + b_1 b_2 + c_1 c_2 & 0 & 0 & a_1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ a_2 & 0 & 0 & 1 \end{bmatrix}$. Our construction is a weighted K_4 depicted in Figure 3.8a. 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

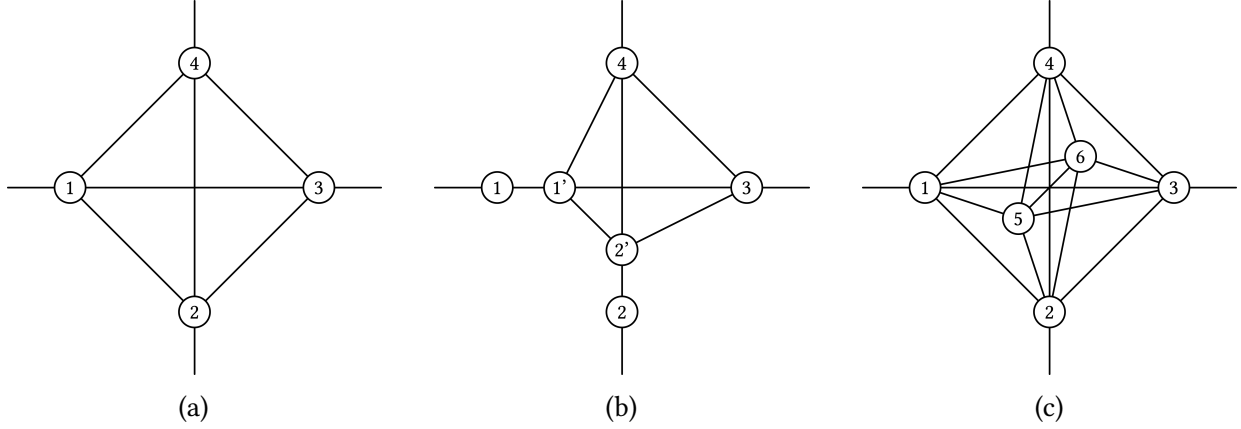


Figure 3.8: 4-ary matchgates.

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 3.8b 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 $\begin{bmatrix} d_1 & 0 & 0 & 1 \\ 0 & b_1 & c_1 & 0 \\ 0 & c_2 & b_2 & 0 \\ b_1 b_2 + c_1 c_2 + d_1 d_2 & 0 & 0 & d_2 \end{bmatrix}$. 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

$$\begin{cases} a_1 a_2 < b_1 b_2 + c_1 c_2 + d_1 d_2 \\ b_1 b_2 < a_1 a_2 + c_1 c_2 + d_1 d_2 \\ c_1 c_2 < a_1 a_2 + b_1 b_2 + d_1 d_2 \\ d_1 d_2 < a_1 a_2 + b_1 b_2 + c_1 c_2 \end{cases} \quad (3.3.1)$$

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 3.8c, 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 signature

with the following parameters

$$\begin{aligned}
a'_1 &= r_1 + p_1 q_2 + p_2 q_1, & a'_2 &= r_2 + p_3 q_4 + p_4 q_3, \\
b'_1 &= s_1 + p_1 q_4 + p_4 q_1, & b'_2 &= s_2 + p_2 q_3 + p_3 q_2, \\
c'_1 &= t_1 + p_1 q_3 + p_3 q_1, & c'_2 &= t_2 + p_2 q_4 + p_4 q_2, \\
d'_1 &= (r_1 r_2 + s_1 s_2 + t_1 t_2) + \\
&\quad (p_3 q_4 + p_4 q_3) r_1 + (p_1 q_2 + p_2 q_1) r_2 + \\
&\quad (p_2 q_3 + p_3 q_2) s_1 + (p_1 q_4 + p_4 q_1) s_2 + \\
&\quad (p_2 q_4 + p_4 q_2) t_1 + (p_1 q_3 + p_3 q_1) t_2. & d'_2 &= 1,
\end{aligned}$$

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 (3.3.1). Our first step is to achieve any relative ratios among the four product values $a_1 a_2, b_1 b_2, c_1 c_2, d_1 d_2$ satisfying (3.3.1); 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 (3.3.1) are as a product $a_1 a_2$, and similarly for $b_1, b_2, c_1, c_2, d_1, d_2$.

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

$$\begin{cases}
A' = p_1 p_2 q_3 q_4 + p_3 p_4 q_1 q_2, & R = r_1 r_2 + r_1(p_3 q_4 + p_4 q_3) + r_2(p_1 q_2 + p_2 q_1), \\
B' = p_1 p_4 q_2 q_3 + p_2 p_3 q_1 q_4, & S = s_1 s_2 + s_1(p_2 q_3 + p_3 q_2) + s_2(p_1 q_4 + p_4 q_1), \\
C' = p_1 p_3 q_2 q_4 + p_2 p_4 q_1 q_3, & T = t_1 t_2 + t_1(p_2 q_4 + p_4 q_2) + t_2(p_1 q_3 + p_3 q_1),
\end{cases} \quad (3.3.2)$$

and define

$$\begin{cases} A &= A' + S + T \\ B &= B' + R + T \\ C &= C' + R + S \\ D &= A' + B' + C'. \end{cases} \quad (3.3.3)$$

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 $r_1, \dots, q_4$ so that A, B, C, D are all positive and satisfy

$$\begin{cases} A &= \frac{1}{2}(b_1 b_2 + c_1 c_2 + d_1 d_2 - a_1 a_2) \\ B &= \frac{1}{2}(a_1 a_2 + c_1 c_2 + d_1 d_2 - b_1 b_2) \\ C &= \frac{1}{2}(a_1 a_2 + b_1 b_2 + d_1 d_2 - c_1 c_2) \\ D &= \frac{1}{2}(a_1 a_2 + b_1 b_2 + c_1 c_2 - d_1 d_2). \end{cases} \quad (3.3.4)$$

Note that, by definition, the left-side of (3.3.4) is precisely the right-side of (3.3.4) when $a_1, \dots, d_2$ are replaced by $a'_1, \dots, d'_2$ respectively. Denote the products $a_1 a_2, b_1 b_2, c_1 c_2, d_1 d_2$ by $a^{**}, b^{**}, c^{**}, d^{**}$ respectively. Then (3.3.4) 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 (3.3.4) 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 (3.3.1) 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

$$\begin{cases} A < B + C + D \\ B < A + C + D \\ C < A + B + D \\ D < A + B + C, \end{cases} \quad (3.3.5)$$

and that A, B, C, D are positive is the same as (3.3.1).

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 (3.3.1). 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 (3.3.5). Thus $(\tilde{A}, \tilde{B}, \tilde{C})$ belongs to the set U in the statement of Lemma 3.3.4.

By Lemma 3.3.4, 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 ($1 \leq i \leq 4$) 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 (3.3.3). 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 ($1 \leq i \leq 4$) 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 (3.3.1) 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 3.8c 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 3.8a to Figure 3.8b). This switches the role of d_1, d_2 with a_2, a_1 and the proof follows. The proofs for $b_1 b_2 = 0$ and $c_1 c_2 = 0$ are symmetric. $\square$

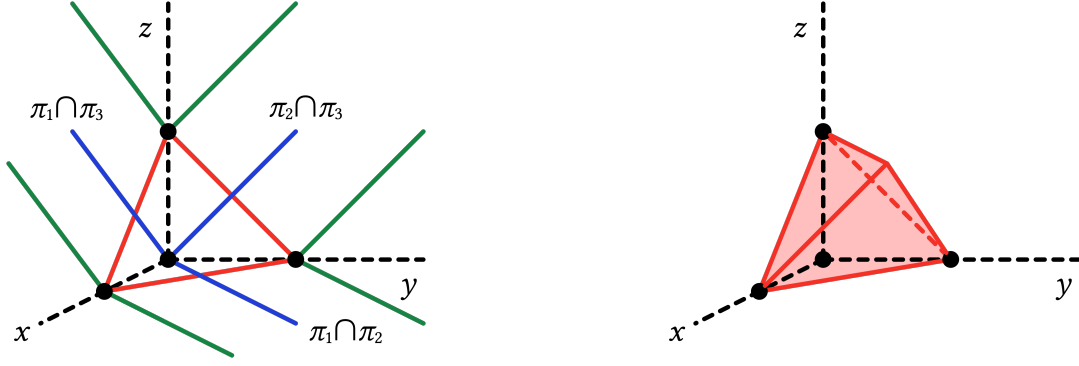
Now we give the crucial geometric lemma.

Lemma 3.3.4. *Let $U = \{(x, y, z) \in \mathbb{R}_{>0}^3 \mid x < y + z + 1, y < x + z + 1, z < x + y + 1, 1 < x + y + z\}$, $V = \{(x, y, z) \in \mathbb{R}_{>0}^3 \mid x + y + z = 1\}$, and $W = \{(x, y, z) \in \mathbb{R}_{>0}^3 \mid y + z > x, x + z > y, x + y > z\}$. Then U is the Minkowski sum of V and W , namely, U consists of precisely those points $\mathbf{u} \in \mathbb{R}^3$, such that $\mathbf{u} = \mathbf{v} + \mathbf{w}$ for some $\mathbf{v} \in V$ and $\mathbf{w} \in W$. 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, $\begin{cases} (\pi_1): y+z \geq x \\ (\pi_2): x+z \geq y \\ (\pi_3): x+y \geq z \end{cases}$, where the planes are defined by equalities. Note that these inequalities imply that $x, y, z \geq 0$, thus this polyhedron has only three facets. We can find the intersection of each pair of the three planes for W as $\begin{cases} \pi_1 \cap \pi_2: x=y \geq 0, z=0 \\ \pi_1 \cap \pi_3: x=z \geq 0, y=0 \\ \pi_2 \cap \pi_3: y=z \geq 0, x=0 \end{cases}$. Note that these intersections lie on the planes $z = 0$, $y = 0$, and $x = 0$, respectively.

Let $\mathbf{e}_1 = (1, 0, 0)$, $\mathbf{e}_2 = (0, 1, 0)$, $\mathbf{e}_3 = (0, 0, 1)$. Then the triangle for V is just the convex hull of $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$. Suppose we shift the origin of W from $\mathbf{0}$ to $\mathbf{e}_1$ and denote the resulting (interior of a) polyhedron by $W^{\mathbf{e}_1}$, then we have the defining inequalities $\begin{cases} (\pi_1^{\mathbf{e}_1}): y+z \geq (x-1) \\ (\pi_2^{\mathbf{e}_1}): (x-1)+z \geq y \\ (\pi_3^{\mathbf{e}_1}): (x-1)+y \geq z \end{cases}$, where the shifted



(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 3.9: A geometric argument.

planes are defined by the corresponding equalities. By symmetry, if we shift the origin of W to e_2 and to e_3 , we have respectively W^{e_2} with $\begin{cases} (\pi_1^{e_2}) : (y-1)+z \geq x \\ (\pi_2^{e_2}) : x+z \geq (y-1) \\ (\pi_3^{e_2}) : x+(y-1) \geq z \end{cases}$, and W^{e_3} with $\begin{cases} (\pi_1^{e_3}) : y+(z-1) \geq x \\ (\pi_2^{e_3}) : x+(z-1) \geq y \\ (\pi_3^{e_3}) : x+y \geq (z-1) \end{cases}$. Note that the shifted planes $\pi_1^{e_1}$, $\pi_2^{e_2}$, and $\pi_3^{e_3}$ contain three distinct facets of U , and they coincide exactly with a facet of W^{e_1} , W^{e_2} , and W^{e_3} , respectively.

By sliding W with its origin along the line $x + y = 1, z = 0$ from e_1 to e_2 , we have a partial coverage of U by the shifting copies of W from W^{e_1} and W^{e_2} :

- The shifted ray of $\pi_1 \cap \pi_2 : x = y \geq 0, z = 0$ moves from $\pi_1^{e_1} \cap \pi_2^{e_1} : (x-1) = y \geq 0, z = 0$ to $\pi_1^{e_2} \cap \pi_2^{e_2} : x = (y-1) \geq 0, z = 0$. 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 $\pi_1^{e_1} \cap \pi_3^{e_1} : (x-1) = z \geq 0, y = 0$ to $\pi_1^{e_2} \cap \pi_3^{e_2} : x = z \geq 0, (y-1) = 0$; the shifted ray of $\pi_2 \cap \pi_3 : y = z \geq 0, x = 0$ moves from $\pi_2^{e_1} \cap \pi_3^{e_1} : y = z \geq 0, (x-1) = 0$ to $\pi_2^{e_2} \cap \pi_3^{e_2} : (y-1) = z \geq 0, x = 0$. Notice that both stay on the plane $x + y - z = 1$ which is $\pi_3^{e_1} = \pi_3^{e_2}$.

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 e_1 to e_2 (which is a side of the triangle V).

Symmetrically, after sliding W with its origin from e_2 to e_3 along the line $y + z = 1, x = 0$ we get the parallel transport from W^{e_2} to W^{e_3} . Also after sliding W with its origin from e_3 back to e_1 along the line segment $x + z = 1, y = 0$ we get the parallel transport from W^{e_3} to W^{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}_{\geq 0}^3 \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}\} - \text{a tetrahedron (Figure 3.9b)}. However this subset can be covered by the rays $\{v + \lambda(1, 1, 1) \mid v \in V, \lambda > 0\}$. 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 $v_n \rightarrow v \in \overline{V}$ and $w_n \rightarrow w \in \overline{W}$, then $u_n = v_n + w_n \in U$, and $u_n \rightarrow v + w$. So $v + w \in \overline{U}$. Conversely, if $u_n \rightarrow u \in \overline{U}$, 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 $v = \lim_{k \rightarrow \infty} v_{n_k} \in \overline{V}$. Then $w_{n_k} = u_{n_k} - v_{n_k}$ also converges to some $w \in \overline{W}$, and then $u = \lim_{k \rightarrow \infty} (v_{n_k} + w_{n_k}) = v + w$, is a sum of points from $\overline{V}$ and $\overline{W}$.

This completes the proof. $\square$

Lemma 3.3.5. 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 $f \in \mathcal{S}_{\leq 2}^E$. In particular, if f satisfies arrow reversal symmetry, $f \in \mathcal{Z}$.

Remark 3.3.3. The last part $d_1 d_2 \leq a_1 a_2 + b_1 b_2 + c_1 c_2$ was proved in [BGJ⁺17, 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 $a_1 a_2 \leq b_1 b_2 + c_1 c_2 + d_1 d_2$.

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

$$f(0011)f(1100) \leq f(0110)f(1001) + f(0101)f(1010) + f(0000)f(1111). \quad (3.3.6)$$

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

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

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 (3.3.6).

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, $\mu : (m_1, m_2) \rightarrow (m_3, m_4)$ is an injection.

To see that μ is weight-preserving, observe that 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,

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

□

3.4 #PERFECTMATCHINGS-hardness

Theorem 3.4.1. *For any four positive numbers $a, b, c, d > 0$ such that $(a, b, c, d) \notin \mathcal{Y}$, the problem $\#8V(a, b, c, d)$ is at least as hard to approximate as counting perfect matchings:*

$$\#PM \leq_{AP} \#8V(a, b, c, d).$$

Remark 3.4.1. 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 Table 3.4. There is a symmetry among a, b, c for the eight-vertex model on general (not necessarily planar) 4-regular graphs, so in this table for simplicity we assume $a \leq b \leq c$.

Table 3.4: Approximation complexity of the eight-vertex model with $(a, b, c, d) \notin \mathcal{Y}$.

	$d = 0$	$d > 0$
$a = b = c = 0$	P-time computable (trivial)	P-time computable (trivial)
$a = b = 0, c > 0$	P-time computable (trivial)	$c = d$: P-time computable [CF17] $c \neq d$: NP-hard (Section 3.5)
$a = 0, b, c > 0$	NP-hard (Section 3.5)	#PM-hard (Section 3.4)
$a, b, c > 0$	NP-hard (Section 2.3)	#PM-hard (Section 3.4)

Corollary 3.4.2. *For any four positive numbers $a, b, c, d > 0$ such that $(a, b, c, d) \in \mathcal{Z} \setminus \mathcal{Y}$,*

$$\#8V(a, b, c, d) \equiv_{AP} \#PM.$$

Note that for the eight-vertex model in the region $\{(a, b, c, d) \mid a + d \leq b + c, \ b + d \leq a + c, \ c + d > a + b\} \cap \mathcal{Z}$, computing $Z_{8V}(a, b, c, d)$ is (1) #P-complete in exact computation [CF17], (2) #PM-equivalent in approximate computation on general 4-regular graphs (Corollary 3.4.2), and (3) admits an FPRAS in approximate computation on planar 4-regular graphs (Section 3.2). To our

best knowledge, these are the first identified problems having these three properties. 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. We note that it is open if approximately computing the partition function in $\mathcal{X} \setminus (\mathcal{Y} \cup \mathcal{Z})$ is #PM-equivalent or not.

Our proof strategy for Theorem 3.4.1 is as follows. In Lemma 3.4.3, 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 (i.e. the even-coloring model) on G , and is achieved using a holographic transformation. In Lemma 3.4.4, we give an approximation-preserving reduction to show that the even-coloring model is at least as hard as a modified version of the even-coloring model where weights at some input originally in the support are dropped off. In Lemma 3.4.5, we establish the equivalence between this modified version of the even-coloring model and the zero-field Ising model. Thus a known result for the Ising model (Proposition 3.4.6) indicates the #PM-equivalence of this modified version of the even-coloring model under certain parameter settings (Corollary 3.4.7). It can be deduced from Lemma 3.4.3, Lemma 3.4.4, and Corollary 3.4.7 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 even-coloring model under approximation-preserving reductions.

Lemma 3.4.3. $2^{|V(G)|} \cdot Z_{8V}(G; a, b, c, d) = \text{Holant} \left(G; \begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & a-b+c-d & a+b-c-d & 0 \\ 0 & a+b-c-d & a-b+c-d & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix} \right).$

The above lemma is essentially the same as Lemma 3.4.3.

Lemma 3.4.4. *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. Figure 3.10 describes a simple gadget construction in the Holant framework called *chaining* which has already been used in Section 2.3. In Figure 3.10 if we place the signature f_1 with $M_{x_i x_j, x_l x_k}(f_1)$ and f_2 with $M_{x_p x_q, x_s x_r}(f_2)$ at the two degree 4 vertices, then the signature f_3 of the 4-ary

construction is $M(f_3) = M_{x_i x_j, x_s x_r}(f_3) = M_{x_i x_j, x_l x_k}(f_1) \cdot M_{x_p x_q, x_s x_r}(f_2)$. Since

$$\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix} = \begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & a-b+c-d & a+b-c-d & 0 \\ 0 & a+b-c-d & a-b+c-d & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

we know that the Holant problem below on the left can be expressed by the Holant problem on the right, and therefore

$$\text{Holant} \left(\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix} \right) \leq_{\text{AP}} \text{Holant} \left(\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & a-b+c-d & a+b-c-d & 0 \\ 0 & a+b-c-d & a-b+c-d & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right).$$

Notice that $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ is the signature matrix of the arity 4 equality function ($=_4$). Therefore, Lemma 3.4.4 is true if we can show

$$\text{Holant} \left(\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & a-b+c-d & a+b-c-d & 0 \\ 0 & a+b-c-d & a-b+c-d & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \right) \leq_{\text{AP}} \text{Holant} \left(\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & a-b+c-d & a+b-c-d & 0 \\ 0 & a+b-c-d & a-b+c-d & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix} \right),$$

which is equivalent to the following in the orientation view according to the proof of Lemma 3.4.3

$$\text{Holant} \left(\neq_2 \mid \begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \right) \leq_{\text{AP}} \text{Holant} \left(\neq_2 \mid \begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix} \right), \quad (3.4.1)$$

by the holographic transformation

$$4(Z^{-1})^{\otimes 4}((=4)) = \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}^{\otimes 4} \left[\begin{bmatrix} 1 \\ 0 \end{bmatrix}^{\otimes 4} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}^{\otimes 4} \right] = \begin{bmatrix} 1 \\ 1 \end{bmatrix}^{\otimes 4} + \begin{bmatrix} -i \\ i \end{bmatrix}^{\otimes 4} = 2[1, 0, 1, 0, 1],$$

a constant multiple of the even parity function, which has the signature matrix $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$.

Next we show how to get (3.4.1). 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 $\#8V(a, b, c, d)$,

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.,

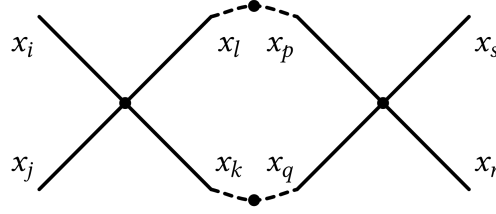


Figure 3.10: 2-chain.

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:

$$0 < d \leq a \leq b \leq c. \quad (3.4.2)$$

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. Following are the details.

Given a signature f with $M(f) = \begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix}$ and $a, b, c \geq 0, d > 0$, we show how to obtain a 4-ary construction with signature $\hat{f}$ such that $M(\hat{f}) = \begin{bmatrix} \hat{d} & 0 & 0 & \hat{a} \\ 0 & \hat{b} & \hat{c} & 0 \\ 0 & \hat{c} & \hat{b} & 0 \\ \hat{a} & 0 & 0 & \hat{d} \end{bmatrix}$ and $\hat{a}, \hat{b}, \hat{c}, \hat{d}$ satisfy the condition $0 < \hat{d} \leq \hat{a} \leq \hat{b} \leq \hat{c} \leq \frac{3}{2}\hat{d}$, using copies of the signature f . (When used in the context of Lemma 3.4.4, the number of copies of f used is a constant.) By a scaling we get

$$1 \leq \hat{d} \leq \hat{a} \leq \hat{b} \leq \hat{c} \leq \frac{3}{2}\hat{d}.$$

We first deal with the case where one of a, b, c is equal to 0. By relabeling we may assume $b = 0$. Then we use the construction G_1 in the proof of Lemma 3.4.4 to get a signature with parameters $(a_1, b_1, c_1, d_1) = (ac, d^2, a^2, cd)$ with no zero entries.

In the following we may assume all entries are positive $a, b, c, d > 0$.

Our first step is to satisfy (3.4.2). If all four entries are equal $a = b = c = d$ then we are done (the signature is also in the affine family and thus the problem is also P-time tractable.) So we may assume at least one of $a, b, c \neq d$. By relabeling we may assume $a \neq d$. We may also assume

at least one of the following three displayed inequalities does not hold, and by relabeling we may assume $a + d < b + c$. Indeed, if all three inequalities

$$a + d \geq b + c, \quad b + d \geq a + c, \quad c + d \geq a + b$$

hold, then by connecting two copies of f using the 4-ary construction in Figure 3.10 and relabeling edges, we get $(a', b', c', d') = (2bc, a^2 + d^2, b^2 + c^2, 2ad)$. This signature has $a' + d' < b' + c'$.

Under a holographic transformation (in the proof of Lemma 3.4.3), in the edge-2-coloring view $\begin{bmatrix} d & 0 & 0 & a \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ a & 0 & 0 & d \end{bmatrix}$ becomes $\frac{1}{2} \begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & a-b+c-d & a+b-c-d & 0 \\ 0 & a+b-c-d & a-b+c-d & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix} = P \cdot \begin{bmatrix} a+d & 0 & 0 & 0 \\ 0 & c-b & 0 & 0 \\ 0 & 0 & a-d & 0 \\ 0 & 0 & 0 & c+b \end{bmatrix} \cdot P^{-1}$, where $P = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$. Connect k copies of f using the 4-ary construction in Figure 3.10, we get the signature f' with $M(f') = \begin{bmatrix} d' & 0 & 0 & a' \\ 0 & b' & c' & 0 \\ 0 & c' & b' & 0 \\ a' & 0 & 0 & d' \end{bmatrix}$ which in the edge-2-coloring view becomes $P \begin{bmatrix} a'+d' & 0 & 0 & 0 \\ 0 & c'-b' & 0 & 0 \\ 0 & 0 & a'-d' & 0 \\ 0 & 0 & 0 & c'+b' \end{bmatrix} P^{-1}$. By construction, it is also equal to

$$P \begin{bmatrix} a+d & 0 & 0 & 0 \\ 0 & c-b & 0 & 0 \\ 0 & 0 & a-d & 0 \\ 0 & 0 & 0 & c+b \end{bmatrix}^k P^{-1} = P \begin{bmatrix} (a+d)^k & 0 & 0 & 0 \\ 0 & (c-b)^k & 0 & 0 \\ 0 & 0 & (a-d)^k & 0 \\ 0 & 0 & 0 & (c+b)^k \end{bmatrix} P^{-1}.$$

It follows that

$$a' = \frac{(a+d)^k + (a-d)^k}{2}, \quad b' = \frac{(b+c)^k - (c-b)^k}{2}, \quad c' = \frac{(b+c)^k + (c-b)^k}{2}, \quad d' = \frac{(a+d)^k - (a-d)^k}{2}.$$

Since $b + c$ is strictly larger than $a + d$, and also $b + c = |b + c| > |c - b|$ and $|a - d|$, one can see that for a large k , we can get f' such that $b' \approx c' \gg a', d'$, and all entries are positive. If furthermore $d' \leq a'$, then we have arrived at a signature f' satisfying (3.4.2) with a possible relabeling. If $d' > a'$, using the construction in Figure 3.10 with two copies of f' (and two $(\neq_2)$ in the middle as we are in the orientation view), we can get f'' with $M(f'') = M(f') \cdot N \cdot M(f') = \begin{bmatrix} 2a'd' & 0 & 0 & a'^2 + d'^2 \\ 0 & 2b'c' & b'^2 + c'^2 & 0 \\ 0 & b'^2 + c'^2 & 2b'c' & 0 \\ a'^2 + d'^2 & 0 & 0 & 2a'd' \end{bmatrix}$. The four parameters for f'' is $(a'', b'', c'', d'') = (a'^2 + d'^2, 2b'c', b'^2 + c'^2, 2a'd')$ which satisfies (3.4.2) (with a possible relabeling) when $b', c' \gg a', d'$.

Now that we have a signature f' (or f'' if the last step is needed) satisfying (3.4.2), we obtain

$\hat{f}$ using 4-ary constructions with f' (or f'' respectively). This is done by repeatedly applying G_2 (defined in Lemma 3.4.4) in which each round starts with the function constructed from the previous round, and possibly relabeling the input values (a, b, c, d) , so that the construction G_2 starts with some (a, b, c, d) satisfying (3.4.2). Note that for any (a, b, c, d) satisfying (3.4.2) and $c > \frac{3}{2}d$, G_2 shrinks the distance between c and d by at least a constant (after normalization). This can be checked by $\frac{c}{d} - \frac{c_2}{d_2} = \frac{(\frac{c}{d}+1)(\frac{c}{d}-1)}{2 \cdot \frac{c}{d}} > \frac{5}{12}$ when $c > \frac{3}{2}d$. Therefore, starting with f' and iterative applying G_2 (constantly many times), we can obtain $\hat{f}$ so that $1 \leq \hat{d} \leq \hat{a} \leq \hat{b} \leq \hat{c} \leq \frac{3}{2}\hat{d}$.

This finishes the construction of a signature satisfying (3.4.2).

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 (3.4.2).

1. G_1 : connect two vertices with signatures f_1 and f_2 respectively as in Figure 3.10. 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 $N = \begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix}$. 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; all of them can be easily verified using (3.4.2):

- d_1 is the weight on sink and source and $0 < d_1 \leq a_1, b_1, c_1$.
- $c_1 = \max(a_1, b_1, c_1, d_1)$.

- $\frac{a_1}{d_1} = \frac{a}{d}, \frac{b_1}{d_1} \leq \frac{b}{d}, \frac{c_1}{d_1} \leq \frac{c}{d}$.
- $c_1 d_1 \leq a_1 b_1$ because $c_1 d_1 - a_1 b_1 = -(b+c)(a-d)(bc-ad) \leq 0$.

2. G_2 : connect two vertices with signatures f_2 as in Figure 3.10. 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 (3.4.2):

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

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 (3.4.2) 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 (3.4.2), as well as $c_1 d_1 \leq a_1 b_1$. 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.,

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

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,

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

Note that c^* may no longer be the largest among a^*, b^*, c^* ; however we will permute them to get $\tilde{a}, \tilde{b}, \tilde{c}$ so that (3.4.2) is still satisfied before proceeding to the next round. This completes the description of our construction in one round which obtains $(\tilde{a}, \tilde{b}, \tilde{c}, \tilde{d})$ 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 permuted triple $\tilde{a}, \tilde{b}, \tilde{c}$ satisfies (3.4.2), the successor sequence as the rounds progress stays with an individual value. E.g., starting from (a, b, c, d) , if after one round $a^* = \max(a^*, b^*, c^*, d^*) = \tilde{c}$, then the successor of $\frac{a}{d}$ after two rounds is $\frac{(\tilde{c})^*}{(\tilde{d})^*}$. 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 $1 \leq i_0 < i_1 < \dots < i_k \leq 3(k+1)$ 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 $i \in \{i_0, \dots, i_k\}$. Since initially we have $1 \leq d \leq a \leq b \leq c \leq \frac{3}{2}d$,

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

Then

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

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

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

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 (3.4.1) is proved. $\square$

Problem : ISING(β).

Instance : Graph $G = (V, E)$.

Output : $Z_{\text{ISING}}(G; \beta) := \sum_{\sigma: V \rightarrow \{0,1\}} \beta^{\text{mono}(\sigma)}$, where $\text{mono}(\sigma)$ denotes the number of edges $\{u, v\}$ such that $\sigma(u) = \sigma(v)$.

Lemma 3.4.5. *The Ising problem ISING $\left(\frac{w}{x}\right)$ is equivalent to the Holant problem Holant $\left(\begin{bmatrix} w & 0 & 0 & x \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ x & 0 & 0 & w \end{bmatrix}\right)$.*

In particular, ISING $\left(\frac{w}{x}\right) \equiv_{\text{AP}} \text{Holant} \left(\begin{bmatrix} w & 0 & 0 & x \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ x & 0 & 0 & w \end{bmatrix}\right)$.

Remark 3.4.2. A non-homogenized form of the Ising model is $\tilde{Z}_{\text{ISING}}(G; x, w) := \sum_{\sigma: V \rightarrow \{0,1\}} w^{\text{mono}(\sigma)} x^{|E| - \text{mono}(\sigma)}$.

If $x \neq 0$ then $\tilde{Z}_{\text{ISING}}(G; x, w) = x^{|E|} Z_{\text{ISING}}(G; \frac{w}{x})$. If $x = 0$ then in $\tilde{Z}_{\text{ISING}}$ all vertices in each component must take the same assignment (all 0 or all 1). In this case both $\tilde{Z}_{\text{ISING}}(G; x, w)$ and the Holant problem in Lemma 3.4.5 are trivially solvable in polynomial time.

Proof. For the problem $\text{Holant} \left(\begin{bmatrix} w & 0 & 0 & x \\ 0 & y & z & 0 \\ 0 & z & y & 0 \\ x & 0 & 0 & w \end{bmatrix} \right)$, 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

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

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

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

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

Given a 4-regular graph $G = (V, E)$ as an instance of $\text{Holant} \left(\begin{bmatrix} w & 0 & 0 & 0 \\ 0 & 0 & z & 0 \\ 0 & z & 0 & 0 \\ 0 & 0 & 0 & w \end{bmatrix} \right)$, we can partition E into a set 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 C is a *crossing circuit* of G . For the graph G , we define its *crossing-circuit graph* $\tilde{G} = (C, \tilde{E})$, with possible multiloops and multiedges, as follows: its vertex set 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 $\tilde{e}_v \in \tilde{E}$ labeled by v . Note that it is possible that $C_1 = C_2$, and for such a self-intersection point the edge $\tilde{e}_v$ is a loop. Each $C \in C$ may have multiple loops, and for distinct circuits C_1 and C_2 there may be multiple edges between them. The edge set $\tilde{E}$ of $\tilde{G}$ is in 1-1 correspondence with V of G .

Observe that the problem $\text{Holant} \left(\begin{bmatrix} w & 0 & 0 & 0 \\ 0 & 0 & z & 0 \\ 0 & z & 0 & 0 \\ 0 & 0 & 0 & w \end{bmatrix} \right)$ 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 $\tilde{e}_v$ if $\sigma'(C_1) = \sigma'(C_2)$ and has local weight z otherwise.

This means

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

Next we show that $\text{ISING} \left(\frac{w}{z} \right)$ can be expressed as $\text{Holant} \left(\begin{bmatrix} w & 0 & 0 & 0 \\ 0 & 0 & z & 0 \\ 0 & z & 0 & 0 \\ 0 & 0 & 0 & w \end{bmatrix} \right)$. 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. $\square$

Proposition 3.4.6 ([GJ08, Lemma 7]). *Suppose $\beta < -1$. Then $\#PM \equiv_{\text{AP}} \text{ISING}(\beta)$.*

Corollary 3.4.7. *Suppose $x \neq 0$ and $\frac{w}{x} < -1$. Then $\#PM \equiv_{\text{AP}} \text{Holant} \left(\begin{bmatrix} w & 0 & 0 & x \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ x & 0 & 0 & w \end{bmatrix} \right)$.*

When $a + d > b + c$ we have $\frac{a+b+c+d}{-a+b+c-d} < -1$, so by Corollary 3.4.7, Lemma 3.4.4, and Lemma 3.4.3, we get $\#PM \equiv_{\text{AP}} \text{Holant} \left(\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix} \right) \leq_{\text{AP}} \text{Holant} \left(\begin{bmatrix} a+b+c+d & 0 & 0 & -a+b+c-d \\ 0 & a-b+c-d & a+b-c-d & 0 \\ 0 & a+b-c-d & a-b+c-d & 0 \\ -a+b+c-d & 0 & 0 & a+b+c+d \end{bmatrix} \right) \equiv_{\text{AP}} \#8V(a, b, c, d)$. By the symmetry of a, b, c , we have proved Theorem 3.4.1.

3.5 NP-hardness

Theorem 3.5.1. *If $(a, b, c, d) \in \overline{\mathcal{X}}$ and at least two of $a, b, c, d > 0$ [¶], then $Z_{8V}(a, b, c, d)$ does not have an FPRAS unless $RP=NP$.*

Remark 3.5.1. For any $(a, b, c, d) \in \overline{\mathcal{X}}$, there are at least two nonzero numbers among a, b, c , and d . The case $d = 0$ and $a, b, c > 0$ was proved in Section 2.3. 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. In this section, we prove the theorem when $d > 0$ and at least one of a, b, c is positive.

Remark 3.5.2. The construction in our proof for the cases when $a > b + c + d$, or $b > a + c + d$, or $c > a + b + d$, is in fact a bipartite graph. This means that approximating $Z_{8V}(a, b, c, d)$ 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 [Yan78]. We reduce 3-MAX CUT to approximating $Z_{8V}(a, b, c, d)$. We first prove the case when $a > b + c + d$, then adapt our proof to the case when $d > a + b + c$. Since the proof of NP-hardness for $Z_{8V}(a, b, c, d)$ is for general (i.e., 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 to the first case.

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.

Let us start with the case when $a > b + c + d$. Recall that we require $d > 0$. First we show how to implement a toy example—a single edge $\{u, v\}$ —by a construction in the eight-vertex model.

[¶]If at most one of a, b, c, d is nonzero, computing the partition function is poly-time tractable.

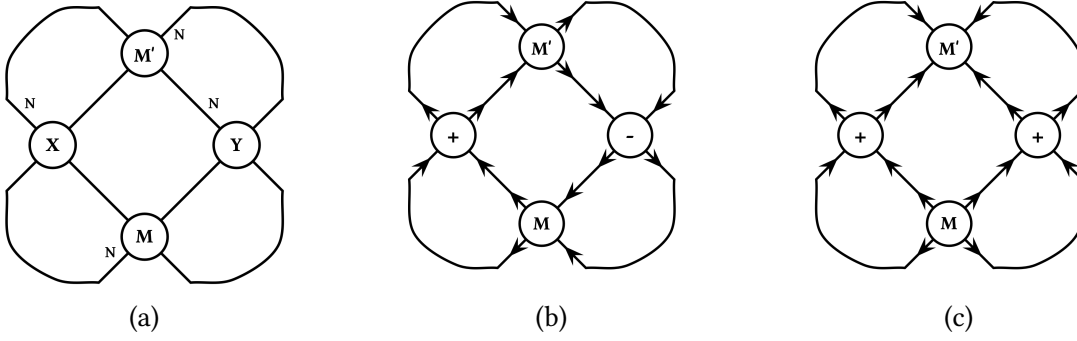


Figure 3.11: A four-way connection implementing a single edge in 3-MAX CUT.

Suppose there are four vertices X, Y, M, M' connected as in Figure 3.11a shows. The order of the 4 edges at each vertex is aligned to Figure 1.1 by a rotation so that the edge marked by “N” corresponds to the north edge in Figure 1.1. 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 1.1-1 or Figure 1.1-2. The same is true for Y . We say X (and similarly Y) is in state $+$ if its local configuration is in Figure 1.1-1 (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 1.1-2 (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 1.1-1 (or both being Figure 1.1-2), with weight a (Figure 3.11b); when (X, Y) is in state $(+, +)$ (or $(-, -)$), M and M' have local configurations both being Figure 1.1-7 or Figure 1.1-8, with weight $d < a$ (Figure 3.11c). This models how two adjacent vertices interact in 3-MAX CUT. We will call the connection pattern described in Figure 3.11a between the set of 4 external edges incident to X and the set of 4 external edges incident to Y (each with two on “top” and two on “bottom”) a *four-way connection*.

To model a vertex of degree 3 in a 3-MAX CUT instance, we use the *locking device* in Figure 3.12a. 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 1.1-1 or Figure 1.1-2. In fact, each locking device has two states, one shown

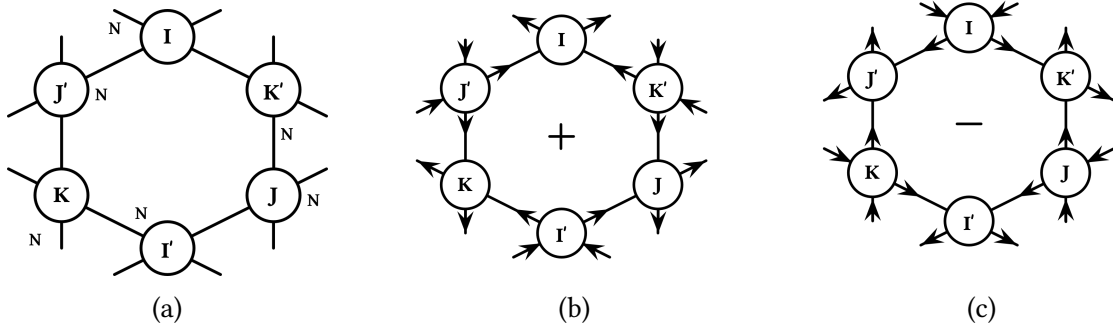


Figure 3.12: A locking device implementing a vertex of degree 3 in 3-MAX CUT.

in Figure 3.12b with every node in configuration Figure 1.1-1 (called the + state) and the other shown in Figure 3.12c with every node in configuration Figure 1.1-2 (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 3.11a), 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.

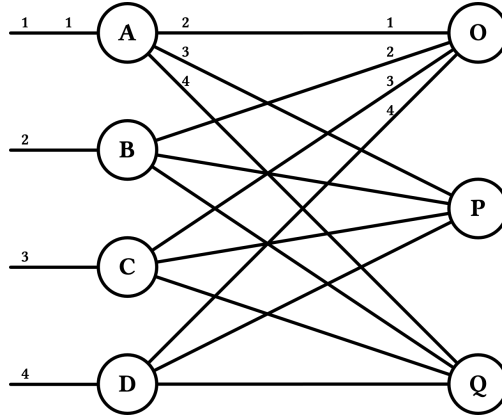


Figure 3.13: A 4-ary construction that amplifies the maximum among a, b, c, d .

Next we show how to enforce the virtual constraint in Figure 3.12a 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 3.13 (input edges of B, C, D labeled similarly as for A and those of P, Q labeled similarly as for O). Starting with (a, b, c, d) on every vertex in Γ (where $a > b + c + d$), 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

$$\begin{aligned} \Lambda(\xi, x, y, z) = & \xi^7 + (3x^4 + 3y^4 + 3z^4 + 4x^2y^2 + 4x^2z^2 + 4y^2z^2)\xi^3 \\ & + (2x^4y^2 + 2x^4z^2 + 2x^2y^4 + 2y^4z^2 + 2x^2z^4 + 2y^2z^4 + 30x^2y^2z^2)\xi. \end{aligned} \quad (3.5.1)$$

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 1.1-1 or Figure 1.1-2. 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 3.5.2 shows that the asymptotic growth rate is exponential in the number of vertices used.

To reduce the problem 3-MAX CUT to approximating $Z_{8V}(a, b, c, d)$, let $\kappa > \lambda \geq 1$ be two constants that will be fixed later. 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 3.12a is created for each $v \in V$, and a four-way connection is made for every $\{u, v\} \in E(G)$, on the external edges corresponding to $\{u, v\}$ as in Figure 3.11a. For each 4-way connection in Figure 3.11a, each of the nodes M, M' is replaced by a $(\lambda \log n)$ -amplifier to boost the ratio of the configurations in Figure 1.1-1 or Figure 1.1-2 over other configurations. For each device in Figure 3.12a, each of the nodes I, I', J, J', K, K' is replaced by a $(\kappa \log n)$ -amplifier to lock in the configurations Figure 3.12b or Figure 3.12c.

Next we argue that the maximum size s of all cuts in G can be recovered from an approximate solution to $Z_{8V}(G'; a, b, c, d)$.

Given a cut (V_+, V_-) of size s in G , we show there is a valid configuration (at the granularity of nodes and edges shown in Figure 3.12a) of weight $\geq (a_{\kappa \log n})^{6n} (a_{\lambda \log n})^{2s} (d_{\lambda \log n})^{2(m-s)}$. For every vertex $u \in V_+$ and every $v \in V_-$ we set the corresponding locking devices in the $+$ state

(Figure 3.12b) and $-$ state (Figure 3.12c) respectively. Consequently, for each edge $\{u, v\}$, the two nodes M and M' in the 4-way connection between the external edges from u and v (two from each) are both in Figure 1.1-1 or Figure 1.1-2 if $\{u, v\}$ is in the cut; they are both in Figure 1.1-7 or Figure 1.1-8 if $\{u, v\}$ is not in the cut. We have defined a valid configuration, and it has weight $\geq \prod_{v \in V} (a_{\kappa \log n})^6 \prod_{e \in V_+ \times V_-} (a_{\lambda \log n})^2 \prod_{e \notin V_+ \times V_-} (d_{\lambda \log n})^2 = (a_{\kappa \log n})^{6n} (a_{\lambda \log n})^{2s} (d_{\lambda \log n})^{2(m-s)}$, where the exponent 6 comes from the 6 nodes I, I', J, J', K, K' in each locking device and 2 comes from the two nodes M, M' in each four-way connection.

We also show that the weighted sum of all configurations is $< \frac{1}{2} (a_{\kappa \log n})^{6n} (a_{\lambda \log n})^{2(s+1)} (d_{\lambda \log n})^{2(m-(s+1))}$, where s is the maximum size of cuts in G . First we bound W_{lock} , the sum of weights for configurations where all nodes labeled I, I', J, J', K, K' are locked. It follows that

$$W_{\text{lock}} \leq 2^n (a_{\kappa \log n})^{6n} \sum_{i=0}^s \binom{m}{i} (a_{\lambda \log n})^{2i} (d_{\lambda \log n})^{2(m-i)} \leq 2^{n+m} (a_{\kappa \log n})^{6n} (a_{\lambda \log n})^{2s} (d_{\lambda \log n})^{2(m-s)},$$

where each locking device has 2 possible states each with weight $(a_{\kappa \log n})^6$, and given a particular $\pm$ assignment of n devices, there can be at most s four-way connections that are between a $+$ device and a $-$ device. Hence $W_{\text{lock}} < \frac{1}{4} (a_{\kappa \log n})^{6n} (a_{\lambda \log n})^{2(s+1)} (d_{\lambda \log n})^{2(m-(s+1))}$ when $\lambda \geq 1$ is large.

It remains to upper-bound the weighted sum of configurations where there is at least one device with some lock broken. This quantity is bounded by

$$\begin{aligned} & 8^{6n} \sum_{i=0}^{6n-1} \binom{6n}{i} (a_{\kappa \log n})^i (b_{\kappa \log n} + c_{\kappa \log n} + d_{\kappa \log n})^{(6n-i)} [2(a_{\lambda \log n} + b_{\lambda \log n} + c_{\lambda \log n} + d_{\lambda \log n})]^{2m} \\ & \leq 2^{24n+2m} (a_{\kappa \log n})^{6n} \left(\frac{b_{\kappa \log n} + c_{\kappa \log n} + d_{\kappa \log n}}{a_{\kappa \log n}} \right) (a_{\lambda \log n} + b_{\lambda \log n} + c_{\lambda \log n} + d_{\lambda \log n})^{2m} \\ & \leq 2^{24n+6m} (a_{\kappa \log n})^{6n} \left(\frac{b_{\kappa \log n} + c_{\kappa \log n} + d_{\kappa \log n}}{a_{\kappa \log n}} \right) (a_{\lambda \log n})^{2m} \\ & \leq 2^{24n+6m} [\Theta(1)]^{mn^{6\lambda}} (a_{\kappa \log n})^{6n} \frac{1}{\beta^{n^\kappa}}, \end{aligned}$$

where we use the fact that $a_{\lambda \log n} \leq a^{64^{\lambda \log n}} = a^{n^{6\lambda}}$ because there are in total $64 = 2^6$ terms in (3.5.1) and $\beta > 1$ by Lemma 3.5.2. This quantity is $< \frac{1}{4} (a_{\kappa \log n})^{6n}$ when $\kappa > \lambda \geq 1$ is sufficiently large.

We have finished the proof for $a > b + c + d$. The case when $b > a + c + d$ or $c > a + b + d$ can be similarly proved. We now adapt our proof to the case when $d > a + b + c$. Since not all $a, b, c = 0$ by our assumption, let us assume without loss of generality that $a > 0$. The amplifier remains exactly the same thanks to its symmetry. For the locking device, we can still lock into two states (with the help of amplifiers on each node in the device): the $+$ state where I, J, K are sources and I', J', K' are sinks; the $-$ state where I, J, K are sinks and I', J', K' are sources. The only difference in the construction is the way that four-way connections are set up (Figure 3.14a). This time, a node marked with I, J, K in one locking device need to be connected to a node marked with I', J', K' (instead of still I, J, K as is the case when $a > b + c + d$) in another locking device to make sure locking devices in contrary states are favored (by setting M and M' as sinks and sources).

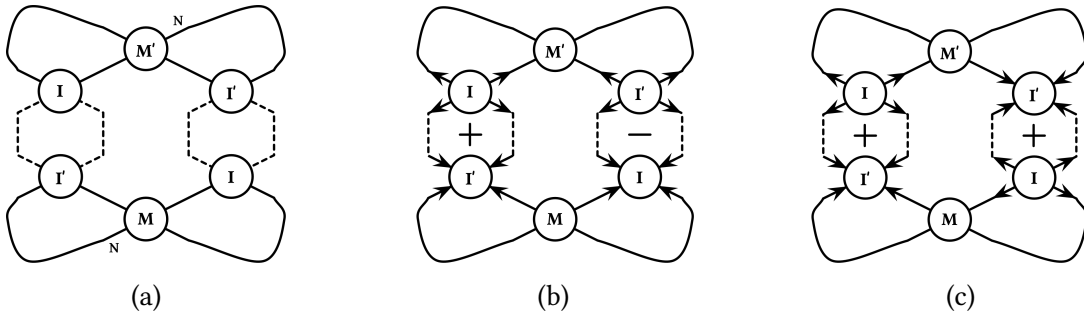


Figure 3.14: Modifying the four-way connection for the case when $d > a + b + c$.

Note that in the case when $a > b + c + d$ (or symmetrically $b > a + c + d$, $c > a + b + d$), the construction G' in the eight-vertex model is bipartite for any (not necessarily bipartite) 3-MAX CUT instance G . To see this, just check that (1) the amplifiers are bipartite and (2) the four way connections and the locking devices are bipartite by setting the nodes marked with M, I, J, K on one side and the nodes marked with M', I', J', K' on the other side. Therefore, approximately computing $Z_{8V}(a, b, c, d)$ in these cases is NP-hard even on bipartite graphs. We remark this is no longer true for the construction of G' in the case when $d > a + b + c$. $\square$

Lemma 3.5.2. *Let $(a_k, b_k, c_k, d_k) = \Lambda^{(k)}(a, b, c, d)$ given by (3.5.1). 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 $k \geq 1$, $\frac{a_k}{b_k + c_k + d_k} \geq \alpha \beta^{2^k}$.*

Proof. Let $(a', b', c', d') = \Lambda(a, b, c, d)$ for any a, b, c, d such that $a, d > 0$ and $b, c \geq 0$. We have $a' > 0, b', c', d' \geq 0$ and $b' + c' + d' > 0$. One can check that $\frac{a'}{b' + c' + d'} > \frac{a}{b + c + d}$ for any such (a, b, c, d) . One can check that

$$\frac{a'}{b' + c' + d'} - \left(\frac{a}{b + c + d} \right)^2 = \frac{a(a - (b + c + d))F}{(b + c + d)^2(b' + c' + d')},$$

where $F = (b^2 + c^2 + d^2 + 2bc + 2bd + 2cd)a^5 + F_4a^4 + F_3a^3 + F_2a^2 + F_1a + F_0$ and F_i ($0 \leq i \leq 4$) are polynomials only in b, c, d (with not necessarily positive coefficients).

Therefore, when $\frac{a}{b+c+d}$ is sufficiently large, $\frac{a'}{b'+c'+d'} - \left(\frac{a}{b+c+d} \right)^2 > 0$. This indicates that the series $\left\{ \frac{a_k}{b_k + c_k + d_k} \right\}_{k \geq 1}$ has a growth rate of 2^k after some finite j such that $\frac{a_j}{b_j + c_j + d_j}$ is sufficiently large. If we normalize $a = 1$, then Λ takes a tuple $(\tilde{b}, \tilde{c}, \tilde{d})$ with $\tilde{b}, \tilde{c}, \tilde{d} \geq 0$ and $0 < \tilde{b} + \tilde{c} + \tilde{d} \leq 1$ and maps it to another tuple $(\tilde{b}', \tilde{c}', \tilde{d}')$ with $\tilde{b}' + \tilde{c}' + \tilde{d}' < \tilde{b} + \tilde{c} + \tilde{d}$. The existence of such a point j is proved in Lemma 3.5.3. This completes the proof by setting $\beta = \frac{a_j}{b_j + c_j + d_j} > 1$. $\square$

Lemma 3.5.3. *Let $\Delta = \{(b, c, d) \mid b, c, d \geq 0, b + c + d \leq 1\}$. Let $s : \Delta \rightarrow \mathbb{R}_+$ be the summation function $s((b, c, d)) = b + c + d$. Suppose $g : \Delta \rightarrow \Delta$ is a continuous function such that for any $x \neq \mathbf{0}$, $s(g(x)) < s(x)$. Then for any $\epsilon > 0$, there exists $N \in \mathbb{Z}_+$ such that for any $x \in \Delta$, $s(g^{(N)}(x)) < \epsilon$.*

Proof. For any $\epsilon > 0$, let $\Delta_\epsilon = \{(b, c, d) \in \Delta \mid b + c + d \geq \epsilon\}$, and consider the continuous function $h(x) = s(x) - s(g(x))$ on Δ_ϵ . Since Δ_ϵ is compact, h reaches its minimum at some x_0 on Δ_ϵ . Since $h(x) > 0$ for any $x \in \Delta_\epsilon$, we have $h(x_0) > 0$. Let $N = \lceil \frac{1}{h(x_0)} \rceil$. Starting from any $x \in \Delta$, we claim that $s(g^{(N)}(x)) < \epsilon$. If not, then $s(g^{(N)}(x)) \geq \epsilon$, and by monotonicity, $g^{(n)}(x) \in \Delta_\epsilon$ for all $0 \leq n \leq N$. But then $s(g^{(N)}(x)) \leq s(x) - Nh(x_0) \leq 0$, a contradiction. $\square$

4 ON PLANAR AND ON BIPARTITE 4-REGULAR GRAPHS

Previous FPRAS results in Section 3.2 are based on the method of *Markov chain Monte Carlo* (MCMC). A nice upper bound on the mixing time of a specific Markov chain can be achieved only in $\mathcal{Z}$ (which is a subregion of the disordered phase $\mathcal{X}$) using a *canonical path argument*. The canonical path argument was introduced for perfect matchings [JS93] and extends well to problems which are believed to have strong connections with perfect matchings [McQ13]. We proved in Section 3.3 that computing $Z_{8v}(a, b, c, d)$ for any $(a, b, c, d) \in \mathcal{Z}$ can be reduced to the problem of counting perfect matchings by expressing the local constraints (a, b, c, d) using “matchgates” whereas such expression provably fails for $(a, b, c, d) \notin \mathcal{Z}$. Moreover, in the ordered phases of the eight-vertex model (the four disjoint parts of $\overline{\mathcal{X}}$), torpid mixing results of natural Markov chains were established on grid/torus graphs (in Chapter 5 and [GR10, Liu18, FR19]).

In this chapter, we give the first FPRAS for $Z_{8v}(a, b, c, d)$ for (a, b, c, d) outside $\mathcal{Z}$ on planar and on bipartite graphs. We simultaneously set up two different kinds of relations between the partition functions of the eight-vertex model and the even-coloring model (introduced in Section 1.2). The first kind of relations exploit the property of planar and of bipartite graphs, and are one-to-one weight-preserving mappings between the states of the eight-vertex model $\mathcal{O}_e(G)$ and the states of the even-coloring model $\mathcal{C}_e(G)$; the second is by the method of *holographic transformation* introduced by Valiant [Val08] which can be thought of as an (exponentially many)-to-(exponentially many) map between the two state spaces (Section 1.3). This, magically, allows us to identify the partition functions of the eight-vertex model on the same graph under *totally different* parameter settings. Interestingly, we show that these maps and their compositions among different parameter settings under which the partition function is preserved have *group structures*. For planar graphs, this group is isomorphic to the *symmetry group* S_3 on three elements (see Section 4.1); for bipartite graphs, this group is isomorphic to the *dihedral group* D_6 of order 12 (the symmetry group of a regular hexagon, see Section 4.2).

Therefore, although the Markov chain on a graph under certain parameter settings outside

$\mathcal{Z}$ is not rapidly mixing, after “mixing up” the state space using a combination of two maps, the Markov chain turns out to be rapidly mixing. Indeed, as a consequence, this “indirect” MCMC leads to FPRAS for new regions in the disordered phase $\mathcal{X}$ and, for the first time, in the ordered phases $\overline{\mathcal{X}}$ for planar graphs and for bipartite graphs.

Theorem 4.0.1. *Let G be a 4-regular plane graph. There is an FPRAS for $Z_{8V}(G; a, b, c, d)$ for (a, b, c, d) in a subregion of $\mathcal{X} \cap \overline{\mathcal{Y}} \cap \overline{\mathcal{Z}}$ and in a subregion of $\overline{\mathcal{X}}$.*

We have proved that on general 4-regular graphs, approximating the eight-vertex model in $\mathcal{X} \cap \overline{\mathcal{Y}}$ is #PM-hard (in Section 3.4) and in $\overline{\mathcal{X}}$ is NP-hard (in Section 2.3 and Section 3.5). Therefore, we have found a family of problems (with parameters ranging in a region of parameter space) having the following provable properties: For the eight-vertex model in the subregion of $\overline{\mathcal{X}}$ given in Theorem 4.0.1 (described more explicitly in Corollary 4.1.3), computing $Z_{8V}(a, b, c, d)$ is

1. NP-hard to approximate on general 4-regular graphs (see Section 2.3 and Section 3.5), and
2. has an FPRAS on planar 4-regular graphs (Section 4.1).

This separation of complexity for general and for planar graphs should be compared and contrasted with the *FKT algorithm* [TF61, Kas61, Kas67] for exact counting of perfect matchings, but here for approximate counting. Previously the combined results of [GvV15] and [HVV15] proved a similar result for k -colorings for general versus planar graphs. We note that it was shown in [GJM15] that approximating the partition functions of many two-state spin systems remain NP-hard on planar graphs. A similar result was shown in [GJ12b] for approximating the Tutte polynomial $T(G; x, y)$ in a large portion of the (x, y) plane. We can also prove that for the subregion, $Z_{8V}(a, b, c, d)$ is #P-complete in exact computation even on planar graphs.

Theorem 4.0.2. *Let G be a 4-regular bipartite graph. There is an FPRAS for $Z_{8V}(G; a, b, c, d)$ for (a, b, c, d) in a subregion of $\mathcal{X} \cap \overline{\mathcal{Y}}$ and in a subregion of $\overline{\mathcal{X}}$.*

Note that the subregions mentioned in Theorem 4.0.2 are disjoint from those mentioned in Theorem 4.0.1. We have proved that on general 4-regular graphs, approximating the eight-vertex

model in $\mathcal{X} \cap \overline{\mathcal{Y}}$ is #PM-hard (in Section 3.4) and in $\overline{\mathcal{X}}$ is NP-hard (in Section 2.3 and Section 3.5). Therefore, we have found a family of problems (with parameters ranging in a region of parameter space) having the following provable properties: For the eight-vertex model in the subregion of $\overline{\mathcal{X}}$ given in Theorem 4.0.2 (described more explicitly in Corollary 4.2.3), computing $Z_{8v}(a, b, c, d)$ is

1. NP-hard to approximate on general 4-regular graphs (see Section 2.3 and Section 3.5), and
2. has an FPRAS on bipartite 4-regular graphs (Section 4.2).

Previously the only problem having similar properties is the antiferromagnetic Ising model on general versus bipartite graphs [SS14, CGG⁺16]. We note that the problem of counting independent sets on bipartite graphs (#BIS) is considered a canonical counting problem of intermediate approximation complexity [DGGJ04]. It is conjectured that #BIS neither has an FPRAS nor is NP-hard to approximate. Many 2-spin systems on bipartite graphs are only known to be #BIS-hard or #BIS-equivalent (below NP-hard) to approximate [CGG⁺16].

To summarize, this chapter establishes the eight-vertex model as the first problem with the provable property that while NP-hard to approximate on general graphs (even #P-hard for planar graphs in exact complexity), it possesses FPRAS on both bipartite graphs and planar graphs in substantial regions of its parameter space.

A key property we use in our proof is that the dual of any planar 4-regular graph is bipartite [Wel69], and hence 2-colorable. Since torus graphs $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z})$ with even m, n also have this property and are bipartite, the results in Theorem 4.0.1 and Theorem 4.0.2 also hold in torus graphs (with even side lengths).

Finally, we note that the techniques introduced in this chapter have other applications. First, the maps between partition functions under different parameter settings established in this chapter are not only useful for giving approximation algorithms. The same maps are useful in our understanding of the exact computational complexity of the eight-vertex model on planar graphs. Second, the techniques are useful for studying the partition functions of other edge-orientation

problems (e.g. other vertex models in statistical physics) or edge-coloring problems (e.g. Holant problems) on planar/bipartite graphs.

4.1 An FPRAS on planar graphs

Lemma 4.1.1. *Let G be a 4-regular plane graph. Then $Z_{8V}(G; a, b, c, d) = Z_{EC}(G; b, a, d, c)$.*

Proof. It is well known that a connected planar graph is Eulerian if and only if its dual is bipartite [Wel69]. Partition functions are multiplicative over connected components, so we may assume G is connected. As G is planar and 4-regular, the dual G^* is bipartite. Hence, we can color the faces of G using two colors, say black and white, so that any two adjacent faces (i.e., they share an edge) are of different colors. For definiteness, we assume that the outer face of G is colored white. See Figure 4.1a for an example. Every edge separates one face colored white and another face colored black, so each edge is on a unique white face. This shows that the binary relation on the set of edges defined by being on the same white face is an equivalence relation.

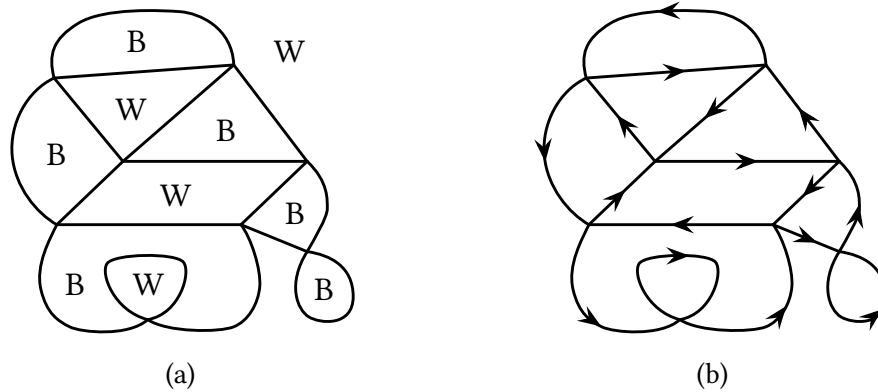


Figure 4.1: A proper 2-coloring of the faces of a planar 4-regular graph and its canonical orientation τ .

Based on the above facts, there is a *canonical orientation* τ which orients all the edges along any (non-outer) white face *clockwise*. This is also the same as to orient edges along every black face *counterclockwise*. See Figure 4.1b for a pictorial illustration. Observe that τ is an Eulerian

orientation of G (at every vertex the in-degree equals the out-degree) and in τ every vertex is in the 5th local configuration in Figure 4.3 (and equivalently the 6th local configuration in Figure 4.4).

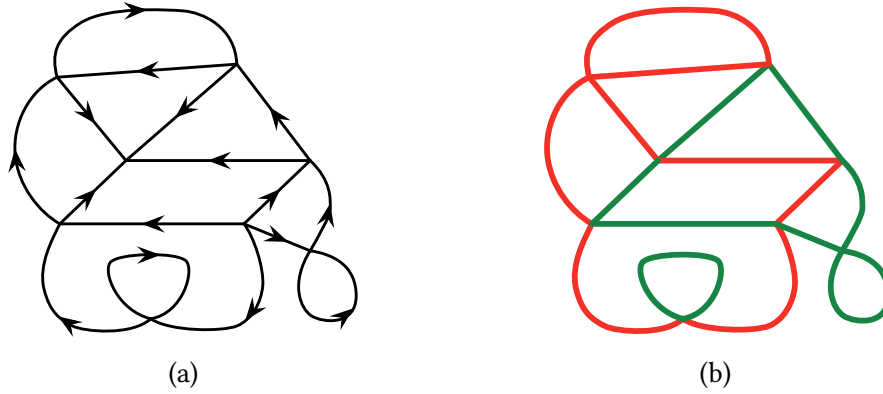


Figure 4.2: An even orientation τ' and its corresponding even coloring.

For an arbitrary orientation τ' of G , we say an edge e is *green* if $\tau'(e) = \tau(e)$, and e is *red* otherwise. Then for any even orientation τ' of G , we can show that this assignment of green-red colors is an *even coloring* of G . (See an illustration in Figure 4.2: Figure 4.2a is an even orientation and Figure 4.2b is its corresponding even coloring.) Every even orientation τ' gives an even coloring because:

- Under this coloring, the canonical orientation τ receives the all-green coloring, which is an even coloring itself.
- For any even orientation τ' , the local configuration of τ' at any vertex differs from the local configuration of τ at the same vertex on an *even* number of edges. These edges receive the red color and the others receive green. Therefore, at each vertex the color assignment will be in one of the states shown in Figure 1.3.

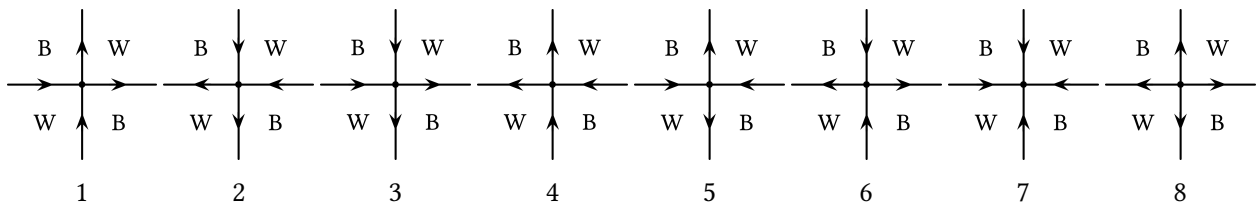


Figure 4.3: Local configurations at one type of vertices in a planar graph.

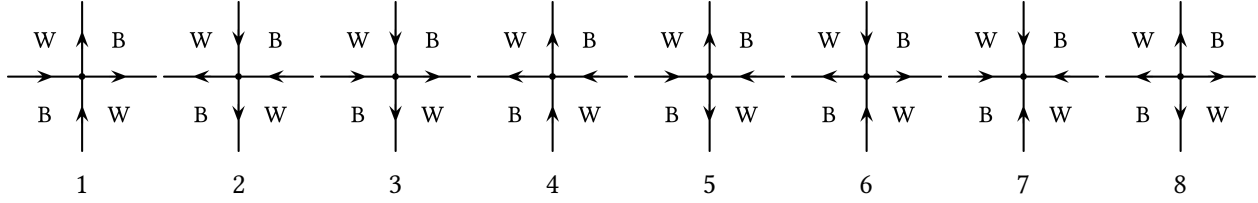


Figure 4.4: Local configurations at the other type of vertices in a planar graph.

We claim that this color assignment on the edges of G gives a bijection M_{PLANAR} from S_{8V} to S_{EC} . Given the black and white coloring of the faces of G , there are two types of vertices in G , either the one in Figure 4.3 or the other one in Figure 4.4. The correspondence of local configurations from Figure 4.3 to Figure 1.3 is $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (3, 4, 1, 2, 7, 8, 5, 6)$; the correspondence of local configurations from Figure 4.4 to Figure 1.3 is $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (4, 3, 2, 1, 8, 7, 6, 5)$. Consequently, M_{PLANAR} induces a one-to-one correspondence between $\mathcal{O}_e(G)$ and $\mathcal{C}_e(G)$.

Although there are two types of maps, both of them map $\{1, 2\}$ to $\{3, 4\}$, $\{3, 4\}$ to $\{1, 2\}$, $\{5, 6\}$ to $\{7, 8\}$, and $\{7, 8\}$ to $\{5, 6\}$. In terms of the weights of local configurations: the arrow reversal symmetry of the eight-vertex model induces an equivalence relation $\{1, 2\}, \{3, 4\}, \{5, 6\}, \{7, 8\}$ in Figure 1.1; the color reversal symmetry of the even-coloring model induces an equivalence relation $\{1, 2\}, \{3, 4\}, \{5, 6\}, \{7, 8\}$ in Figure 1.3. Thus there is a uniform and consistent way to assign a mapping of parameters. One can see that M_{PLANAR} is weight-preserving if the even-coloring model has parameter setting $(w, x, y, z) = (b, a, d, c)$. It follows that

$$Z_{8V}(G; a, b, c, d) = Z_{\text{EC}}(G; b, a, d, c).$$

□

Now that we have set up multiple equations (Lemma 1.3.2, Lemma 1.3.3, and Lemma 4.1.1) between Z_{8V} and Z_{EC} under different mappings in terms of the parameter settings, we can combine them and obtain equations between Z_{8V} under different parameter settings.

Before that, we make the following observation. Since in any even orientation of a 4-regular graph G , the number of sinks (Figure 1.1-7) must be equal to the number of sources (Figure 1.1-8)

and thus their sum is always even, we know that in the eight-vertex model under parameter setting (a, b, c, d) , the weight of a state is unchanged if we flip d to $-d$. Therefore, we have

$$Z_{8V}(G; a, b, c, d) = Z_{8V}(G; a, b, c, -d). \quad (4.1.1)$$

(In particular, for non-negative a, b, c, d , even though $-d$ makes an appearance on the right-hand-side, this equation says the value $Z_{8V}(G; a, b, c, -d) \geq 0$.) Obviously the approximation complexity for computing $Z_{8V}(a, b, c, -d)$ is the same as that for $Z_{8V}(a, b, c, d)$.

Notation. Given a set of 4-tuples S , let $N_d(S) = \{(a, b, c, -d) \mid (a, b, c, d) \in S\}$.

Notation. Given two 4×4 invertible matrices M_1 and M_2 , denote by $\langle M_1, M_2 \rangle$ the group of matrices generated by M_1 and M_2 , where the group operation is matrix multiplication.

Theorem 4.1.2. *Let G be a 4-regular plane graph and let $M_Z^{\text{Pl}} = \frac{1}{2} \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$, $M_{HZ}^{\text{Pl}} = \frac{1}{2} \begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix}$. Then for any $M \in \langle M_Z^{\text{Pl}}, M_{HZ}^{\text{Pl}} \rangle$, $Z_{8V}(G; a, b, c, d) = Z_{8V}(G; a', b', c', d')$ where $\begin{bmatrix} a' \\ b' \\ c' \\ d' \end{bmatrix} = M \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$.*

Remark 4.1.1. One can check that $\langle M_Z^{\text{Pl}}, M_{HZ}^{\text{Pl}} \rangle$ is isomorphic to the symmetry group S_3 and the group elements are shown in Table 4.1.

Proof. First we prove the theorem for M_Z^{Pl} . From Lemma 1.3.2, we know that $Z_{8V}(G; a, b, c, d) = Z_{\text{EC}}(G; w, x, y, z)$ where $\begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$. From Lemma 4.1.1, we know that $Z_{\text{EC}}(G; w, x, y, z) = Z_{8V}(G; a', b', c', d')$ where $\begin{bmatrix} a' \\ b' \\ c' \\ d' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} w \\ x \\ y \\ z \end{bmatrix}$. Therefore, we have $Z_{8V}(G; a, b, c, d) = Z_{8V}(G; a', b', c', d')$ where $\begin{bmatrix} a' \\ b' \\ c' \\ d' \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = M_Z^{\text{Pl}} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$.

The proof for M_{HZ}^{Pl} is similar. Instead of combining Lemma 1.3.2 and Lemma 4.1.1, we simply need to combine Lemma 1.3.3 and Lemma 4.1.1 and notice that $M_{HZ}^{\text{Pl}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$.

Since the theorem is proved for the two invertible matrices M_Z^{Pl} and M_{HZ}^{Pl} , it is also true for the group of matrices generated by these two matrices using their inverse and matrix multiplication.

□

Table 4.1: Elements of $\langle M_Z^{\text{Pl}}, M_{HZ}^{\text{Pl}} \rangle$ and preimages of $\mathcal{Y} = \mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD}$ under corresponding maps. A substantial subregion of each preimage admits FPRAS on planar 4-regular graphs. The last column lists the approximation complexity of the eight-vertex model on general 4-regular graphs in the corresponding preimage regions.

Element	Matrix	Preimage	Approximation
I_4	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD}$	FPRAS in $\mathcal{Z}$
M_Z^{Pl}	$\frac{1}{2} \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{bmatrix}$	$N_d(\mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}} \cap \mathcal{C})$	#PM-hard
$(M_Z^{\text{Pl}})^2$	$\frac{1}{2} \begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ -1 & -1 & 1 & -1 \end{bmatrix}$	$\overline{\mathcal{C}}$	NP-hard
M_{HZ}^{Pl}	$\frac{1}{2} \begin{bmatrix} 1 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \end{bmatrix}$	$\mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}} \cap \mathcal{C}$	#PM-hard
$M_Z^{\text{Pl}} M_{HZ}^{\text{Pl}}$	$\frac{1}{2} \begin{bmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$	$N_d(\overline{\mathcal{C}})$	NP-hard
$(M_Z^{\text{Pl}})^2 M_{HZ}^{\text{Pl}}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$	$N_d(\mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD})$	FPRAS in $N_d(\mathcal{Z})$

Corollary 4.1.3. *Let G be a 4-regular plane graph and let M_Z^{Pl} and M_{HZ}^{Pl} be defined as in Theorem 4.1.2. Then for any $M \in \langle M_Z^{\text{Pl}}, M_{HZ}^{\text{Pl}} \rangle$, there is an FPRAS for $Z_{8V}(G; a, b, c, d)$ if $M \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathcal{Y} \cap \mathcal{Z}$.*

Thus we know that for any $M \in \langle M_Z^{\text{Pl}}, M_{HZ}^{\text{Pl}} \rangle$, there is an FPRAS for $Z_{8V}(G; a, b, c, d)$ for (a, b, c, d) in a subregion of $M^{-1}(\mathcal{Y})$. Note that $\mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}} \cap \mathcal{C} \subset \mathcal{X} \cap \overline{\mathcal{Y}}$ and $\overline{\mathcal{C}} \subset \overline{\mathcal{X}}$. With the help of Table 4.1, one can see that Theorem 4.0.1 is implied by Corollary 4.1.3^{||}.

4.2 An FPRAS on bipartite graphs

Lemma 4.2.1. *Let G be a bipartite graph. Let G' denote its edge-vertex incidence graph. Suppose f satisfies arrow reversal symmetry. Then $\text{Holant}(G'; \neq_2 \mid f) = \text{Holant}(G'; =_2 \mid f) = \text{Holant}(G; f)$. In particular, if G is 4-regular, then $Z_{8V}(G; a, b, c, d) = Z_{EC}(G; a, b, c, d)$.*

^{||}In fact, to prove Theorem 4.0.1 we need $(M_{HZ}^{\text{Pl}})^{-1}(\mathcal{Y} \cap \mathcal{Z}) \subset \mathcal{AD} \cap \mathcal{BD} \cap \overline{\mathcal{CD}} \cap \mathcal{C} \cap \overline{\mathcal{Z}}$ and one can check that this is true.

Proof. For any bipartite graph $G = (L, R, E)$, there is a *canonical orientation* τ which is to orient all the edges from R to L . Let G be a 4-regular bipartite graph. Observe that in τ every vertex in L has local configuration Figure 1.1-7 and every vertex in R has local configuration Figure 1.1-8.

For an arbitrary orientation τ' of G , we say an edge e is *green* if $\tau'(e) = \tau(e)$, and e is *red* otherwise. This coloring assignment on the edges of G gives a bijection $M_{\text{BIPARTITE}}$ from S_{8V} to S_{EC} . For the vertices in L , this can be seen in the entry-wise correspondence $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (1, 2, 3, 4, 5, 6, 7, 8)$ from Figure 1.1 to Figure 1.3; for the vertices in R , the correspondence of local configurations from Figure 1.1 to Figure 1.3 is $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (2, 1, 4, 3, 6, 5, 8, 7)$. Consequently $M_{\text{BIPARTITE}}$ defines a one-to-one correspondence between $\mathcal{O}_e(G)$ and $C_e(G)$. Again because both maps respect the same equivalence relation induced by the arrow reversal symmetry, this one-to-one correspondence $M_{\text{BIPARTITE}}$ is weight-preserving if the even-coloring model has the parameter setting $(w, x, y, z) = (a, b, c, d)$. It follows that

$$Z_{8V}(G; a, b, c, d) = Z_{\text{EC}}(G; a, b, c, d).$$

The idea of the above mapping can be easily extended to general (not necessarily 4-regular) graphs. For a bipartite graph $G = (L, R, E)$ and its edge-vertex incidence graph $G' = (V_E, L \cup R, E')$, every vertex $v_e \in V_E$ has degree 2 and connects a vertex $l \in L$ with a vertex $r \in R$. To see a one-to-one weight-preserving mapping from valid configurations in $\text{Holant}(G'; \neq_2 | f)$ to valid configurations in $\text{Holant}(G'; =_2 | f)$, one simply flips the assignment on every edge $\{v_e, r\}$ such that $v_e \in V_E$ and $r \in R$. $\square$

For any bipartite regular graph $G = (L, R, E)$, we know that $|L| = |R|$ and hence the total number of vertices is always even. In the eight-vertex model under parameter setting (a, b, c, d) , the weight of a state is unchanged if we flip the sign of the weight on every vertex. Therefore, for bipartite 4-regular graphs, in addition to (4.1.1), we also have

$$Z_{8V}(G; a, b, c, d) = Z_{8V}(G; -a, -b, -c, -d). \quad (4.2.1)$$

Notation. Given a set of 4-tuples S , let $N(S) = \{(-a, -b, -c, -d) \mid (a, b, c, d) \in S\}$.

Theorem 4.2.2. *Let G be a 4-regular bipartite graph and let $M_Z^{\text{Bi}} = \frac{1}{2} \begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$, $M_{HZ}^{\text{Bi}} = \frac{1}{2} \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$. Then for any $M \in \langle M_Z^{\text{Bi}}, M_{HZ}^{\text{Bi}} \rangle$, $Z_{8V}(G; a, b, c, d) = Z_{8V}(G; a', b', c', d')$ where $\begin{bmatrix} a' \\ b' \\ c' \\ d' \end{bmatrix} = M \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$.*

Remark 4.2.1. One can check that $\langle M_Z^{\text{Bi}}, M_{HZ}^{\text{Bi}} \rangle$ is isomorphic to the dihedral group D_6 and the group elements are shown in Table 4.2.

Proof. The proof is similar to that of Theorem 4.1.2. Instead of combining the holographic maps in Lemma 1.3.2, Lemma 1.3.3 with the planar map in Lemma 4.1.1, we need to combine them with the bipartite map in Lemma 4.2.1. □

Corollary 4.2.3. *Let G be a 4-regular bipartite graph and let M_Z^{Bi} and M_{HZ}^{Bi} be defined as in Theorem 4.2.2. Then for any $M \in \langle M_Z^{\text{Bi}}, M_{HZ}^{\text{Bi}} \rangle$, there is an FPRAS for $Z_{8V}(G; a, b, c, d)$ if $M \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \in \mathcal{Y} \cap \mathcal{Z}$.*

Thus we know that for any $M \in \langle M_Z^{\text{Bi}}, M_{HZ}^{\text{Bi}} \rangle$, there is an FPRAS for $Z_{8V}(G; a, b, c, d)$ for (a, b, c, d) in a subregion of $M^{-1}(\mathcal{Y})$. Note that $\overline{\mathcal{AD}} \cap \overline{\mathcal{BD}} \cap \overline{\mathcal{CD}} \cap \mathcal{D} \subset \mathcal{X} \cap \overline{\mathcal{Y}}$ and $\overline{\mathcal{D}} \subset \overline{\mathcal{X}}$. With the help of Table 4.2, one can see that Theorem 4.0.2 is implied by Corollary 4.2.3.

Table 4.2: Elements of $\langle M_Z^{\text{Bi}}, M_{HZ}^{\text{Bi}} \rangle$ and preimages of $\mathcal{Y} = \mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD}$ under corresponding maps. A substantial subregion of each preimage admits FPRAS on bipartite 4-regular graphs. The last column lists the approximation complexity of the eight-vertex model on general 4-regular graphs in the corresponding preimage regions.

Element	Matrix	Preimage	Approximation
I_4	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD}$	FPRAS in $\mathcal{Z}$
M_Z^{Bi}	$\frac{1}{2} \begin{bmatrix} -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$	$N_d(\overline{\mathcal{AD}} \cap \overline{\mathcal{BD}} \cap \overline{\mathcal{CD}} \cap \mathcal{D})$	#PM-hard
$(M_Z^{\text{Bi}})^2$	$\frac{1}{2} \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$	$N(\overline{\mathcal{D}})$	NP-hard
$(M_Z^{\text{Bi}})^3$	$-I_4$	$N(\mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD})$	FPRAS in $N(\mathcal{Z})$
$(M_Z^{\text{Bi}})^4$	$-M_Z^{\text{Bi}}$	$N(N_d(\overline{\mathcal{AD}} \cap \overline{\mathcal{BD}} \cap \overline{\mathcal{CD}} \cap \mathcal{D}))$	#PM-hard
$(M_Z^{\text{Bi}})^5$	$-(M_Z^{\text{Bi}})^2$	$\overline{\mathcal{D}}$	NP-hard
M_{HZ}^{Bi}	$\frac{1}{2} \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$	$\overline{\mathcal{AD}} \cap \overline{\mathcal{BD}} \cap \overline{\mathcal{CD}} \cap \mathcal{D}$	#PM-hard
$M_Z^{\text{Bi}} M_{HZ}^{\text{Bi}}$	$\frac{1}{2} \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & -1 & 1 \\ -1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$	$N(N_d(\overline{\mathcal{D}}))$	NP-hard
$(M_Z^{\text{Bi}})^2 M_{HZ}^{\text{Bi}}$	$\begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$N(N_d(\mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD}))$	FPRAS in $N(N_d(\mathcal{Z}))$
$(M_Z^{\text{Bi}})^3 M_{HZ}^{\text{Bi}}$	$-M_{HZ}^{\text{Bi}}$	$N(\overline{\mathcal{AD}} \cap \overline{\mathcal{BD}} \cap \overline{\mathcal{CD}} \cap \mathcal{D})$	#PM-hard
$(M_Z^{\text{Bi}})^4 M_{HZ}^{\text{Bi}}$	$-M_Z^{\text{Bi}} M_{HZ}^{\text{Bi}}$	$N_d(\overline{\mathcal{D}})$	NP-hard
$(M_Z^{\text{Bi}})^5 M_{HZ}^{\text{Bi}}$	$-(M_Z^{\text{Bi}})^2 M_{HZ}^{\text{Bi}}$	$N_d(\mathcal{AD} \cap \mathcal{BD} \cap \mathcal{CD})$	FPRAS in $N_d(\mathcal{Z})$

5 ON THE SQUARE LATTICE

In the first part of this chapter, we show the following two theorems in Section 5.1 and Section 5.2, respectively. Let Λ be a square region on the square lattice.

Theorem 5.0.1 (Ferroelectric phase). *The directed-loop algorithm for the six-vertex model under parameter settings (a, b, c) with $a > b + c$ or $b > a + c$ (i.e. the whole ferroelectric phase) mixes torpidly on Λ with periodic boundary conditions.*

Remark 5.0.1. We note that for *periodic boundary conditions* Glauber dynamics is *not* irreducible, so we do not consider that.

Theorem 5.0.2 (Anti-ferroelectric phase). *Both Glauber dynamics and the directed-loop algorithm for the six-vertex model under parameter settings (a, b, c) with $c \geq 2.639 \max(a, b)$ (in the anti-ferroelectric phase) mix torpidly on Λ with free boundary conditions and periodic boundary conditions.*

Parameter settings covered by the above two theorems are all in the ordered phases, i.e., in the low temperature regime. Given that the F model in statistical mechanics is a special case of the six-vertex model when $a = b = 1$ [Lie67a], Theorem 5.0.2 holds for the F model with $c \geq 2.639$.

Our proofs build on the equivalence between small conductance and torpid mixing by Jerrum and Sinclair [SJ89]. When arguing Markov chains for the six-vertex model in the anti-ferroelectric phase have small conductance, we switch our view between finite regions of the square lattice and their *medial lattices*. This transposition allows us to adopt a *Peierls argument* which has been used in statistical physics to prove the existence of phase transitions (e.g., [Pei36, BKW73]), and in theoretical computer science to prove the torpid mixing of Markov chains (e.g., [Ran06, BGRT13]).

In the proof of Theorem 5.0.2, we introduce a version of the *fault line argument* for the six-vertex model. Fault line arguments are introduced by Dana Randall [Ran06] for the lattice hardcore gas and latter adapted in [LPW06] for the lattice ferromagnetic Ising, which proves torpid mixing

of Markov chains via topological obstructions. The constant 2.639 comes from an upper bound for the connective constant for the square lattice self-avoiding walks [GC01].

Although MCMC methods fail at low temperatures, there is still hope to approximately count and sample efficiently. In the second part of this chapter, we give the first efficient approximate counting and sampling algorithms for the six- and the eight-vertex models on finite regions of the square lattice $\mathbb{Z}^2$ with free or periodic boundary conditions in all the ordered phases at sufficiently *low temperatures*. Our work uses cluster expansions, an old tool in statistical physics, combined with Barvinok's method to approximate partition functions via Taylor expansions [Bar16, PR17], a recent approach pioneered by Helmuth, Perkins, and Regts [HPR19]. Being the first to combine cluster expansion and Barvinok's method, [HPR19] gave efficient approximation algorithms for the low-temperature ferromagnetic Potts model and the high-fugacity hard-core model on finite regions of the lattice $\mathbb{Z}^d$ with special boundary conditions and on the torus $(\mathbb{Z}/n\mathbb{Z})^d$ when natural Markov chains are known to mix torpidly.

Definition 5.0.1. Suppose $f : \Sigma^* \rightarrow \mathbb{R}$ is a function mapping problem instances to real numbers. A *fully polynomial-time approximation scheme (FPTAS)* for a problem is a deterministic algorithm that takes as input an instance x and $\epsilon > 0$, running in time polynomial in n (the input length) and ϵ^{-1} , and outputs a number Z such that $(1 - \epsilon)f(x) \leq Z \leq (1 + \epsilon)f(x)$.

Definition 5.0.2. An ϵ -*approximate sample* from a probability measure μ on Ω is a configuration drawn according to $\hat{\mu}$ with total variational distance $\|\hat{\mu} - \mu\|_{\text{TV}} < \epsilon$. Suppose μ_n is a sequence of probability measures indexed by a parameter n . An *efficient sampling algorithm* is a randomized algorithm that returns an ϵ -approximate sample to μ_n and runs in time polynomial in n and ϵ^{-1} .

Our results can be summarized by the following theorem, in which we assume the energy parameters $\epsilon_1, \epsilon_3, \epsilon_5, \epsilon_7$ are rational numbers.

Theorem 5.0.3. *For any parameters $\epsilon_1, \epsilon_3, \epsilon_5, \epsilon_7$ (with a unique minimum value), there exists $\beta_* = \beta_*(\epsilon_1, \epsilon_3, \epsilon_5, \epsilon_7) > \beta_c$ such that for all inverse temperature $\beta > \beta_*$ there is an FPTAS and an efficient*

sampling algorithm for the zero-field eight-vertex model on the torus $(\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$ or on any finite region of the square lattice $\mathbb{Z}^2$ with free boundary conditions.

Remark 5.0.2. Our algorithms work for the eight-vertex model with inverse temperature $\beta > \beta_* > \beta_c$. In terms of parameters $a, b, c, d \geq 0$, there exists some λ such that there is an FPTAS and an efficient sampling algorithm when $a > \lambda \max(b, c, d)$, $b > \lambda \max(a, c, d)$, $c > \lambda \max(a, b, d)$, or $d > \lambda \max(a, b, c)$. These regions are all in the ordered (ferroelectric or anti-ferroelectric) phases.

Our proof has several ingredients.

In Section 5.3, we revisited the even-coloring model (introduced in Section 1.2) to unify the low-temperature eight-vertex model under different ordered phases so that the ground states of the even-coloring model correspond to the ground states of the eight-vertex model. A by-product of this unified view is that the torpid mixing result [GR10] for a natural Markov chain when d is sufficiently large translates to the torpid mixing when a or b or c is sufficiently large (Theorem 5.3.1, Corollary 5.3.2). This is achieved via a set of Markov chain isomorphisms that transform the states and the moves in a one-to-one fashion.

In Section 5.4, from the even-coloring model we introduce an abstract polymer model equivalent to it on the square lattice with free boundary conditions. This is done by representing the deviations from the two ground states and such “contour” representations have been used in statistical physics to understand phase transitions [BKW73]. Thus the partition function of the polymer model is the same as that of the even-coloring model (up to an easily computable factor). On the torus, however, the above polymer model partition function contains extra terms in addition to the terms corresponding to the even-coloring model partition function. In Section 5.7, we present four different polymer models possessing the same set of underlying polymers but with different weight functions (signs on some polymers are altered). As a consequence, the partition function of the even-coloring model can be expressed as the sum of the partition functions of the four polymer models (up to some factor) which we can approximate. To guarantee that an FPTAS for each of the four quantities gives an FPTAS for their sum, we further use a *Peierls argument* to show that

the terms corresponding to the even-coloring model partition function are non-negligible in each polymer model partition function.

In Section 5.5, we prove that the partition function of the abstract polymer model, as a function of the inverse temperature, does not vanish outside a disc in the complex plane. We do this by using a general condition from statistical physics [KP86, FV17] ensuring the convergence of its cluster expansion. We also apply *Cauchy–Hadamard theorem* and the *analytic implicit function theorem* from complex analysis.

In Section 5.6, we apply the framework recently developed by Helmuth, Perkins, and Regts [HPR19] to give efficient approximation algorithms for abstract polymer models when cluster expansion converges. Concretely, the FPTAS is obtained by expressing the log partition function as a Taylor series, using the absence of zeros (i.e. convergence of cluster expansion) to bound the error of the truncated Taylor series, and computing the low-order coefficients of the Taylor series efficiently [Bar16, PR17].

We note that all these other successful attempts described above are for models whose partition functions are only proved to be #BIS-hard to approximate (#BIS denotes the problem of counting independent sets on bipartite graphs; its approximate complexity is widely believed to be below NP-hardness), but not known to be NP-hard**. On the other hand, it was proved in Section 3.5 that approximately computing the partition function of the eight-vertex model is NP-hard even on bipartite graphs in the ordered phases when $a > b + c + d$, $b > a + c + d$, or $c > a + b + d$. Thus, the eight-vertex model under parameter settings described in Theorem 5.0.3 when a , b , or c is large are the *first* problems that are known to be NP-hard on bipartite graphs but admit FPTAS on square lattice graphs.

We also note that the polymers introduced in [HPR19] and used in other aforementioned works for approximation purpose are restricted to connected components in the original input lattice/graph. We show a more flexible use of the method by going beyond their formulation, as our polymers are defined on the superposition of two lattices dual to each other for which the

**These include ferromagnetic Potts on general graphs [GJ12a], counting independent sets and q -colorings on bipartite graphs [DGGJ04].

original $\mathbb{Z}^2$ is their *medial lattice* (see Section 5.4).

5.1 Torpid mixing for the ferroelectric six-vertex model

In this section we prove Theorem 5.0.1 that $\mathcal{M}_D$ in the directed-loop algorithm for the six-vertex model in the ferroelectric phase is torpid mixing on Λ_n with periodic boundary conditions. The directed-loop algorithm Markov chain, denoted by $\mathcal{M}_D$, is formally defined in Section 3.2 for the eight-vertex model on general 4-regular graphs. Here we describe $\mathcal{M}_D$ for the six-vertex model on square lattice regions at a high level. We also briefly introduce the way to bound the mixing time via a *conductance argument*.

Directed-loop algorithm

The state space of $\mathcal{M}_D$ is not only Ω , the “perfect” Eulerian orientations, but also the set of all “near-Eulerian” orientations, denoted by Ω' . For example, in Figure 5.1 the state τ_{ur} is in Ω and all other five states are in Ω' . We think of each edge in Λ_n as the two half-edges cut in the middle, and each of the half edge can be oriented independently. We say an orientation of all the half-edges is *perfect* (in Ω) if every pair of half-edges is oriented consistently and the ice-rule is satisfied at every vertex (except for boundary vertices under free boundary conditions); an orientation is *near-Eulerian* (in Ω') if there are exactly two pairs of half-edges p_1 and p_2 not oriented consistently and the ice-rule is satisfied at every vertex (except for boundary vertices under free boundary conditions), with the restriction that if two half-edges in p_1 are oriented toward each other then in p_2 the two half-edges must be oriented against each other and vice versa.

The transitions in $\mathcal{M}_D$ are Metropolis moves among “neighboring” states. An Ω state τ and an Ω' state τ' are neighboring if τ' can be transformed from τ by picking two half-edges e_1, e_2 incident to a vertex v with one pointing inwards v and the other pointing outwards v (or two half-edges e_1, e_2 on the boundary with one pointing towards the boundary and the other pointing against the boundary), and reverse the direction of e_1 and e_2 together. For instance, in Figure 5.1

$\{\tau_{ur}, \tau_{1b}\}$ and $\{\tau_{ur}, \tau_{1c}\}$ are two pairs of neighboring states. An Ω' state τ'_1 and another Ω' state τ'_2 are neighboring if τ'_2 can be transformed from τ'_1 by “shifting” one pair of conflicting half-edges one step away, while fixing the other pair of conflicting half-edges. For example, in Figure 5.1 τ_{1c} and τ_2 are neighboring to each other. $\mathcal{M}_D$ can be proved to be ergodic and converges to the Gibbs measure on $\Omega \cup \Omega'$ with both free boundary conditions and periodic boundary conditions (Section 3.2).

Mixing time

The *mixing time* t_{mix} measures the time required by a Markov chain to evolve to be close to its stationary distribution, in terms of *total variation distance*. (The definition of mixing time can be found in [LPW06].) We say a Markov chain is torpid mixing if the mixing time is exponentially large in the input size. A common technique to bound the mixing time is via bounding conductance, defined by Jerrum and Sinclair [SJ89].

Let π denote the stationary distribution of an *ergodic* and *time reversible* ($\pi(x)P(x, y) = \pi(y)P(y, x)$ for any $x, y \in \Omega$) Markov chain $\mathcal{M}$ on a finite state space Ω , with transition probabilities $P(x, y)$, $x, y \in \Omega$. The *conductance* of $\mathcal{M}$ is defined by

$$\Phi = \Phi(\mathcal{M}) = \min_{\substack{S \subset \Omega \\ 0 < \pi(S) \leq \frac{1}{2}}} \frac{Q(S, \bar{S})}{\pi(S)},$$

where $Q(S, \bar{S})$ denotes the sum of $Q(x, y) = \pi(x)P(x, y)$ over edges in the transition graph of $\mathcal{M}$ with $x \in S$, and $y \in \bar{S} = \Omega \setminus S$.

In order to show a Markov chain mixes torpidly, we only need to prove that the conductance is (inverse) exponentially small due to the following bound [LPW06]:

$$t_{\text{mix}} = t_{\text{mix}}\left(\frac{1}{4}\right) \geq \frac{1}{4\Phi}.$$

As is usually assumed, Markov chains studied in this dissertation are all *lazy* ($P(x, x) = \frac{1}{2}$ for

any $x \in \Omega$) and transition probabilities $(P(x, y))$ for $x, y \in \Omega$ between neighboring states (where $P(x, y) > 0$) are at least inverse polynomially large. Therefore, armed with the above bound, we can prove the torpid mixing of a Markov chain if we can establish the following:

1. Partition the state space Ω into three subsets $\Omega_{\text{LEFT}} \cup \Omega_{\text{MIDDLE}} \cup \Omega_{\text{RIGHT}}$ as a disjoint union.
2. Show that for any state $\tau_l \in \Omega_{\text{LEFT}}$ and $\tau_r \in \Omega_{\text{RIGHT}}$, $P(\tau_l, \tau_r) = 0$. Under the assumption that the Markov chain is irreducible (i.e., the transition graph is strongly connected), this indicates that in order to go from states in Ω_{LEFT} to states in Ω_{RIGHT} , the Markov process has to go through the “middle states” Ω_{MIDDLE} .
3. Demonstrate that $\pi(\Omega_{\text{MIDDLE}})$ is exponentially small (compared with $\min(\pi(\Omega_{\text{LEFT}}), \pi(\Omega_{\text{RIGHT}}))$) in the input size. This means that starting from any state in Ω_{LEFT} , the probability of going through Ω_{MIDDLE} (and consequently to any state in Ω_{RIGHT} and reach stationarity) is exponentially small. Hence the conclusion of torpid mixing.

Now we are ready to prove Theorem 5.0.1. For any parameter setting (a, b, c) in the ferroelectric phase, either $a > b + c$ or $b > a + c$. By symmetry, without loss of generality, suppose $a > b + c$. This implies that vertex configurations as shown in Figure 1.1-1 and Figure 1.1-2 have higher weights than others. Under the periodic boundary condition, there is a state τ_{ur} in which every vertical edge points upwards and every horizontal edge points to the right (Figure 5.1a), i.e., every vertex on Λ_n is in local configuration shown in Figure 1.1-1. The total weight of τ_{ur} is a^{n^2} as there are n^2 vertices on Λ_n .

For $\mathcal{M}_D$, the three-way partition of the state space $\Omega \cup \Omega'$ is as follows. Denote by T_i the states that can be reached from τ_{ur} in at most i steps of transitions where i is a nonnegative integer. Write $\partial T_i = T_i \setminus T_{i-1}$ for $i \geq 1$. Let $\Omega_{\text{LEFT}} = T_{n-1}$, $\Omega_{\text{MIDDLE}} = \partial T_n$, and $\Omega_{\text{RIGHT}} = (\Omega \cup \Omega') \setminus (\Omega_{\text{LEFT}} \cup \Omega_{\text{MIDDLE}})$. It is obvious that $\Omega \cup \Omega' = \Omega_{\text{LEFT}} \cup \Omega_{\text{MIDDLE}} \cup \Omega_{\text{RIGHT}}$ is a partition of the state space. Clearly $\tau_{ur} \in \Omega_{\text{LEFT}}$, thus the total weight of Ω_{LEFT} is no less than a^{n^2} , the weight of τ_{ur} .

Before proving the total weight of Ω_{MIDDLE} is exponentially small compared with that of Ω_{LEFT} or Ω_{RIGHT} , let us look at what is in T_i with $0 \leq i \leq n$. T_0 is just $\{\tau_{ur}\}$. ∂T_1 consists of all the states

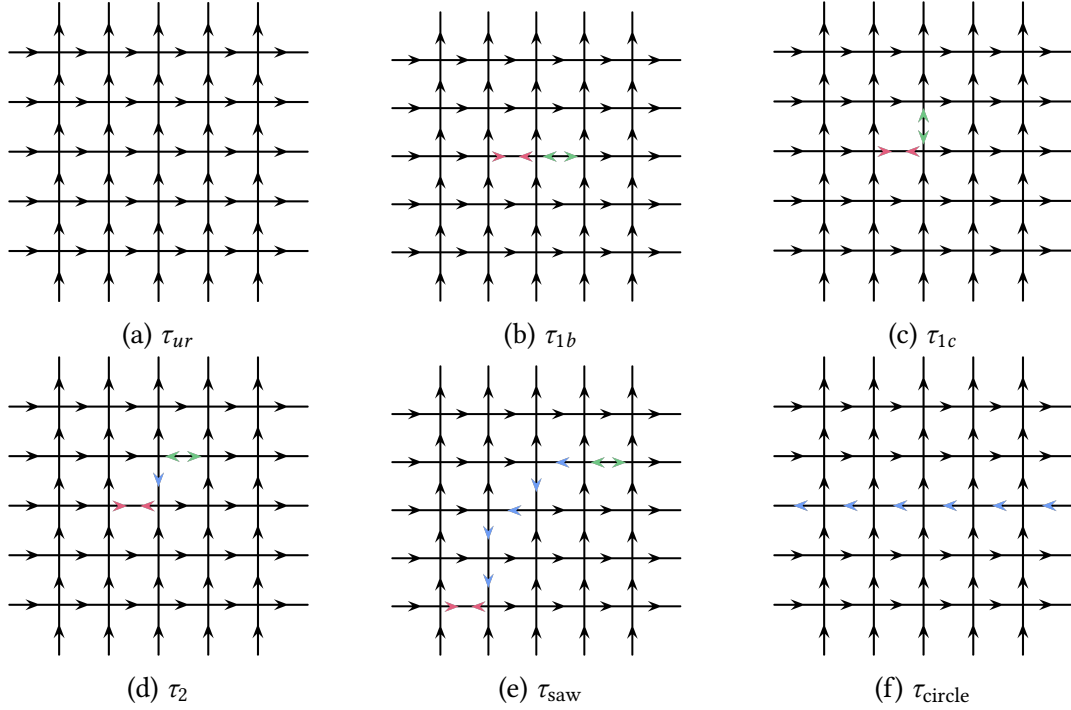


Figure 5.1: Some states in the state space of $\mathcal{M}_D$.

evolved from τ_{ur} by picking a vertex v on Λ_n and two incident half-edges (one pointing towards v and the other away from v), and then reversing the orientations on these two edges. After such a transition, two pairs of conflicting half-edges are created, so $\partial T_1 \subseteq \Omega'$.

For example, the states shown in Figure 5.1b (state τ_{1b}) and Figure 5.1c (state τ_{1c}) are in ∂T_1 . The weight of τ_{1b} is $a^{n^2-1}b$ and that of τ_{1c} is $a^{n^2-1}c$. For every state in ∂T_1 obtained by transitions from τ_{ur} , there is exactly one vertex v^* on Λ_n no longer in the local configuration Figure 1.1-1. (Of course no vertex can be in state Figure 1.1-2.) Actually, depending on whether the two pairs of conflicting half-edges are: (1) both vertical, (2) both horizontal, or (3) one horizontal and the other vertical, the vertex v^* is in configuration shown in (1) Figure 1.1-3, (2) Figure 1.1-4, or (3) Figure 1.1-5/Figure 1.1-6, respectively. Therefore, every state in ∂T_1 has weight $a^{n^2-1}b$ in case (1) and case (2) or $a^{n^2-1}c$ in case (3).

Transitions from states in ∂T_1 to states in ∂T_2 are composed of “shifting” one of the two conflicting pairs of half-edges to a neighboring edge on Λ_n . For example, the state in Figure 5.1d is in ∂T_2 . This process will result in exactly two vertices on Λ_n not in local configuration Figure 1.1-1

(nor in Figure 1.1-2). As a consequence, the weight of any state in ∂T_2 is among $a^{n^2-2}b^2$, $a^{n^2-2}bc$, and $a^{n^2-2}c^2$. The state shown in Figure 5.1d has weight $a^{n^2-2}c^2$.

This line of argument can be extended for ∂T_i for $1 \leq i \leq n$, in any state of which there are exactly i vertices on Λ_n not in local configuration Figure 1.1-1 (nor in Figure 1.1-2). When two conflicting pairs of half-edges are created in ∂T_1 , one of them is above or to the right of another (or both). Denote the former by p_{ur} (the green pair in Figure 5.1) and the latter by p_{dl} (the red pair in Figure 5.1). Observe that as the Markov chain evolves, by a single step, from a state in ∂T_i to another in ∂T_{i+1} (where $1 \leq i \leq n-1$), either p_{ur} is “pushed” up or to the right, or p_{dl} down or to the left. For example, from Figure 5.1c to Figure 5.1d, p_{ur} is pushed to the right. By induction, p_{ur} is always above or to the right of p_{dl} (when $i < n$). A direct consequence is that there can be no state containing a closed circuit formed by the reversed edges (with regard to τ_{ur}) in ∂T_i until $i = n$. Therefore, the edges reversed in any state in ∂T_i ($1 \leq i \leq n$) can be seen as either a *self-avoiding walk* between the middle points of the two pairs of conflicting half-edges (e.g. Figure 5.1e) or a *self-avoiding circuit* (e.g. Figure 5.1f). In fact, when the reversed edges form a circuit, the circuit must “go straightforward” at each step. This circuit is a circle parallel or perpendicular to the torus equatorial plane. The weight of any state in ∂T_i is $a^{n^2-i}b^j c^k$ with $i = j + k$, where the values of j and k depend on how many “turnarounds” are there in the self-avoiding walk.

Therefore, the total weight of states in ∂T_n is at most $n^2 \cdot a^{n^2-n}(b+c)^n$, where n^2 is an upper bound on all the possible starting points for self-avoiding walks, and each monomial in $(b+c)^n$ is from a unique self-avoiding walk. Combining with the fact that total weight of Ω_{LEFT} is at least a^{n^2} (the weight of τ_{ur}) and is at most that of Ω_{RIGHT} (because there is a weight-preserving injective map from Ω_{LEFT} to Ω_{RIGHT} by reversing orientations of all the edges), we know that the conductance of $\mathcal{M}_D$ is at most $\frac{n^2 \cdot a^{n^2-n}(b+c)^n}{a^{n^2}} = n^2 \left(\frac{b+c}{a}\right)^n$. This is exponentially small in n since $a > b+c$ are fixed constants in the ferroelectric phase.

5.2 Torpid mixing for the anti-ferroelectric six-vertex model

In this section, we introduce Glauber dynamics $\mathcal{M}_G$ for the six-vertex model and prove Theorem 5.2.1 which is part of Theorem 5.0.2. After proving Theorem 5.2.1, we state the ideas needed to extend it to Theorem 5.0.2.

Glauber dynamics

Denote by Λ_n a square lattice region where there are n vertices of degree 4 on each row and each column. Λ_n is in *periodic boundary condition* if it forms a two-dimensional torus; the *free boundary condition* can be formulated in the following way: there are $n + 2$ vertices on each row and each column, where the “boundary vertices” are of degree 1 and don’t need to satisfy the ice-rule (and don’t take weights) in a valid six-vertex configuration. For convenience, we assume there are “virtual edges” connecting every two boundary vertices with unit distance on $\mathbb{Z}^2$. A virtual edge does not have orientations, serving only the purpose that every unit square inside the $(n + 1) \times (n + 1)$ region is closed.

Let Ω be the set of all valid configurations of the six-vertex model (Eulerian orientations) on Λ_n . The Glauber-dynamics Markov chain, which we will denote by $\mathcal{M}_G$, has state space Ω . To move from one configuration to another, this chain selects a unit square (a *face*) s on Λ_n (together with the virtual edges) uniformly at random. If all the non-virtual edges along the unit square s are oriented consistently (clockwise or counter-clockwise), the chain picks a direction d (clockwise or counter-clockwise) and reorients the non-virtual edges along s according to the Gibbs measure.

One can easily check that such transitions take valid configurations to valid configurations. Actually, this Markov chain is equivalent to that in [GMP04] for sampling three-colorings on the faces of Λ_n . The ergodicity of that chain translates straightforwardly to the ergodicity of $\mathcal{M}_G$ (with free boundary conditions) thanks to the equivalence between Eulerian orientations and three-colorings on $\mathbb{Z}^2$. Besides, the heat-bath move indicates that the stationary distribution of

$\mathcal{M}_G$ is the Gibbs distribution for the six-vertex model.

As we did in the ferroelectric phase, the intuition behind our proof for the anti-ferroelectric phase is to find a partition $\Omega = \Omega_{\text{LEFT}} \cup \Omega_{\text{MIDDLE}} \cup \Omega_{\text{RIGHT}}$ of the state space of $\mathcal{M}_G$, i.e., all the Eulerian orientations on Λ_n . However, the strategy is different from that used in Section 5.1 — here the subset Ω_{MIDDLE} is determined in terms of a topological obstruction.

Theorem 5.2.1 (Anti-ferroelectric phase). *Glauber dynamics for the six-vertex model under parameter settings (a, b, c) with $c \geq 2.639 \max(a, b)$ mix torpidly on Λ_n with free boundary conditions.*

Observe that there are two states in Ω with maximum weights: τ_G (Figure 5.2a) and τ_R (Figure 5.2b) where every vertex is in local configuration Figure 1.1-5 or Figure 1.1-6, and thus has vertex weight c . Since τ_G and τ_R are total reversals of each other in edge orientations, for any edge in any state $\tau \in \Omega$, it is oriented either as in τ_G or as in τ_R . Let us call an edge to be *green* if it is oriented as is in τ_G and *red* otherwise. Observe that in order to satisfy the ice-rule (2-in-2-out), the number of green (and thus also two red) edges incident to any vertex (except for the boundary vertices) is always even (0, 2, or 4), and if there are two green (or red) edges they must be *rotationally adjacent* to each other. See Figure 5.2c for an example. Also note that the four edges along a unit square on $\mathbb{Z}^2$ are all red edges or all green edges if and only if they are oriented consistently, hence flippable by a single move of $\mathcal{M}_G$.

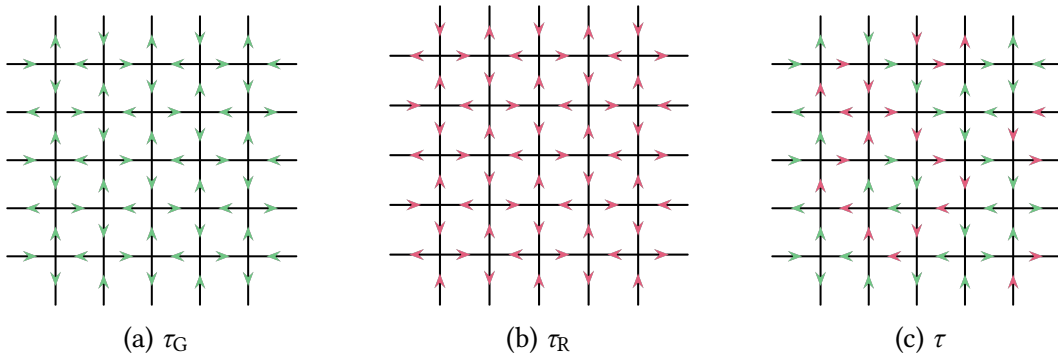


Figure 5.2: Some states in the state space of $\mathcal{M}_G$.

We say a simple path from a horizontal edge on the left boundary of Λ_n to a horizontal edge on the right boundary of Λ_n is a *horizontal green (or red) bridge* if the path consists of only green

(or red, respectively) edges; a *vertical green (or red) bridge* is defined similarly. A state $\tau \in \Omega$ has a *green cross* if it has both a green horizontal bridge and a green vertical bridge; a *red cross* is defined similarly. Let $C_G \subset \Omega$ denote the states having a green cross and C_R the states having a red cross. In the following lemma, we prove that $C_G \cap C_R = \emptyset$.

Lemma 5.2.2. *A green cross and a red cross cannot coexist.*

Proof. It suffices to show that a green horizontal bridge precludes a red vertical bridge. Consider a virtual point v_L sitting to the left of Λ_n connected by an edge to every (external) vertices of Λ_n on the left boundary, and another virtual point v_R connected by an edge to every vertex of Λ_n on the right boundary. Connect v_L and v_R by an edge below Λ_n .

If there is a green horizontal bridge, then by definition there is a continuous closed curve C formed by the bridge and some edges we added (Figure 5.3a). According to the *Jordan Curve Theorem*, C separates the plane into two disjoint regions, the inside and the outside. Vertices of Λ_n that are on the bottom boundary are inside; vertices on the top boundary are outside. Therefore, in order to have a red vertical bridge, there must be a simple red path going across C . That is to say, a red vertical bridge must cross the green horizontal bridge.

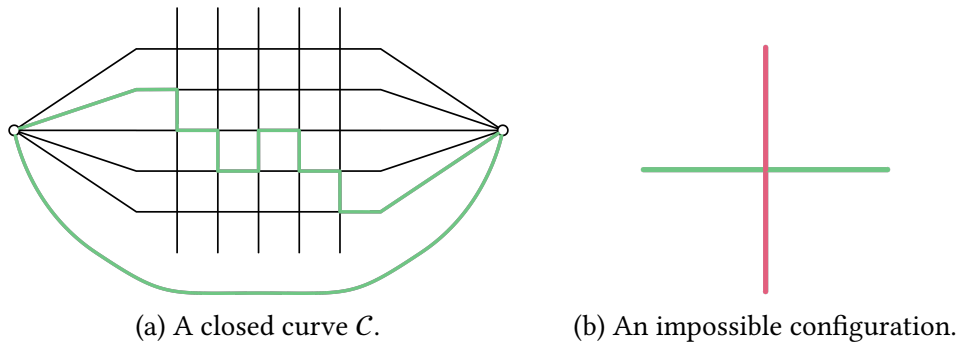


Figure 5.3: A topological argument.

However, this is impossible. Clearly, being of different colors, a red bridge and a green bridge cannot share any edge. Since the local configuration shown in Figure 5.3b means that the four edges incident to a vertex (i, j) on Λ_n are all pointing inwards (when $i + j$ is even) or all pointing outwards (when $i + j$ is odd), it is not allowed in any valid states of six-vertex configurations.

Similarly, the local configuration of a vertex surrounded by two red horizontal edges and two green vertical edges (a 90 degree rotation of Figure 5.3b) is also not allowed. $\square$

Next we characterize the states in $\Omega \setminus (C_G \cup C_R)$. Define a shifted lattice^{††} L to be $\mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2})$ where two points (a, b) and (c, d) in L are neighbors if $|a - c| = |b - d| = 1$ (a, b, c , and d are all half integers), i.e., they are at the center of a square in $\mathbb{Z}^2$ and are connected by “diagonal” edges of length $\sqrt{2}$. An example of L and its relationship with $\mathbb{Z}^2$ is shown in Figure 5.4a. L is not connected — it is composed of two sub-lattices L_0 and L_1 (depicted with different colors in Figure 5.4b). Denote by L_n the restriction of L on the finite region inside Λ_n . Note that in graph theoretical terms, the square lattice Λ_n is planar and 4-regular, and thus can be seen as the *medial graph* of two planar graphs. In fact, they are L_0 and L_1 (restricted onto L_n).

From now on, we use Λ_n -vertices/edges as an abbreviation for vertices/edges in Λ_n ; and we use L_n -vertices/edges and other similar notations whenever it has a clear meaning in the context. For any state $\tau \in \Omega$, there is a subset L_τ of L_n -edges associated with τ . Each L_n -edge e “goes diagonally through” exactly one Λ_n -vertex, denoted by v_e . We say e is in L_τ if and only if the four Λ_n -edges incident to v_e are 2-green-2-red and e separates the two green edges from the two red edges (Remember that in this case edges in the same color must be rotationally adjacent to each other). See Figure 5.4c for an instance of a state τ and its associated L_τ . In the following we abuse the notation and use L_τ as its induced subgraph of L . This view was adopted by [BKW73] for establishing the existence of the spontaneous staggered polarization in the anti-ferroelectric phase of the six-vertex model.

For any $\tau \in \Omega$ and L_τ , we make the following observations:

- There is always an even number of L_τ -edges meeting at any L_n -vertex, except for the L_n -vertices on the boundary. Because this number is equal to the number of times for the color change on the four Λ_n -edges surrounding the L_n -vertex, if we start from any one of the four Λ_n -edges and go rotationally over the four Λ_n -edges, which is even.

^{††}Strictly speaking, a lattice is a discrete subgroup of $\mathbb{R}^n$. A shifted copy of a lattice does not contain 0.

- For any Λ_n vertex, there can be at most one L_τ -edge going through which is either in L_0 or in L_1 .
- If $\bar{\tau} \in \Omega$ is the state by a total edge reversal of τ , then $L_{\bar{\tau}} = L_\tau$.

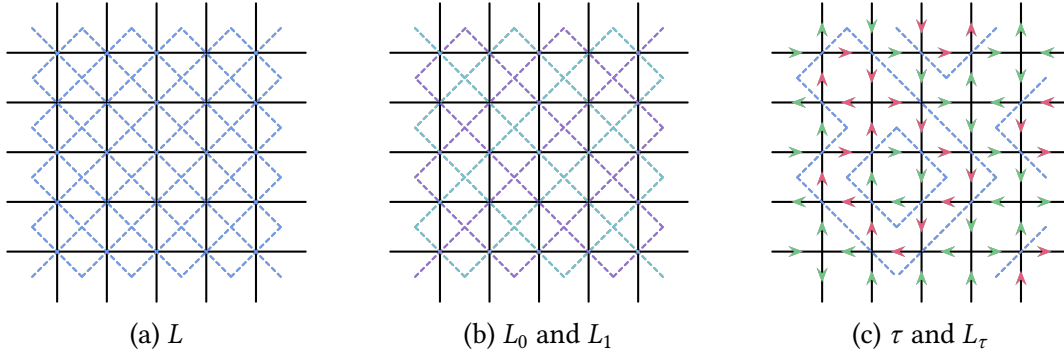


Figure 5.4: An auxiliary diagonal lattice.

For a state $\tau \in \Omega$, we say τ has a *horizontal (or vertical) fault line* if there is a self-avoiding path in L_τ connecting a L_n -vertex on the left (top, respectively) boundary of L_n to a L_n -vertex on the right (bottom, respectively) boundary of L_n . See Figure 5.5c for an example where a state has both a horizontal fault line and a vertical fault line. Denote by C_{FL} the set of states containing a horizontal fault line or a vertical fault line. Since a fault line separates green edges from red edges, a vertical (horizontal) fault line precludes any horizontal (vertical, respectively) monochromatic bridge (the proof is basically the same as Lemma 5.2.2). This is to say, C_G , C_{FL} , and C_R are pairwise disjoint. Next we show the following lemma and its direct implication (Corollary 5.2.4).

Lemma 5.2.3. *If in a state τ there is no monochromatic cross, then there is a fault line.*

Proof. If there is no monochromatic cross (i.e., a green cross or a red cross) in τ , we can assume that

$$\text{there is no green horizontal bridge and there is no red horizontal bridge.} \quad (*)$$

- Suppose τ has a green horizontal bridge. There is no red vertical bridge since it cannot “cross” the green horizontal bridge; there is no green vertical bridge since there is no green cross. Therefore, this case is symmetric to $(*)$, switching horizontal for vertical.

- Suppose τ has a red horizontal bridge. This case is similar to the above case, and thus is also symmetric to (*).

Next we show there is a vertical fault line if there is no monochromatic horizontal bridge. We introduce another graph M_n that is the *medial graph* of Λ_n where every vertex of M_n corresponds to an edge of Λ_n , i.e., two M_n -vertices are neighboring if the two corresponding Λ_n -edges are rotationally adjacent. Note that M_n is part of another shifted square lattice. An example of M_n and its relationship with $\mathbb{Z}^2$ is shown in Figure 5.5a. For any state $\tau \in \Omega$, there is a subset M_τ of M_n -edges associated with τ . An M_n -edge e is in M_τ if the two vertices that e is incident to (as Λ_n -edges) have the same color (both green or both red). See Figure 5.5d for an example.

Observe that the correspondence between Λ_n -edges and M_n -vertices translates into the correspondence between simple monochromatic paths in Λ_n to simple connected paths in M_τ . In fact, the connected components in M_τ , capturing the notion of monochromatic regions of Λ_n -edges, and connected components in L_τ , capturing the notion of separation between regions of Λ_n -edges of different colors, are in a *dual* relationship.

This duality is depicted in Figure 5.5e and helps us find a fault line. Let V_M be the collection of M_n -vertices that can be reached from the left boundary of M_n by a simple path in M_τ . Since there is no monochromatic horizontal bridge, V_M does not contain any M_n -vertex on the right boundary. As a consequence, there is a *cutset* in M_n separating V_M from the right boundary. This cutset, composed of M_n -edges, corresponds to a vertical fault line. For instance, in Figure 5.5e the blue solid L_τ -path γ is a fault line defined by the above argument. $\square$

Corollary 5.2.4. $\Omega = C_G \cup C_{FL} \cup C_R$ is a partition of the state space.

Before moving on to prove Theorem 5.2.1, we introduce the notion of *almost fault lines*. A *horizontal (or vertical) almost fault line* is a self-avoiding L_n -path connecting a L_n -vertex on the left boundary of L_n to a L_n -vertex on the right boundary of L_n where all edges except for one are in L_τ . Denote by C_{AFL} the set of states containing an almost fault line. Let ∂C_G be the set of states outside C_G which are one-flip away from C_G in the state space of $\mathcal{M}_G$.

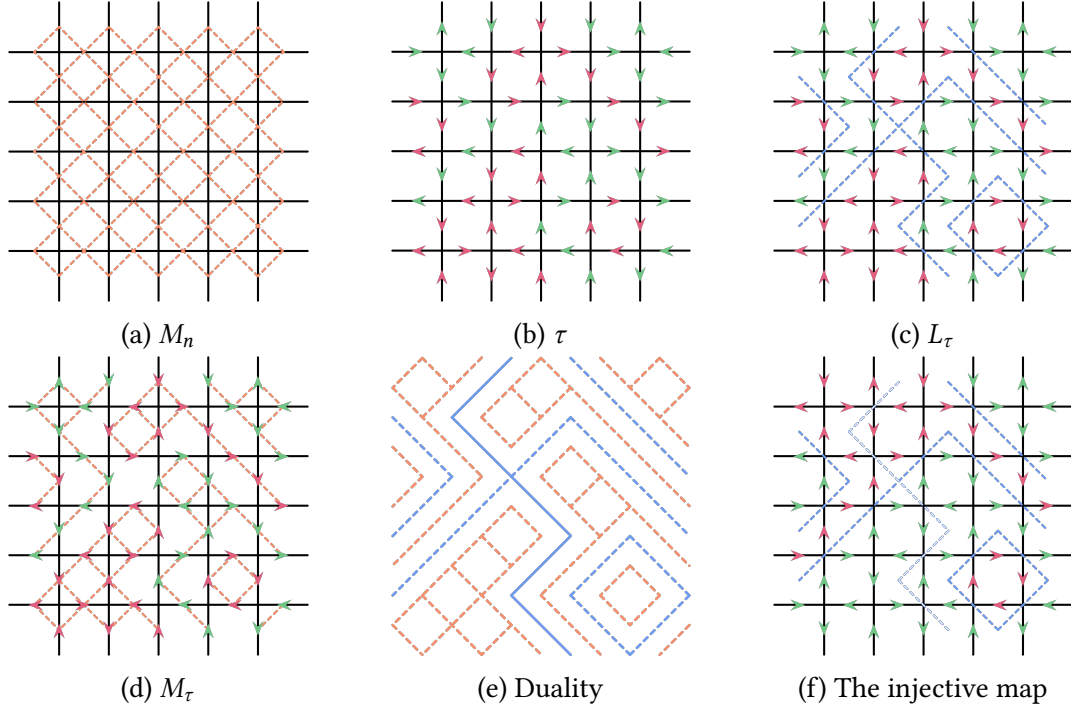


Figure 5.5: A fault line argument.

Lemma 5.2.5. $\partial C_G \subset C_{\text{FL}} \cup C_{\text{AFL}}$.

Proof. If a state $\tau_\partial \in \partial C_G$ is not in C_{FL} (does not have a fault line), then by Corollary 5.2.4 $\tau_\partial \in C_R$ since by definition it is outside of C_G . Because τ_∂ is one move away from C_G , there exists a state $\tau_G \in C_G$ such that flipping four monochromatic edges along a unit square s on $\mathbb{Z}^2$ yields τ_∂ . We know that in τ_G there exists a green cross and no red cross; in τ_∂ there exists a red cross and no green cross. Then it must be true that the four edges along s are all green in τ_G and all red in τ_∂ . See Figure 5.6 for a pictorial illustration. Moreover, any green cross in τ_G must contain edges along s and so is any red cross in τ_∂ ; otherwise, a green horizontal (vertical) bridge in τ_G must go across a red vertical (or horizontal, respectively) bridge in τ_∂ at some other place on Λ_n , which is impossible.

Therefore, there exists a simple green path Γ_G from some vertices on s to the top boundary of Λ_n and a simple red path Γ_R from some vertices on s to the top boundary of Λ_n . In the following, we prove that the above conditions suffice to show that there exists an L_τ -path from s to the top boundary of L_n . Similar conclusions can be made for the existence of L_τ -paths from s to the

bottom boundary of L_n . By adding at most one L_n -edge, we can concatenate these paths to obtain a vertical almost fault line.

Γ_G and Γ_R cannot cross each other (Figure 5.3). Without loss of generality, suppose Γ_G is to the left of Γ_R . Then we use the medial lattice view M_{τ_∂} as in the proof of Lemma 5.2.3. Γ_G corresponds to a connected component m_G in M_{τ_∂} . Denote by V_G the set of M_n -vertices which can be reached by m_G in M_{τ_∂} , and let V'_G be V_G together with the set of M_n -vertices which are separated from the right boundary of M_n by V_G . Then there is a *cutset* in M_n separating V'_G from the right boundary. This cutset, composed of M_n -edges, corresponds to a dual L_τ -path from s to the top boundary of L_n . $\square$



Figure 5.6: A step in $\mathcal{M}_G$.

Now we are ready to show Theorem 5.2.1. Let $\Omega_{\text{LEFT}} = C_G$, $\Omega_{\text{MIDDLE}} = C_{\text{FL}} \cup (C_{\text{AFL}} \cap C_R)$, and $\Omega_{\text{RIGHT}} = C_R \setminus C_{\text{AFL}}$. Theorem 5.2.1 is a consequence of the following lemma which uses a *Peierls argument* to show that $\pi(C_{\text{FL}} \cup C_{\text{AFL}})$ is exponentially small.

Lemma 5.2.6. $\pi(C_{\text{FL}} \cup C_{\text{AFL}}) \leq O(n) \left(\frac{2.639 \max(a,b)}{c} \right)^n$.

Proof. For a self-avoiding path γ in L_n connecting a vertex on the top boundary to a vertex on the bottom boundary, denote by F_γ the set of states in Ω that contain γ as vertical fault line or almost fault line. Reversing directions of all the edges to the left side of γ defines an injective mapping from F_γ to $\Omega \setminus F_\gamma$ that magnifies probability by a factor of at least $\frac{\min(a,b)}{c} \cdot \left(\frac{c}{\max(a,b)} \right)^{|\gamma|-1}$. This is because: if γ is a fault line in a state, every Λ_n -vertex sitting on γ would have four incident edges in the same color after the map, which increase its weight to c (from a or b); if γ is an almost

fault line in a state, the above is true except that for one Λ_n -vertex, its weight decrease from c to a or b after the map, as orientations on half of the four monochromatic edges are reversed. This indicates that $\pi(F_\gamma) \leq \frac{c}{\min(a,b)} \cdot \left(\frac{\max(a,b)}{c}\right)^{|\gamma|-1}$. See Figure 5.5f for an example. The same goes for horizontal (almost) fault lines.

Since every (almost) fault line is a self-avoiding walk on L , the number of fault lines of length l is upper bounded by $2n$ times the number of self-avoiding walks of that length starting at a vertex on the left or bottom boundary. The latter can be bounded by an well-studied estimate μ^l on the number of self-avoiding walks of length l on $\mathbb{Z}^2$, where μ is called the *connective constant*. The best proved bound is $\mu \approx 2.638158 \dots$ [GC01]. Summing this over fault lines of length from n to n^2 completes the proof. $\square$

We have proved the torpid mixing of Glauber dynamics for the six-vertex model on the lattice region Λ_n with free boundary conditions. Next we state the idea to extend the proof for the case when the Markov chain is $\mathcal{M}_D$ and the case when the boundary of $\tilde{\Lambda}_n$ is periodic. Theorem 5.0.2 is a combination of Theorem 5.2.1 and the extensions.

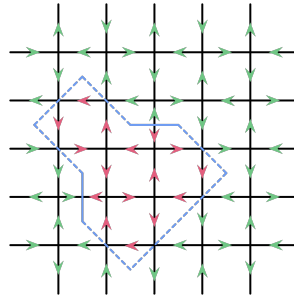


Figure 5.7: A state in $\mathcal{M}_D$.

To extend Theorem 5.2.1 to hold for the directed-loop algorithm $\mathcal{M}_D$ whose state space is $\Omega \cup \Omega'$, we need to pay extra attention for the states in Ω' , the near-perfect Eulerian orientations. For a state $\tau' \in \Omega'$, there are two “defects” on the edges (Figure 5.7). Apart from the diagonal L_n -edges, there are two $(\mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2}))$ -edges separating green (half-)edges from red (half-)edges. The adaption we make is to put such $(\mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2}))$ -edges also into the set $L_{\tau'}$. Notice that a connected component in $L_{\tau'}$ could possibly lie on L_0 -edges as well as L_1 -edges. Everything we prove is still

correct if we allow fault lines to have $(\mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2}))$ -edges. Due to the possible positions of such two defects, the weight of $C_{\text{FL}} \cup C_{\text{AFL}}$ only increases by a polynomial factor in n , hence not affecting the fact of torpid mixing.

To extend Theorem 5.2.1 to hold for $\widetilde{\Lambda}_n$ with periodic boundary condition (i.e., a 2-dimensional torus), we make the following modification. When n is even, there still are two states with maximum weights c^{n^2} (similar to τ_{G} and τ_{R} in Figure 5.2). Again, for any state τ , $\widetilde{\Lambda}_n$ -edges can be classified as green or red, and its associated L_τ separates $\widetilde{\Lambda}_n$ -edges of different colors. For the 2-dimensional torus T^2 , the homology group $H_1(T^2) \cong \mathbb{Z} \times \mathbb{Z}$. We say a state τ has a *green (or red) cross* if there are two *non-contractable* cycles of green (or red, respectively) edges of *homology classes* (a_1, b_1) and (a_2, b_2) with $\det \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix} \neq 0$; τ has a *pair of fault lines* if there is a pair of non-contractable cycles of L_τ -edges. (By parity, if there is one L_τ -cycle there must be two.) Then the proofs in this section can be naturally adapted for the torus case, and the torpid mixing result follows.

5.3 The even-coloring model revisited

Assuming the underlying graph is a finite region of the square lattice $\mathbb{Z}^2$ or a torus $(\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$, we can set up several one-to-one onto mappings between S_{8V} and S_{EC} (with m, n being even for M_{III} and M_{IV}).

M_{I} Given a vertex, color an incident edge e green if the orientation on e is the same as in Figure 1.1-1 and red otherwise. Every local orientations in S_{8V} differs from Figure 1.1-1 at an even number of edges, and thus is mapped to a local coloring in S_{EC} . This color assignment induces a one-to-one onto mapping from S_{8V} to S_{EC} with $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (7, 8, 5, 6, 3, 4, 1, 2)$. Under this mapping, $Z_{8V}(a, b, c, d) = Z_{\text{EC}}(d, c, b, a)$.

M_{II} Given a vertex, color an incident edge e green if the orientation on e is the same as in Figure 1.1-3 and red otherwise. This induces a one-to-one onto mapping from S_{8V} to S_{EC}

with $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (5, 6, 7, 8, 1, 2, 3, 4)$. Under this mapping, we have $Z_{8V}(a, b, c, d) = Z_{EC}(c, d, a, b)$.

M_{III} Pick a vertex (i_0, j_0) on the square lattice to be an odd vertex, and call a vertex (i, j) to be *odd* if $(i - i_0) + (j - j_0) \equiv 0 \pmod{2}$ and *even* otherwise^{‡‡}. Given an odd (or even) vertex, color an incident edge e green if the orientation on e is the same as in Figure 1.1-5 (or Figure 1.1-6, respectively) and red otherwise. This color assignment is well-defined, because in any orientation an edge has to either agree with Figure 1.1-5 as an edge incident to the odd vertex and agree with Figure 1.1-6 as an edge incident to the even vertex at the same time or disagree with both at the same time. Therefore, it induces a one-to-one onto mapping from S_{8V} to S_{EC} with $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (3, 4, 1, 2, 7, 8, 5, 6)$ for the odd vertices and $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (4, 3, 2, 1, 8, 7, 6, 5)$ for the even vertices. Under this mapping, $Z_{8V}(a, b, c, d) = Z_{EC}(b, a, d, c)$.

M_{IV} Again classify the vertices into *odd* vertices and *even* vertices based on their positions on the square lattice. Given an odd (or even) vertex, color an incident edge e green if the orientation on e is the same as in Figure 1.1-7 (or Figure 1.1-8, respectively) and red otherwise. This induces a one-to-one onto mapping from S_{8V} to S_{EC} with $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (1, 2, 3, 4, 5, 6, 7, 8)$ for the odd vertices and $(1, 2, 3, 4, 5, 6, 7, 8) \rightarrow (2, 1, 4, 3, 6, 5, 8, 7)$ for the even vertices. Under this mapping, $Z_{8V}(a, b, c, d) = Z_{EC}(a, b, c, d)$.

The above mappings provide a unified way to study the eight-vertex model in the order phases when one of a, b, c, d dominates — convert the eight-vertex model in ferroelectric phases with a (or b) large into the even-coloring model with ψ large using M_I (or M_{III} , respectively); convert the eight-vertex model in anti-ferroelectric phases with c (or d) large into the even-coloring model with ψ large using M_{III} (or M_{IV} , respectively). We note that when ψ is large, i.e., the local configuration with all edges of the same color has dominating weight, the system favors

^{‡‡}This is why we need m and n to be even for torus $(\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$.

states with large monochromatic area. In other words, our mappings unify the ferroelectric and anti-ferroelectric eight-vertex models into a ferromagnetic-Ising-like model.

In addition to the relationship between partition functions of the two models, the above mappings also relate the behaviors of local Markov chains for them.

First we describe *Glauber dynamics for the eight-vertex model* on square regions of $\mathbb{Z}^2$, $\mathcal{M}_{8V}$. The state space of this Markov chain is the set of all even orientations $\mathcal{O}_e$. In each step, the chain uniformly chooses a cycle of length four in the lattice (i.e. a face) and either reverses the orientations of all the edges around that cycle or keeps them all unchanged, according to the correct conditional probabilities dictated by the *Gibbs distribution*. We note that if the change happens, the local configuration at any of the four vertices around the cycle is still in S_{8V} as arrows on exactly two incident edges are flipped. $\mathcal{M}_{8V}$ is *irreducible, aperiodic*, and satisfies the *detailed balance condition* even when the sink (Figure 1.1-7) and source (Figure 1.1-8) configurations are disallowed [GMP04]. It was shown that $\mathcal{M}_{8V}$ mix torpidly on square regions of $\mathbb{Z}^2$ with free boundary conditions under parameter settings (a, b, c, d) when (1) $a = b = c = 1, d > 384$ [GR10] or (2) $d = 0, c \geq 2.639 \max(a, b)$ (in Section 5.2 or [Liu18]) or (3) $d = 0, ac + bc + 3ab < c^2$ [FR19]. These regions are all in the ordered phases.

Here we introduce *Glauber dynamics for the even-coloring model* on square regions of $\mathbb{Z}^2$, $\mathcal{M}_{EC}$. The state space of this Markov chain is the set of all even-colorings $\mathcal{C}_e$. In each step, the chain uniformly chooses a cycle of length four in the lattice (i.e. a face) and either changes the colors of all the edges around that cycle or keeps them all unchanged, according to the correct conditional probabilities dictated by the *Gibbs distribution*. We note that if the change happens, the local configuration at any of the four vertices around the cycle is still in S_{EC} as colors on exactly two incident edges are flipped.

Using $\mathcal{M}_{EC}$ as an intermediate Markov chain, we can correlate the behaviors of $\mathcal{M}_{8V}$ under different parameter settings.

Theorem 5.3.1. *$\mathcal{M}_{8V}$ is rapidly (or torpidly) mixing on square regions of $\mathbb{Z}^2$ with free boundary conditions under parameter setting (a, b, c, d) if and only if it is rapidly (or torpidly, respectively)*

mixing under parameter setting (b, a, d, c) . The statement is also true when we replace (b, a, d, c) by (c, d, a, b) , or by (d, c, b, a) .

Proof. On the one hand, based on the one-to-one weight-preserving mapping M_I (M_{II} , or M_{III}) introduced earlier, there is an equivalence between the state spaces as well as transitions of $\mathcal{M}_{8V}$ with parameters (a, b, c, d) and $\mathcal{M}_{EC}$ with parameters (d, c, b, a) ((c, d, a, b) , or (b, a, d, c) , respectively).

On the other hand, based on the one-to-one weight-preserving mapping M_{IV} introduced earlier, there is an equivalence between the state spaces as well as transitions of $\mathcal{M}_{8V}$ with parameters (a, b, c, d) and $\mathcal{M}_{EC}$ with parameters (a, b, c, d) .

The following diagram illustrates the mappings described above:

$$\mathcal{M}_{8V}(a, b, c, d) \xrightarrow{M_I} \mathcal{M}_{EC}(d, c, b, a) \xrightarrow{M_{IV}^{-1}} \mathcal{M}_{8V}(d, c, b, a).$$

□

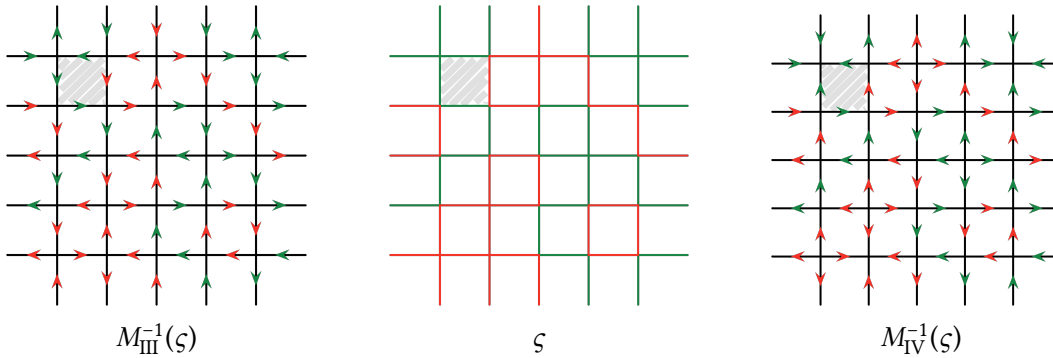


Figure 5.8: A state ζ in the even-coloring model from two different states in the eight-vertex model under M_{III} and M_{IV} , respectively.

Here we give an example (Figure 5.8) to illustrate the correlation between $\mathcal{M}_{8V}$ when c is large and $\mathcal{M}_{8V}$ when d is large via $\mathcal{M}_{EC}$. In Figure 5.8, assuming the top leftmost degree-4 vertex is indexed as an odd vertex, two different even orientations are mapped to the same even coloring ζ using M_{III} and M_{IV} , respectively. Thus we denote them by $M_{III}^{-1}(\zeta)$ and $M_{IV}^{-1}(\zeta)$. In order to demonstrate the equivalence described in Theorem 5.3.1, we assume $M_{III}^{-1}(\zeta)$ is from the eight-vertex

model under parameter setting $(a_1, b_1, c_1, d_1) = (\nu, \mu, \psi, \phi)$, and $M_{\text{IV}}^{-1}(\zeta)$ is from the eight-vertex model under parameter setting $(a_2, b_2, c_2, d_2) = (\mu, \nu, \phi, \psi)$. Under such settings, the two eight-vertex models both correspond to the even-coloring model under parameter setting (μ, ν, ϕ, ψ) . One can check that $M_{\text{III}}^{-1}(\zeta)$ and $M_{\text{IV}}^{-1}(\zeta)$ have the same weights (as they both correspond to the same state ζ in the even-coloring model). Suppose $\mathcal{M}_{8V}$ is about to flip the marked face in Figure 5.8. Prior to the flip, enumerating clockwise starting from the top left vertex, the four vertices in $M_{\text{III}}^{-1}(\zeta)$ are in local configurations Figure 1.1-5, Figure 1.1-2, Figure 1.1-3 and Figure 1.1-3, with weights c_1, a_1, b_1 , and b_1 , respectively; after the flip, the four vertices are in local configurations Figure 1.1-1, Figure 1.1-7, Figure 1.1-8, Figure 1.1-5, with weights a_1, d_1, d_1, c_1 , respectively. At the same time in $M_{\text{IV}}^{-1}(\zeta)$, prior to the flip, the four vertices are in Figure 1.1-7, Figure 1.1-4, Figure 1.1-1, Figure 1.1-1 with weights d_2, b_2, a_2, a_2 ; after the flip, they are in Figure 1.1-3, Figure 1.1-5, Figure 1.1-6, Figure 1.1-7 with weights b_2, c_2, c_2, d_2 . Since $a_1 = b_2 = \nu$, $b_1 = a_2 = \mu$, $c_1 = d_2 = \psi$, and $d_1 = c_2 = \phi$, the weights of both sides equal to each other both before and after the flip. Therefore, the Markov chains will have the same behavior given the same random bits.

The following corollary follows from the results in Section 5.2, [GR10, Liu18, FR19], and Theorem 5.3.1.

Corollary 5.3.2. *$\mathcal{M}_{8V}$ mix torpidly on square regions of $\mathbb{Z}^2$ with free boundary conditions under parameter settings (a, b, c, d) when*

- $b = c = d = 1, a > 384$, or $b = 0, a \geq 2.639 \max(c, d)$, or $b = 0, ac + ad + 3cd < a^2$.
- $a = c = d = 1, b > 384$, or $a = 0, b \geq 2.639 \max(c, d)$, or $a = 0, bc + bd + 3cd < b^2$.
- $a = b = d = 1, c > 384$, or $d = 0, c \geq 2.639 \max(a, b)$, or $d = 0, ac + bc + 3ab < c^2$.
- $a = b = c = 1, d > 384$, or $c = 0, d \geq 2.639 \max(a, b)$, or $c = 0, ad + bd + 3ab < d^2$.

Remark 5.3.1. We note that these regions are all in the ordered phases.

5.4 An abstract polymer model

The cluster expansion is a powerful tool in statistical mechanics which provides a method for studying the logarithm of a partition function, when the model under consideration has a partition function that can be written in a particular form [FV17]. In particular, cluster expansion is developed for models that can be described in terms of geometrical objects called polymers (rather than the original microscopic components) such that

1. each configuration was set in one-to-one correspondence with a family of *pairwise disjoint* polymers;
2. once expressed in terms of polymers, the Boltzmann weight split into a *product of weights* associated to the polymers.

Here is the formal definition of polymer models.

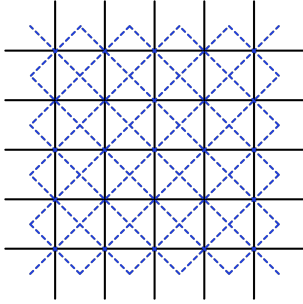
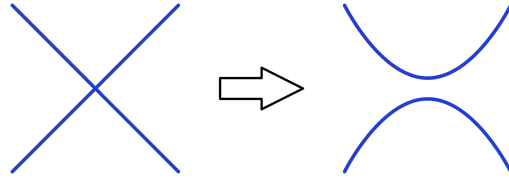
Definition 5.4.1. Let $G = (V, E)$ be a finite graph. Define a *polymer* γ in G to be a connected subgraph of G . A *polymer model* consists of a set $\mathcal{P}(G)$ of polymers along with *weight functions* $w(\gamma, \cdot) : \mathbb{C} \rightarrow \mathbb{C}$ for each polymer γ .

Let $\mathcal{C}(G)$ be the set of all finite sets of polymers from $\mathcal{P}(G)$ that are *pairwise compatible* (whose definition depends on the specific model), including the empty set of polymers. The (*polymer*) *partition function* is defined by

$$Z_{\text{POLYMER}}(G; z) = \sum_{\Gamma \in \mathcal{C}(G)} \prod_{\gamma \in \Gamma} w(\gamma, z) \quad (5.4.1)$$

where the term corresponding to the empty set of polymers $\Gamma = \emptyset$ is 1 by convention. We think of $Z_{\text{POLYMER}}(G; z)$ as a function of one complex variable z .

Next we define the polymer model representation of the even-coloring model on the square lattice with free boundary conditions. The polymer models corresponding to the even-coloring model on a torus of finite side length will be introduced in Section 5.7.

Figure 5.9: $\mathbb{Z}^2$ and L .Figure 5.10: The deformation rule at $L \setminus \mathbb{Z}^2$ -vertices.

Define a lattice L to be $\mathbb{Z}^2 \cup (\mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2}))$ where two points (x_1, y_1) and (x_2, y_2) in L are neighbors if $|x_1 - x_2| = |y_1 - y_2| = \frac{1}{2}$. In other words, the edges of L are the half-diagonals of squares in $\mathbb{Z}^2$. An example of L and its relationship with $\mathbb{Z}^2$ is shown in Figure 5.9. Given Λ , a finite (simply connected) region on the square lattice with the free boundary condition, denote by $\mathbb{Z}_\Lambda^2$ (or L_Λ) the restriction to $\mathbb{Z}^2$ (or L , respectively) on the finite region inside Λ . The polymers are the set of connected subgraphs of L_Λ in which each vertex has degree two or four (except for boundary vertices on $L \setminus \mathbb{Z}^2$ which have degree one or two), with the following further restrictions:

- At each vertex of $\mathbb{Z}^2$, if a polymer has degree two, the two half-diagonals must be in diagonally opposite quadrants. (Figure 5.12 shows all the valid configurations at a $\mathbb{Z}^2$ -vertex.)
- At each vertex of $L \setminus \mathbb{Z}^2$, if a polymer has degree four, we virtually group the four half-diagonals into two pairs such that the upper two are detached from the lower two.
- Two polymers are compatible if they are vertex-disjoint, or at each $L \setminus \mathbb{Z}^2$ -vertex they share, one polymer has two half-diagonals in the upper two quadrants and the other has two in the lower two quadrants.

Note that by definition polymers can have degree-4 vertices only at $\mathbb{Z}^2$ -vertices. We define the size $|\gamma|$ of a polymer to be the number of $\mathbb{Z}_\Lambda^2$ -vertices it goes through as a connected component.

Now we describe the mapping from the even-coloring model to the above polymer model. Given an even coloring ς , consider the four quadrants adjacent to a $\mathbb{Z}^2$ -vertex. As the lattice edges

of L are half-diagonals each connecting a $\mathbb{Z}^2$ -vertex to an $L \setminus \mathbb{Z}^2$ -vertex, for each half-diagonal there are two edges on the $\mathbb{Z}^2$ lattice that are neighboring to it and share the same $\mathbb{Z}^2$ -vertex. Define $B(\zeta)$ to be the set of half-diagonals whose two neighboring edges have opposite colors. In order to decompose $B(\zeta)$ into a set Γ of compatible polymers, we use a *deformation rule* on $L \setminus \mathbb{Z}^2$ -vertices. To this end, we first remark that the number of half-diagonals in $B(\zeta)$ meeting at any interior vertex of L_Λ is always 0, 2, or 4. When this number is 4 and the vertex is on $L \setminus \mathbb{Z}^2$, we deform $B(\zeta)$ using the rule in Figure 5.10. See Figure 5.15 for an example.

We briefly argue that the above mapping is an exact *two-to-one onto correspondence* between $C_e(\Lambda)$, the even colorings on Λ , and $C(L_\Lambda)$, compatible subsets of polymers in L_Λ . It is “two-to-one” because two even colorings having totally opposite color configurations are mapped to the same polymer configuration. As a special case, the empty polymer set corresponds to both the all-green and the all-red colorings. To find one of the two even-colorings given a compatible subset of polymers, we fix an arbitrary edge to be green, then give the same color to all edges cyclically adjacent to it unless a polymer edge (on the L lattice) is crossed (in which case we switch the color), and follow the above rule to propagate colors until the boundary is reached.

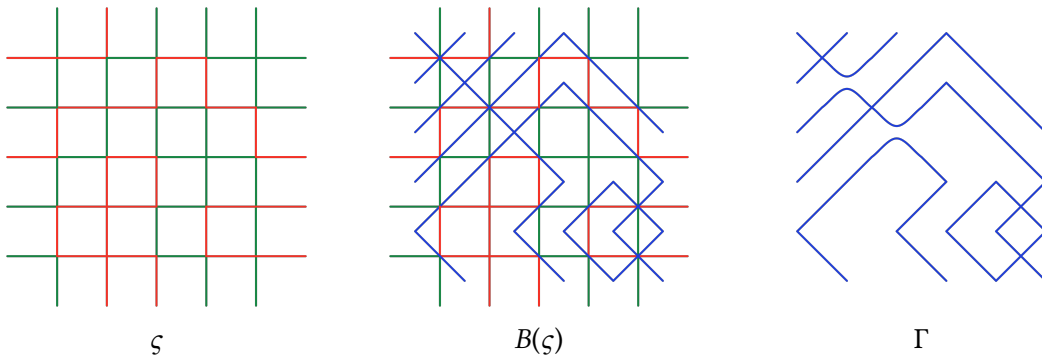


Figure 5.11: An example of the correspondence between an even coloring and compatible polymers.

In order to set up the equivalence between Z_{EC} and $Z_{POLYMER}$, we need to choose suitable weights for the polymers. To do this, the weight function $w(\gamma, z)$ is defined to be the product of the weights on all $\mathbb{Z}_\Lambda^2$ -vertices that γ goes through, where the weights for the local configurations 1, 2, and 3 in Figure 5.12 are $\frac{\mu}{\psi}$, $\frac{\nu}{\psi}$, and $\frac{\phi}{\psi}$, respectively. We note that μ, ν, ϕ, ψ are parameters of the even-coloring

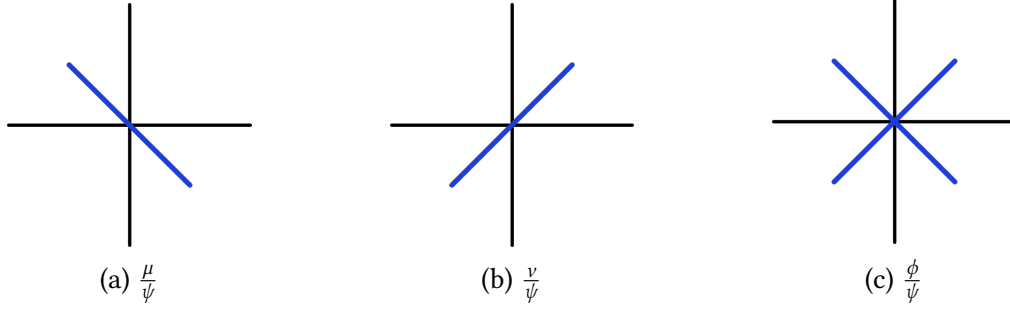


Figure 5.12: The three valid local configurations of a polymer at $\mathbb{Z}^2$ -vertices.

model as in (1.2.1). Here we simplify the notation in (1.2.1) and define

$$\epsilon_\mu = \epsilon'_1 = \epsilon'_2, \quad \epsilon_\nu = \epsilon'_3 = \epsilon'_4, \quad \epsilon_\phi = \epsilon'_5 = \epsilon'_6, \quad \epsilon_\psi = \epsilon'_7 = \epsilon'_8.$$

Under the assumption that they are all rational numbers, so are the differences $\epsilon_\mu - \epsilon_\psi$, $\epsilon_\nu - \epsilon_\psi$, and $\epsilon_\phi - \epsilon_\psi$, denoted by $\frac{p_1}{q}$, $\frac{p_2}{q}$, and $\frac{p_3}{q}$, respectively, where p_1, p_2, p_3, q are positive integers and $q = \text{lcm}(p_1, p_2, p_3)$. If we let $z = e^{-\beta/q}$ then the weights on polymers are monomials in the variable z . We note that z is viewed as a complex variable in order to apply Theorem 5.5.1 and Theorem 5.6.2. Thus we can write

$$\begin{aligned} w(\gamma, z) &= \left(\frac{\mu}{\psi} \right)^{n'_1} \left(\frac{\nu}{\psi} \right)^{n'_2} \left(\frac{\phi}{\psi} \right)^{n'_3} \\ &= e^{-\beta[n'_1(\epsilon_\mu - \epsilon_\psi) + n'_2(\epsilon_\nu - \epsilon_\psi) + n'_3(\epsilon_\phi - \epsilon_\psi)]} \\ &= z^{n'_1 p_1 + n'_2 p_2 + n'_3 p_3} \end{aligned}$$

where n'_i ($i = 1, 2, 3$) is the number of local configurations of type i shown in Figure 5.12. In fact, the local configuration Figure 5.12a in the polymer model corresponds to the local configurations Figure 1.3-1 and Figure 1.3-2 in the even-coloring model; Figure 5.12b corresponds to Figure 1.3-3 and Figure 1.3-4; Figure 5.12c corresponds to Figure 1.3-5 and Figure 1.3-6.

Therefore, we can write the partition function of the even-coloring model on Λ as

$$\begin{aligned}
Z_{\text{EC}}(\Lambda; \mu, \nu, \phi, \psi) &= \sum_{\zeta \in C_e(\Lambda)} \mu^{n_1+n_2} \nu^{n_3+n_4} \phi^{n_5+n_6} \psi^{n_7+n_8} \\
&= \psi^{|\mathbb{Z}_\Lambda^2|} \sum_{\zeta \in C_e(\Lambda)} \left(\frac{\mu}{\psi}\right)^{n_1+n_2} \left(\frac{\nu}{\psi}\right)^{n_3+n_4} \left(\frac{\phi}{\psi}\right)^{n_5+n_6} \\
&= 2\psi^{|\mathbb{Z}_\Lambda^2|} \sum_{\Gamma \in C(L_\Lambda)} \prod_{\gamma \in \Gamma} w(\gamma, z) \\
&= 2\psi^{|\mathbb{Z}_\Lambda^2|} Z_{\text{POLYMER}}(L_\Lambda; z).
\end{aligned}$$

The factor of 2 comes from the two-to-one correspondence from $C_e(\Lambda)$ to $C(L_\Lambda)$.

5.5 Convergence of the cluster expansion

In order to apply the algorithm in Section 5.6, we have to find a disc centered at the origin such that the polymer partition function $Z_{\text{POLYMER}}(G; z)$ is not zero for any G and any z inside the disc. This can be proved by showing $\log Z_{\text{POLYMER}}(G; z)$ is not infinity for any G and any z inside the disc. The cluster expansion provides an explicit expansion for $\log Z_{\text{POLYMER}}$, in the form of a series [FV17, HPR19].

Define the *Ursell function* φ on ordered families of polymers $(\gamma_1, \dots, \gamma_m)$ by $\varphi(\gamma_1) := 1$, when $m = 1$ and

$$\varphi(\gamma_1, \dots, \gamma_m) := \frac{1}{m!} \sum_{\substack{G \in K_m \\ \text{connected}}} \prod_{\{i,j\} \in G} \zeta(\gamma_i, \gamma_j),$$

when $m \geq 2$, where K_m is the complete graph on m vertices, and $\zeta(\gamma_i, \gamma_j) = 0$ if γ_i and γ_j are compatible and -1 otherwise. Then the cluster expansion gives the following

$$\log Z_{\text{POLYMER}}(G; z) = \sum_{m \geq 1} \sum_{\gamma_1} \cdots \sum_{\gamma_m} \varphi(\gamma_1, \dots, \gamma_m) \prod_{i=1}^m w(\gamma_i, z). \quad (5.5.1)$$

Under suitable conditions, (5.5.1) is an absolutely convergent power series representation.

Theorem 5.5.1 (Kotecký and Preiss [KP86]). *Assume that there exists $h : \mathcal{P}(G) \rightarrow \mathbb{R}_{>0}$ such that, for each $\gamma^* \in \mathcal{P}(G)$,*

$$\sum_{\gamma \nabla \gamma^*} |w(\gamma, z)| e^{h(\gamma)} \leq h(\gamma^*) \quad (5.5.2)$$

where the sum is over all polymers γ incompatible with γ^* . Then the cluster expansion for $Z_{\text{POLYMER}}(G; z)$ in (5.5.1) converges absolutely. In particular, $Z_{\text{POLYMER}}(G; z) \neq 0$.

Now our task becomes the following: given $\epsilon_\mu, \epsilon_\nu, \epsilon_\phi, \epsilon_\psi$, find a suitable $h(\cdot)$, and the range of values for z such that (5.5.2) holds. The following two theorems in complex analysis will be used.

Theorem 5.5.2 (Cauchy–Hadamard Theorem). *Consider the formal power series in one complex variable z of the form*

$$f(z) = \sum_{n=0}^{\infty} c_n (z - a)^n$$

where $a, c_n \in \mathbb{C}$. Then the radius of convergence R of f at the point a is given by

$$\frac{1}{R} = \limsup_{n \rightarrow \infty} |c_n|^{1/n}.$$

Theorem 5.5.3 (Analytic Implicit Function Theorem). *Suppose $F(x, y) : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}$ is analytic in a neighborhood of (x_0, y_0) . Furthermore, suppose $F(x_0, y_0) = 0$, and that $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$. Then there is one and only one solution $y = f(x)$ of $F(x, y) = 0$ near (x_0, y_0) that is an analytic function of x near x_0 . Thus, the radius of convergence of the associated power series is nonzero.*

Lemma 5.5.4. *There exists β_* such that for all inverse temperature $\beta > \beta_*$, $Z_{\text{POLYMER}}(\Lambda; z) \neq 0$ for any simply connected finite region Λ on the square lattice with free boundary conditions.*

Proof. Our strategy is to give an upper bound for the sum of (the absolute values) of the weights of all polymers of a given size n starting from a vertex u in direction l . Recall that a polymer can be viewed as a connected component on L which can have degree-4 vertices only at $\mathbb{Z}^2$ -vertices. Instead of working directly with the set of such connected components (possibly with cycles), we will work with a corresponding set of trees (without cycles) by constructing a weight-preserving

injection from the former into the latter by duplicating some $L \setminus \mathbb{Z}^2$ -vertices: if there exist cycles in such a connected component, we virtually “split” an arbitrary $L \setminus \mathbb{Z}^2$ -vertex v into two vertices v', v'' in a cycle and repeat until no cycle exists. Considering the starting vertex u as the root, this procedure results in a tree with $\mathbb{Z}^2$ -vertices and $L \setminus \mathbb{Z}^2$ -vertices alternating and can have one or three children at $\mathbb{Z}^2$ -vertices and can have only one child at $L \setminus \mathbb{Z}^2$ -vertices. In such a tree, edges may not repeat. We define the weight of a tree to be the product of weights on its $\mathbb{Z}^2$ -vertices as in Figure 5.12. It remains unchanged compared to the original polymer, because the vertices we “split” do not include any $\mathbb{Z}^2$ -vertex.

To give an upper bound for the sum of weights of such trees, from now on we consider the set of trees in which edges may repeat (but still with the restriction on degrees). In fact, such trees form a superset of the original set of trees coming from the polymers. The size of a tree is defined as the number of $\mathbb{Z}^2$ -vertices it contains (with repetition).

For convenience, instead of working directly with $\epsilon_\mu, \epsilon_\nu, \epsilon_\phi, \epsilon_\psi$, we define $r = \max \left(\left| \frac{\mu}{\psi} \right|, \left| \frac{\nu}{\psi} \right|, \left| \frac{\phi}{\psi} \right| \right)$. We note that when $\epsilon_\psi < \epsilon_\mu, \epsilon_\nu, \epsilon_\phi$, $r < 1$ and r becomes smaller as β increases, i.e. temperature drops. We denote by A_n the sum of weights of all trees (described above) of size n rooted at an $L \setminus \mathbb{Z}^2$ -vertex with a fixed direction, and denote by B_n the sum of weights of trees of size n rooted at a $\mathbb{Z}^2$ -vertex with a fixed direction (the starting vertex does not contribute to the size n). Then we can write down the follow recurrence relations for $n \geq 1$

$$A_n \leq rB_{n-1} + r \sum_{\substack{i,j,k \geq 0 \\ i+j+k=n-1}} B_i B_j B_k,$$

$$B_n = 3A_n.$$

and the base case is $B_0 = 1$ and consequently $A_1 \leq 2r$. The first recurrence relation comes from the fact that there are two possible ways to extend a tree at a $\mathbb{Z}^2$ -vertex, yielding a weight at most r . For example, consider all possible trees of size n rooted at an $L \setminus \mathbb{Z}^2$ -vertex going towards the southeast direction as in Figure 5.13. The first option is to have a single sub-tree of size $n - 1$ also going southeast (shown in Figure 5.13a). The local contribution to the weight is $\frac{\mu}{\psi}$

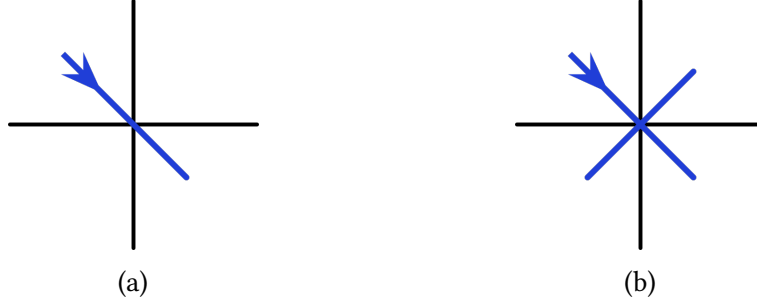


Figure 5.13: Two possible ways to extend a tree at a $\mathbb{Z}^2$ -vertex.

(see Figure 5.12a). The second option is to have three sub-trees with total size $n - 1$ (shown in Figure 5.13b). The local weight contribution is $\frac{\phi}{\psi}$ (see Figure 5.12c). The situations for the other three initial directions rooted at the same $L \setminus \mathbb{Z}^2$ -vertex are symmetric. The second recurrence

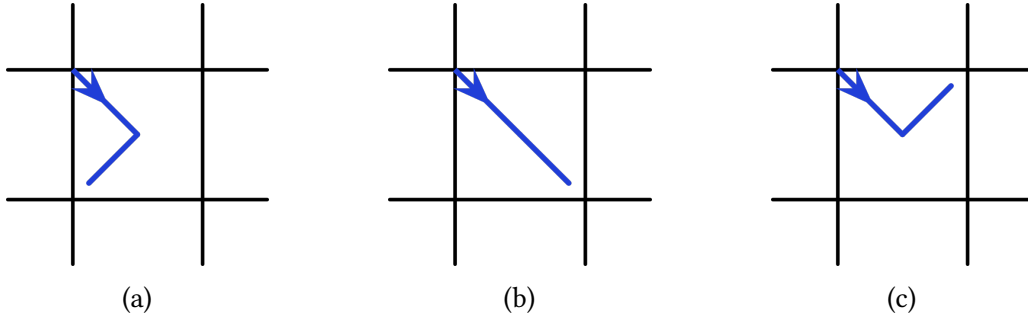


Figure 5.14: Three possible ways to extend a tree at an $L \setminus \mathbb{Z}^2$ -vertex.

relation comes from the fact that there are three possible ways to extend a tree at an $L \setminus \mathbb{Z}^2$ -vertex (Figure 5.14).

Next we give an upper bound for A_n . Let $A_0 = \frac{1}{3}$ and for $n \geq 1$ we have

$$A_n \leq 3r \left(A_{n-1} + 9 \sum_{\substack{i,j,k \geq 0 \\ i+j+k=n-1}} A_i A_j A_k \right),$$

or equivalently

$$\frac{A_n}{(3r)^n} \leq \frac{A_{n-1}}{(3r)^{n-1}} + 9 \sum_{\substack{i,j,k \geq 0 \\ i+j+k=n-1}} \frac{A_i}{(3r)^i} \frac{A_j}{(3r)^j} \frac{A_k}{(3r)^k}. \quad (5.5.3)$$

For convenience we work with the recurrence equation $A'_n = A'_{n-1} + 9 \sum_{\substack{i,j,k \geq 0 \\ i+j+k=n-1}} A'_i A'_j A'_k$ with $A'_0 = \frac{1}{3}$. We remark that $A_n \leq (3r)^n A'_n$. Let the formal power series $f(x) = \sum_{n \geq 0} A'_n x^n$ be the

generating function. Then based on the recurrence, the generating function must satisfy the functional equation

$$f(x) - \frac{1}{3} = xf(x) + 9xf^3(x).$$

Define $F(x, y) = \frac{1}{3} + (x-1)y + 9xy^3$ and let $x_0 = 0, y_0 = \frac{1}{3}$. We have $F(x_0, y_0) = 0$ and $\frac{\partial F}{\partial y}(x_0, y_0) = -1 \neq 0$. According to Theorem 5.5.3, the functional equation determines a unique *analytic* (implicit) function $y = f(x)$, satisfying the initial condition $y_0 = f(x_0)$. It has a power series expansion at x_0 with nonzero radius of convergence $R > 0$. Then by Theorem 5.5.2, the coefficients of the power series $f(x)$ can be upper bounded by $A'_n \leq t^n$ for some constant $t > 0$. Combine this with (5.5.3), we know A_n and B_n are both $O((3rt)^n)$.

Recall that given a finite region of the square lattice Λ , the size of a polymer $|\gamma|$ is the number of $\mathbb{Z}^2$ -vertices it goes through, although γ is sitting on L . Observe that the number of L -vertices a polymer γ goes through is at most $4|\gamma| + 1$, and that observe that $\gamma \neq \gamma^*$ only if γ and γ^* share some L -vertex in common. Let us define, for all γ , $h(\gamma) = 4|\gamma| + 1$. Then, the sum in (5.5.2) can be bounded by

$$\begin{aligned} \sum_{\gamma \neq \gamma^*} |w(\gamma, z)| e^{h(\gamma)} &\leq (4|\gamma^*| + 1) \max_{l \in \gamma^*} \sum_{\gamma \ni l} |w(\gamma, z)| e^{4|\gamma|+1} \\ &\leq (4|\gamma^*| + 1) e \max_{l \in \gamma^*} \sum_{n=1}^{\infty} e^{4n} \sum_{\gamma \ni l, |\gamma|=n} |w(\gamma, z)|. \end{aligned}$$

Note that the set of trees of size n that goes through a vertex l can be viewed as a set of all possible combinations of

- $(l \in L \setminus \mathbb{Z}^2)$ a single tree of size n with 4 possible starting directions; or
- $(l \in L \setminus \mathbb{Z}^2)$ two trees of size i ($1 \leq i \leq n-1$) and $n-i$ with 6 possible starting directions; or
- $(l \in \mathbb{Z}^2)$ two trees of size i ($0 \leq i \leq n-1$) and $n-1-i$ with 2 possible starting directions; or
- $(l \in \mathbb{Z}^2)$ four trees of size i, j, k ($i, j, k \geq 0, i+j+k \leq n-1$) and $n-1-i-j-k$ with 1 possible starting directions.

This means

$$\sum_{\gamma \ni l, |\gamma|=n} |w(\gamma, z)| \leq \max \left(4A_n + 6 \sum_{i=1}^{n-1} A_i A_{n-i}, \quad 2 \sum_{i=0}^{n-1} B_i B_{n-1-i} + \sum_{\substack{i,j,k \geq 0 \\ i+j+k \leq n-1}} B_i B_j B_k B_{n-1-i-j-k} \right) \leq O(n^3 \cdot (3rt)^n).$$

Therefore, in order to satisfy (5.5.2) and thereby apply Theorem 5.5.1, we need

$$O \left(\sum_{n=1}^{\infty} n^3 (3e^4 rt)^n \right) \leq 1.$$

This is satisfied when r is sufficiently small, i.e. the temperature is sufficiently low. $\square$

5.6 An FPTAS at low temperatures

In this section, we apply the framework recently developed by Helmuth, Perkins, and Regts [HPR19] to give efficient approximation algorithms for abstract polymer models when cluster expansion converges. Concretely, the FPTAS is obtained by expressing the log partition function as a Taylor series, using the absence of zeros (i.e. convergence of cluster expansion) to bound the error of the truncated Taylor series, and computing the low-order coefficients of the Taylor series efficiently [Bar16, PR17].

The *Taylor series* for $\log Z_{\text{POLYMER}}(G; z)$ around $z = 0$ is

$$\log Z_{\text{POLYMER}}(G; z) = \sum_{k \geq 1} \frac{z^k}{k!} \frac{\partial^k}{\partial z^k} \log Z_{\text{POLYMER}}(G; 0). \quad (5.6.1)$$

Denote the partial sums of the Taylor series by

$$T_m(G; z) := \sum_{k \geq 1}^m \frac{z^k}{k!} \frac{\partial^k}{\partial z^k} \log Z_{\text{POLYMER}}(G; 0).$$

The following proposition indicates where to truncate the Taylor series to get a good approximation.

Proposition 5.6.1 ([PR17, Lemma 2.2]). *Suppose the degree of the polynomial $Z_{\text{POLYMER}}(G; z)$ is at most N and suppose that $Z_{\text{POLYMER}}(G; z) \neq 0$ for all $|z| \leq \delta$. Then for every $\epsilon > 0$ and every $|z| < \delta$, we have*

$$e^{-\epsilon} \leq \frac{Z_{\text{POLYMER}}(G; z)}{T_m(G; z)} \leq e^{\epsilon}$$

for all $m \geq \frac{\log(N/\epsilon)}{1-|z|/\delta}$.

Based on the observation by Dobrushin [Dob96], the cluster expansion (5.5.1) and Taylor series (5.6.1) are the same power series in z , [HPR19] computed the coefficients of the Taylor series via the cluster expansion

$$\frac{\frac{\partial^k}{\partial z^k} \log Z_{\text{POLYMER}}(G; 0)}{k!} = \sum_{j \geq 1} \sum_{\gamma_1} \cdots \sum_{\gamma_j} \varphi(\gamma_1, \dots, \gamma_j) \frac{1}{k!} \frac{\partial^k}{\partial z^k} \left(\prod_{i=1}^j w(\gamma_i, z) \right)_{z=0}.$$

and proved the following theorem.

Assumption 1. The weight functions $w(\gamma, z)$ are analytic functions of z in a neighborhood of the origin of the complex plane, and there is an absolute constant $\rho > 0$ such that for each $\gamma \in \mathcal{P}(G)$ the first non-zero term in the Taylor series expansion of $w(\gamma, z)$ around zero is of order $k \geq |\gamma|\rho$.

Theorem 5.6.2 ([HPR19, Theorem 2.2]). *Fix Δ and let $\mathcal{G}$ be a set of graphs of degree at most Δ . Suppose:*

1. *There is a constant C so that $Z_{\text{POLYMER}}(G; z)$ is a polynomial in z of degree at most $C|G|$ for all $G \in \mathcal{G}$.*
2. *The weight functions satisfy Assumption 1, and we can compute $w(\gamma, z)$ up to order m for all $G \in \mathcal{G}$ and all $\gamma \in \mathcal{P}(G)$ in time $\exp(O(m + \log |G|))$.*
3. *For every connected subgraph G' of every $G \in \mathcal{G}$, we can list all polymers $\gamma \in \mathcal{P}(G)$ with $\gamma = G'$ in time $\exp(O(|G'|))$.*
4. *There exists $\delta > 0$ so that for all $|z| < \delta$ and all $G \in \mathcal{G}$, $Z_{\text{POLYMER}}(G; z) \neq 0$.*

Then for every z with $|z| < \delta$, there is an FPTAS for $Z_{\text{POLYMER}}(G; z)$ for all $G \in \mathcal{G}$.

Remark 5.6.1. Roughly speaking, condition (1) is to make sure the N in Proposition 5.6.1 is polynomial in the input size, and consequently the number of coefficients (in the Taylor series) needed is logarithmic in the graph size. Condition (2) and condition (3) guarantee that the coefficients can be computed efficiently using the cluster expansion. Condition (4) makes sure that Proposition 5.6.1 applies.

Proof of Theorem 5.0.3. We will verify that our polymer model representation (introduced in Section 5.4) for the even-coloring model (introduced in Section 1.2 and Section 5.3) satisfies all the conditions in Theorem 5.6.2. Observe that in our polymer model, $\Delta = 4$.

1. This is obvious since, given any Λ , each term in $Z_{\text{POLYMER}}(L_\Lambda; z)$ has a factor of at most $z^{\max(p_1, p_2, p_3)}$ on each of the Λ -vertices (where p_1, p_2, p_3 are constants coming from the rational energy parameters), and $|\Lambda|$ is about half of $|L_\Lambda|$.
2. By definition, Assumption 1 is satisfied as the exponent of z in $w(\gamma, z)$ grows linearly in $|\gamma|$. Also, we can compute $w(\gamma, z)$ using linear time (in $|\gamma|$).
3. This is also trivial.
4. This is proved in Lemma 5.5.4.

Therefore, Theorem 5.6.2 gives the approximate counting algorithm for the even-coloring model when the parameters satisfy the condition in Lemma 5.5.4. Under the maps M_I , M_{II} , M_{III} , and M_{IV} in Section 5.3, it directly translates into approximate counting algorithms for the eight-vertex model in all the ordered phases.

The approximate sampling algorithm can be given by [HPR19, Theorem 5.1] as similar conditions are needed. □

5.7 Extension to the torus

In Section 5.4, we defined the polymer model representation of the even-coloring model on the square lattice with free boundary conditions. Since there are no “boundary” vertices on a torus Λ , the polymers on the torus are the same as in the free boundary case (connected subgraphs of L_Λ in which each vertex has degree two or four with the same restrictions), but with no exception for boundary vertices. The mapping from $C_e(\Lambda)$, the even-colorings on Λ , to $C(L_\Lambda)$, the compatible subsets of polymers on L_Λ , is still well-defined and is two-to-one. However it is no longer *onto*, i.e., although for any even-coloring (and for the even-coloring obtained by flipping all colors) there is a unique corresponding compatible subset of polymers (obtained by separating the green edges from the red ones), there exist compatible subsets of polymers that do not correspond to any even-colorings under this mapping. See Figure 5.15a for an example.

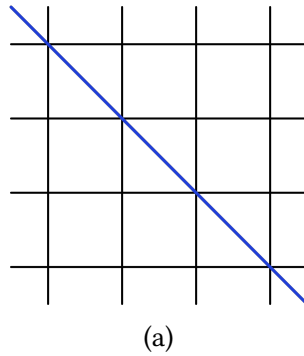


Figure 5.15: A polymer configuration on $(\mathbb{Z}/4\mathbb{Z})^2$ that does not correspond to any even-colorings.

Now we give a characterization for the image of $C_e(\Lambda)$ under this mapping in $C(L_\Lambda)$. Suppose Λ is a torus $(\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$. Given a compatible subset of polymers $\Gamma \in C(L_\Lambda)$ and a real number $c \in [0, n)$, let $\text{Col}(\Gamma, c) = 1$ if the number of polymer edges cut by the vertical line $x = c$ is even and -1 if it is odd. If $c \in \mathbb{Z}$ and the line $x = c$ passes through a $\mathbb{Z}^2$ -vertex of Γ with degree four, then we consider two polymer edges are cut. If c is a half integer and the line $x = c$ passes through a $L \setminus \mathbb{Z}^2$ -vertex of Γ , and if the two incident edges are on the same side of $x = c$, then we consider two (or zero) polymer edges are cut. We claim that $\text{Col}(\Gamma, c)$ is constant for all c . One can imagine a vertical line sweeping towards the right, starting at c . During this process, the number of polymer

edges cut by the vertical line does not change except when going through some integral or half integral points. However, at these points since all the polymers have degree 2 or 4 and due to the restriction on the allowed configurations of Γ there, the parity of this number does not change. Therefore, we can write $\text{Col}(\Gamma)$ for $\text{Col}(\Gamma, c)$ for any c . Similarly let $\text{Row}(\Gamma) = 1$ if the number of polymer edges cut by a horizontal line is even and -1 otherwise. We also write $\text{Col}(\gamma) = \text{Col}(\{\gamma\})$. It is not hard to see that $\text{Col}(\Gamma) = \prod_{\gamma \in \Gamma} \text{Col}(\gamma)$ and similarly for Row .

We claim that $\Gamma \in C(L_\Lambda)$ is in the image of $C_e(\Lambda)$ if and only if $\text{Row}(\Gamma) = \text{Col}(\Gamma) = 1$. The “only if” direction is because that the colors on any even-coloring has to alternate an even number of times both horizontally and vertically and each alternation corresponds to crossing a polymer edge. The “if” direction follows an argument similar to that for the free boundary case in Section 5.4 and notice that a torus is identified with a rectangle region in which we identify the top and bottom boundary edges, and also the right and left boundary edges. So when the two edges in each identified boundary edge pair have the same color, we get a valid even-coloring of the torus.

We inherit the weight function $w(\gamma, z)$ on polymers from Section 5.4. For $b_R, b_C \in \{0, 1\}$, define

$$S_{b_R b_C} = \sum_{\substack{\Gamma \in C(L_\Lambda) \\ \text{Row}(\Gamma)=1-2b_R, \text{Col}(\Gamma)=1-2b_C}} \prod_{\gamma \in \Gamma} w(\gamma, z).$$

We note that S_{00} is the sum of weights of all polymers corresponding to the set of even-colorings.

We have

$$\begin{aligned} Z_{\text{EC}}(\Lambda; \mu, \nu, \phi, \psi) &= 2\psi^{|\mathbb{Z}_\Lambda^2|} \sum_{\Gamma \in C(L_\Lambda)} \frac{1 + \text{Row}(\Gamma)}{2} \frac{1 + \text{Col}(\Gamma)}{2} \prod_{\gamma \in \Gamma} w(\gamma, z) \\ &= \frac{1}{2} \psi^{|\mathbb{Z}_\Lambda^2|} \sum_{\Gamma \in C(L_\Lambda)} \left(1 + \prod_{\gamma \in \Gamma} \text{Row}(\gamma) \right) \left(1 + \prod_{\gamma \in \Gamma} \text{Col}(\gamma) \right) \prod_{\gamma \in \Gamma} w(\gamma, z) \\ &= \frac{1}{2} \psi^{|\mathbb{Z}_\Lambda^2|} \sum_{\Gamma \in C(L_\Lambda)} \left(\prod_{\gamma \in \Gamma} w(\gamma, z) + \prod_{\gamma \in \Gamma} \text{Row}(\gamma) w(\gamma, z) \right. \\ &\quad \left. + \prod_{\gamma \in \Gamma} \text{Col}(\gamma) w(\gamma, z) + \prod_{\gamma \in \Gamma} \text{Row}(\gamma) \text{Col}(\gamma) w(\gamma, z) \right) \\ &= \frac{1}{2} \psi^{|\mathbb{Z}_\Lambda^2|} (Z + Z_R + Z_C + Z_{RC}) \end{aligned}$$

where

$$\begin{cases} Z &= \sum_{\Gamma \in \mathcal{C}(L_\Lambda)} \prod_{\gamma \in \Gamma} w(\gamma, z) &= S_{00} + S_{01} + S_{10} + S_{11}, \\ Z_R &= \sum_{\Gamma \in \mathcal{C}(L_\Lambda)} \prod_{\gamma \in \Gamma} \text{Row}(\gamma) w(\gamma, z) &= S_{00} + S_{01} - S_{10} - S_{11}, \\ Z_C &= \sum_{\Gamma \in \mathcal{C}(L_\Lambda)} \prod_{\gamma \in \Gamma} \text{Col}(\gamma) w(\gamma, z) &= S_{00} - S_{01} + S_{10} - S_{11}, \\ Z_{RC} &= \sum_{\Gamma \in \mathcal{C}(L_\Lambda)} \prod_{\gamma \in \Gamma} \text{Row}(\gamma) \text{Col}(\gamma) w(\gamma, z) &= S_{00} - S_{01} - S_{10} + S_{11}. \end{cases}$$

Define $w_R(\gamma, z) = \text{Row}(\gamma)w(\gamma, z)$, $w_C(\gamma, z) = \text{Col}(\gamma)w(\gamma, z)$, and $w_{RC}(\gamma, z) = \text{Row}(\gamma)\text{Col}(\gamma)w(\gamma, z)$.

Then Z , Z_R , Z_C , and Z_{RC} are the partition functions of four abstract polymer models with weight functions w , w_R , w_C , and w_{RC} , respectively. We can approximately compute Z using the same techniques in Section 5.5 and Section 5.6. Since the norm of the four weight functions are identical, the condition in Theorem 5.5.1 and the proof of Lemma 5.5.4 also works for w_R , w_C , and w_{RC} under the same condition on energies and temperature.

Although we have FPTAS for the four quantities Z , Z_R , Z_C , Z_{RC} , this does not mean we can approximately compute their sum $4S_{00}$ under Definition 5.0.1. The problem is that it is possible that four quantities with large norms can add up to a number whose norm is very small, in which case the relative error guarantee for the four quantities do not translate to a good approximation to the sum. We demonstrate that this is not a problem by proving S_{00} is not negligible compared with S_{01} , S_{10} , and S_{11} using a Peierls argument.

From now on we assume $m \geq n$. The cases where $n \geq m$ can be proved using a symmetric argument. Consider a polymer γ^* with $\text{Row}(\gamma^*) = -1$. Define $F_{\gamma^*} = \{\Gamma \in \mathcal{C}(L_\Lambda) \mid \gamma^* \in \Gamma, \text{Row}(\Gamma) = -1\}$. Define an injection from F_{γ^*} to $\{\Gamma \in \mathcal{C}(L_\Lambda) \mid \text{Row}(\Gamma) = 1\}$ by removing γ^* . Using the same notation $r = \max\left(\left|\frac{\mu}{\psi}\right|, \left|\frac{\nu}{\psi}\right|, \left|\frac{\phi}{\psi}\right|\right)$ as in Section 5.4, we have

$$\sum_{\Gamma \in F_{\gamma^*}} \prod_{\gamma \in \Gamma} w(\gamma, z) \leq r^{|\gamma^*|} \sum_{\Gamma \in F_{\gamma^*}} \prod_{\gamma \in \Gamma \setminus \{\gamma^*\}} w(\gamma, z) \leq r^{|\gamma^*|} \sum_{\substack{\Gamma \in \mathcal{C}(L_\Lambda) \\ \text{Row}(\Gamma)=1}} \prod_{\gamma \in \Gamma} w(\gamma, z) = r^{|\gamma^*|} (S_{00} + S_{01}).$$

Using techniques similar to those in Section 5.5, one can show that the number of polymers of size l starting at a given L -vertex is at most t^l for some constant t . It is not hard to see that on a torus $(\mathbb{Z}/n\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$, in order to have $\text{Row}(\gamma^*) = -1$, it must be true that $|\gamma^*| \geq m$. Indeed, if

$|\gamma^*| < m$, then there must be row $y = r$ which does not intersect γ^* , resulting in an even number of cuts, and thus $\text{Row}(\gamma^*, r) = 1$. As there are at most n possible starting vertex, the total number of polymers of size l with $\text{Row}(\gamma^*) = -1$ is at most $n \cdot t^l$. Since every term in $S_{10} + S_{11}$ has at least one polymer γ with $\text{Row}(\gamma) = -1$, the sum is upper bounded by

$$S_{10} + S_{11} \leq \sum_{l=m}^{mn} n \cdot t^l \cdot r^l (S_{00} + S_{01}) \leq \frac{n(rt)^m}{1 - rt} (S_{00} + S_{01}).$$

Since $m \geq n$, when the temperature is sufficiently low, i.e. $rt < 1$, both S_{10} and S_{11} are negligible in $Z = S_{00} + S_{01} + S_{10} + S_{11}$. In particular we can have $S_{10} < o(1/m)(S_{00} + S_{01})$.

A symmetric argument shows that

$$S_{01} \leq S_{01} + S_{11} \leq \frac{m(rt)^n}{1 - rt} (S_{00} + S_{10}).$$

When $rt < 1$, $S_{01} < O(m)(S_{00} + S_{10}) = O(m)S_{00} + o(1)S_{01}$, so $S_{01} < O(m)S_{00}$. Thus, up to a factor $\Theta(m)$, S_{00} is no less than S_{01} , and also S_{10} , S_{11} are both negligible in S_{00} . Hence, an FPTAS for S_{00} can be obtained by approximating Z , Z_R , Z_C , and Z_{RC} with a sufficiently large inverse polynomial (in m and n) accuracy.

BIBLIOGRAPHY

- [AR05] David Allison and Nicolai Reshetikhin. Numerical study of the 6-vertex model with domain wall boundary conditions. *Annales de l'Institut Fourier*, 55(6):1847–1869, 2005.
- [Bar16] Alexander Barvinok. *Combinatorics and Complexity of Partition Functions*, volume 30 of *Algorithms and Combinatorics*. Springer, 2016.
- [Bax71] R. J. Baxter. Eight-vertex model in lattice statistics. *Phys. Rev. Lett.*, 26:832–833, Apr 1971.
- [Bax72] R. J. Baxter. Partition function of the eight-vertex lattice model. *Annals of Physics*, 70(1):193 – 228, 1972.
- [Bax82] R. J. Baxter. *Exactly Solved Models in Statistical Mechanics*. Academic Press Inc., San Diego, CA, USA, 1982.
- [BCG16] Alexei Borodin, Ivan Corwin, and Vadim Gorin. Stochastic six-vertex model. *Duke Math. J.*, 165(3):563–624, 02 2016.
- [BGJ⁺17] 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.
- [BGRT13] Antonio Blanca, David Galvin, Dana Randall, and Prasad Tetali. Phase coexistence and slow mixing for the hard-core model on $\mathbb{Z}^2$. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM)*, pages 379–394, 2013.
- [BKW73] H. J. Brascamp, H. Kunz, and F. Y. Wu. Some rigorous results for the vertex model in statistical mechanics. *Journal of Mathematical Physics*, 14(12):1927–1932, 1973.
- [BN98] G. T. Barkema and M. E. J. Newman. Monte carlo simulation of ice models. *Phys. Rev. E*, pages 1155–1166, 1998.

- [CC17] Jin-Yi Cai and Xi Chen. *Complexity Dichotomies for Counting Problems*, volume 1. Cambridge University Press, 2017.
- [CF17] Jin-Yi Cai and Zhiguo Fu. Complexity classification of the eight-vertex model. *CoRR*, abs/1702.07938, 2017.
- [CFS17] Jin-Yi Cai, Zhiguo Fu, and Shuai Shao. A complexity trichotomy for the six-vertex model. *CoRR*, abs/1704.01657, 2017.
- [CFX18] Jin-Yi Cai, Zhiguo Fu, and Mingji Xia. Complexity classification of the six-vertex model. *Information and Computation*, 259:130 – 141, 2018.
- [CGG⁺16] Jin-Yi Cai, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, Mark Jerrum, Daniel Štefankovič, and Eric Vigoda. #BIS-hardness for 2-spin systems on bipartite bounded degree graphs in the tree non-uniqueness region. *Journal of Computer and System Sciences*, 82(5):690–711, 2016.
- [CL20] 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, page TBD, 2020.
- [CLL19] 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, Philadelphia, PA, USA, 2019. Society for Industrial and Applied Mathematics.
- [CLLY20] Jin-Yi Cai, Tianyu Liu, Pinyan Lu, and Jing Yu. Approximability of the eight-vertex model. In *Proceedings of the 35th Computational Complexity Conference*, CCC ’20, page TBD, 2020.
- [DFK91] 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.

- [DGGJ04] Martin Dyer, Leslie Ann Goldberg, Catherine Greenhill, and Mark Jerrum. The relative complexity of approximate counting problems. *Algorithmica*, 38(3):471–500, Mar 2004.
- [Dim18] Evgeni Dimitrov. Six-vertex Models and the GUE-corners Process. *International Mathematics Research Notices*, 04 2018.
- [Dob96] R. L. Dobrushin. Estimates of semi-invariants for the Ising model at low temperatures topics in statistical and theoretical physics. *Translations of the American Mathematical Society-Series 2*, 177:59–81, 1996.
- [Elo99] K. Eloranta. Diamond Ice. *Journal of Statistical Physics*, 96:1091–1109, September 1999.
- [FR19] Matthew Fahrbach and Dana Randall. Slow mixing of Glauber dynamics for the six-vertex model in the ordered phases. In Dimitris Achlioptas and László A. Végh, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)*, volume 145 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 37:1–37:20, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.
- [FV17] Sacha Friedli and Yvan Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*. Cambridge University Press, 2017.
- [FW70] Chungpeng Fan and F. Y. Wu. General lattice model of phase transitions. *Phys. Rev. B*, 2:723–733, Aug 1970.
- [GC01] A.J. Guttmann and A.R. Conway. Square lattice self-avoiding walks and polygons. *Annals of Combinatorics*, 5(3):319–345, Dec 2001.
- [GJ08] Leslie Ann Goldberg and Mark Jerrum. Inapproximability of the Tutte polynomial. *Information and Computation*, 206(7):908 – 929, 2008.
- [GJ12a] Leslie Ann Goldberg and Mark Jerrum. Approximating the partition function of the ferromagnetic Potts model. *J. ACM*, 59(5):25:1–25:31, November 2012.

- [GJ12b] Leslie Ann Goldberg and Mark Jerrum. Inapproximability of the Tutte polynomial of a planar graph. *Computational Complexity*, 21(4):605–642, Dec 2012.
- [GJM15] Leslie Ann Goldberg, Mark Jerrum, and Colin McQuillan. Approximating the partition function of planar two-state spin systems. *Journal of Computer and System Sciences*, 81(1):330 – 358, 2015.
- [GJS76] M.R. Garey, D.S. Johnson, and L. Stockmeyer. Some simplified NP-complete graph problems. *Theoretical Computer Science*, 1(3):237 – 267, 1976.
- [GMP04] Leslie Ann Goldberg, Russell Martin, and Mike Paterson. Random sampling of 3-colorings in $\mathbb{Z}^2$. *Random Structures & Algorithms*, 24(3):279–302, 2004.
- [GR10] Sam Greenberg and Dana Randall. Slow mixing of Markov chains using fault lines and fat contours. *Algorithmica*, 58(4):911–927, Dec 2010.
- [GvV15] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. *J. ACM*, 62(6), December 2015.
- [GW13] Heng Guo and Tyson Williams. The complexity of planar Boolean #CSP with complex weights. In *Proceedings of the 40th International Colloquium on Automata, Languages, and Programming*, ICALP ’13, pages 516–527, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.
- [HL16] Sangxia Huang and Pinyan Lu. A dichotomy for real weighted Holant problems. *Computational Complexity*, 25(1):255–304, March 2016.
- [HPR19] Tyler Helmuth, Will Perkins, and Guus Regts. Algorithmic Pirogov-Sinai theory. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC ’19, pages 1009–1020, New York, NY, USA, 2019. ACM.

- [HVV15] 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.
- [Jae90] F. Jaeger. On transition polynomials of 4-regular graphs. In Geña Hahn, Gert Sabidussi, and Robert E. Woodrow, editors, *Cycles and Rays*, pages 123–150. Springer Netherlands, Dordrecht, 1990.
- [Jer03] Mark Jerrum. *Counting, Sampling and Integrating: Algorithm and Complexity*. Birkhäuser, Basel, 2003.
- [JS89] Mark Jerrum and Alistair Sinclair. Approximating the permanent. *SIAM Journal on Computing*, 18(6):1149–1178, 1989.
- [JS93] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the ising model. *SIAM Journal on Computing*, 22(5):1087–1116, 1993.
- [JSV04] Mark Jerrum, Alistair Sinclair, and Eric Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *J. ACM*, 51(4):671–697, July 2004.
- [JVV86] 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.
- [Kas61] 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.
- [Kas67] P.W. Kasteleyn. Graph theory and crystal physics. In F. Harary, editor, *Graph Theory and Theoretical Physics*, pages 43–110. Academic Press, 1967.
- [KL83] 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.
- [KL17] Rick Keesman and Jules Lamers. Numerical study of the f model with domain-wall boundaries. *Phys. Rev. E*, 95:052117, May 2017.
 - [KP86] R. Kotecký and D. Preiss. Cluster expansion for abstract polymer models. *Communications in Mathematical Physics*, 103(3):491–498, Sep 1986.
 - [KS18] David Keating and Ananth Sridhar. Random tilings with the GPU. *Journal of Mathematical Physics*, 59(9):091420, 2018.
 - [Kup96] Greg Kuperberg. Another proof of the alternative-sign matrix conjecture. *International Mathematics Research Notices*, 1996(3):139–150, 1996.
 - [Lic97] W.B.Raymond Lickorish. *An Introduction to Knot Theory*. Springer New York, 1997.
 - [Lie67a] Elliott H. Lieb. Exact solution of the F model of an antiferroelectric. *Phys. Rev. Lett.*, 18:1046–1048, Jun 1967.
 - [Lie67b] Elliott H. Lieb. Exact solution of the two-dimensional Slater KDP model of a ferroelectric. *Phys. Rev. Lett.*, 19:108–110, Jul 1967.
 - [Lie67c] Elliott H. Lieb. Residual entropy of square ice. *Phys. Rev.*, 162:162–172, Oct 1967.
 - [Liu18] Tianyu Liu. Torpid mixing of Markov chains for the six-vertex model on $\mathbb{Z}^2$. In Eric Blais, Klaus Jansen, José D. P. Rolim, and David Steurer, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2018)*, volume 116 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 52:1–52:15, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

- [LKR18] I. Lyberg, V. Korepin, G. A. P. Ribeiro, and J. Viti. Phase separation in the six-vertex model with a variety of boundary conditions. *Journal of Mathematical Physics*, 59(5):053301, 2018.
- [LKV17] I Lyberg, V Korepin, and J Viti. The density profile of the six vertex model with domain wall boundary conditions. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(5):053103, 2017.
- [LPW06] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, 2006.
- [LRS01] Michael Luby, Dana Randall, and Alistair Sinclair. Markov chain algorithms for planar lattice structures. *SIAM Journal on Computing*, 31(1):167–192, 2001.
- [McQ13] Colin McQuillan. Approximating Holant problems by winding. *CoRR*, abs/1301.2880, 2013.
- [MW96] M. Mihail and P. Winkler. On the number of Eulerian orientations of a graph. *Algorithmica*, 16(4):402–414, Oct 1996.
- [Pau35] Linus Pauling. The structure and entropy of ice and of other crystals with some randomness of atomic arrangement. *Journal of the American Chemical Society*, 57(12):2680–2684, 1935.
- [Pei36] R. Peierls. Statistical theory of adsorption with interaction between the adsorbed atoms. *Mathematical Proceedings of the Cambridge Philosophical Society*, 32(3):471–476, 1936.
- [PR17] V. Patel and G. Regts. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM Journal on Computing*, 46(6):1893–1919, 2017.
- [PS01] Nikolay Prokof’ev and Boris Svistunov. Worm algorithms for classical statistical models. *Phys. Rev. Lett.*, 87:160601, Sep 2001.

- [Ran06] Dana Randall. Slow mixing of Glauber dynamics via topological obstructions. In *Proceedings of the Seventeenth Annual ACM-SIAM Symposium on Discrete Algorithm*, SODA '06, pages 870–879, Philadelphia, PA, USA, 2006. Society for Industrial and Applied Mathematics.
- [RS72] 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.
- [RT00] Dana Randall and Prasad Tetali. Analyzing Glauber dynamics by comparison of Markov chains. *Journal of Mathematical Physics*, 41(3):1598–1615, 2000.
- [Sin92] Alistair Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1:351–370, 1992.
- [SJ89] Alistair Sinclair and Mark Jerrum. Approximate counting, uniform generation and rapidly mixing markov chains. *Information and Computation*, 82(1):93 – 133, 1989.
- [SS14] Allan Sly and Nike Sun. Counting in two-spin models on d-regular graphs. *The Annals of Probability*, 42(6):2383–2416, 2014.
- [Sut67] Bill Sutherland. Exact solution of a two-dimensional model for hydrogen-bonded crystals. *Phys. Rev. Lett.*, 19:103–104, Jul 1967.
- [Sut70] Bill Sutherland. Two dimensional hydrogen bonded crystals without the ice rule. *Journal of Mathematical Physics*, 11(11):3183–3186, 1970.
- [SZ04] Olav F. Syljuåsen and M. B. Zvonarev. Directed-loop monte carlo simulations of vertex models. *Phys. Rev. E*, 70:016118, Jul 2004.
- [TF61] 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.

- [Tut54] W. T. Tutte. A contribution to the theory of chromatic polynomials. *Canad. J. Math.*, 6(0):80–91, January 1954.
- [Val02] Leslie Valiant. Quantum circuits that can be simulated classically in polynomial time. *SIAM Journal on Computing*, 31(4):1229–1254, 2002.
- [Val08] Leslie G. Valiant. Holographic algorithms. *SIAM J. Comput.*, 37(5):1565–1594, February 2008.
- [Ver88] Michel Las Vergnas. On the evaluation at $(3, 3)$ of the Tutte polynomial of a graph. *Journal of Combinatorial Theory, Series B*, 45(3):367 – 372, 1988.
- [Wel69] D.J.A. Welsh. Euler and bipartite matroids. *Journal of Combinatorial Theory*, 6(4):375 – 377, 1969.
- [Yan78] 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.
- [YN79] A. Yanagawa and J.F. Nagle. Calculations of correlation functions for two-dimensional square ice. *Chemical Physics*, 43(3):329 – 339, 1979.
- [Zei96] Doron Zeilberger. Proof of the alternating sign matrix conjecture. *Electronic Journal of Combinatorics*, 3(2), 1996.