Approximate counting, phase transitions and geometry of polynomials

Jingcheng Liu

Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2019-110
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2019/EECS-2019-110.html

August 1, 2019

Approximate counting, phase transitions and geometry of polynomials

by

Jingcheng Liu

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Computer Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Alistair Sinclair, Chair
Professor Prasad Raghavendra
Professor Nike Sun

Summer 2019

Approximate counting, phase transitions and geometry of polynomials

by

Jingcheng Liu

Abstract

Approximate counting, phase transitions and geometry of polynomials

by

Jingcheng Liu
Doctor of Philosophy in Computer Science

University of California, Berkeley
Professor Alistair Sinclair, Chair

In classical statistical physics, a phase transition is understood by studying the geometry (the zero-set) of an associated polynomial (the partition function). In this thesis, we will show that one can exploit this notion of phase transitions algorithmically, and conversely exploit the analysis of algorithms to understand phase transitions.

As applications, we give efficient deterministic approximation algorithms (FPTAS) for counting q-colorings, and for computing the partition function of the Ising model.

Contents

Acknowledgments

This thesis would not have been possible without the support of many people along the journey in the past five years. The names included here does not form a complete list of the people I am grateful to, and I apologize in advance if I failed to mention your name here: rest assured that I really appreciate the support of those who're reading these lines.

First I would like to thank my Ph.D advisor, Alistair Sinclair, for his constant encouragement, assistance and guidance in the past five years, both in a technical sense and in other aspects of life. I remain in awe of his energy and enthusiasm, and I thank his patience and support as I develop myself. Alistair, working with you was a great learning experience, and I am immensely grateful and feel very lucky to have had you as my advisor. I'd also like to thank my thesis committee, Prasad Raghavendra, Alistair, Nike Sun and Luca Trevisan. Their feedback during my qualifying exam was already very helpful, and was the starting point of the last two parts of this thesis. Additionally, I would like to thank my undergraduate advisors Pinyan Lu and Yong Yu, for introducing me to theoretical computer science and the continual support. Also thanks to Pinyan Lu, for hosting me as an intern at Microsoft Research, Asia during 2013-2014, where I enjoyed a wonderful time and collaboration.

The results in this thesis are derived largely in collaboration with Piyush Srivastava and Alistair. Piyush has visited Berkeley many times during the past five years, and I have been fortunate to have received significant advice and support from him, in addition to the joyous discussions on many problems, research or otherwise. His uncompromising style and attention to details remain a constant inspiration to me. Without his introduction to the Lee-Yang theory, this thesis would have been completely different. Also thanks to Nikhil Srivastava, the course *Geometry of Polynomials* you offered introduced me to the many tools and problems in the field. Last but not the least, I would also like to thank the many (anonymous) reviewers for their comments on early versions of the results in this thesis. Special thanks goes to Guus Regts and an anonymous reviewer for pointing out a simplification to an early version of the Fisher zeros result, which inspired the current further simplifcation presented in this thesis; also thanks to Alexander Barvinok and Guus Regts for motivating discussions on the Potts model, and explaining their existing results, which was the starting point of the last part of the thesis.

I feel incredibly lucky and a great honor to have worked closely with many brilliant collaborators in the past: Pinyan Lu, Alistair Sinclair, Piyush Srivastava, Heng Guo, Chihao Zhang, Mark Jerrum, Kunal Talwar, Matthew Green, Peihan Miao, Ian Miers, Alessandro Chiesa, Chengyu Lin, Pratyush Mishra. Additionally, I am also indebted to numerous inspiring discussions with Shayan Oveis Gharan, Leslie Ann Goldberg, Daniel Štefankovič, Yitong Yin and Yumeng Zhang. I will surely miss the enlightening discussions that we had.

I would like to thank all my fellow theory students and postdocs at Berkeley for our shared lunches, retreats, basketball and inspiring talks. I especially thank Nima Anari, Frank Ban, Antonio Blanca, Jonah Brown-Cohen, Fotis Iliopoulos, Tarun Kathuria, Pasin Manurangsi, Peihan Miao, Aviad Rubinstein, Manuel Sabin, Aaron Schild, Tselil Schramm, Akshayaram Srinivasan, Di Wang, Sam Wong, Elizabeth Yang and Richard Zhang, for our fun and inspirational moments. Also thanks to the amazing theory courses offered by Alessandro Chiesa, Sanjam Garg, Russell Impagliazzo, Lap Chi Lau and Luca Trevisan, from which I have benefited a lot.

To Kunal Talwar, thanks for hosting me as an intern at Google during the summer of 2018. The internship was a tremendous learning experience to me, and our collaboration was enjoyable. In addition, I was also fortunate to learn from Vitaly Feldman, Jerry Li and many others at Google.

Lastly, I would like to thank my family, my mom and dad, for their continual love and support. I am very fortunate to have you with me throughout my life.

Chapter 1

Introduction

Counting problems arise in numerous fields of study, including statistical physics, combinatorics, statistical inference and estimation, volume computation and machine learning. In a counting problem, the goal is to compute or to estimate a weighted sum or integral. Canonical examples include estimating probabilities, or the expectation of a random variable, or drawing samples from a given probability distribution. A central object in all these problems is the *partition function*, a weighted sum over configurations, which generalizes classical combinatorial counting problems. A number of examples come from spin systems in statistical physics. Such a system is defined on a graph $G = (V, E)$, so that the entities comprising the system correspond to the vertices V and their pairwise interactions to the edges E. A *configuration* of the system is an assignment $\boldsymbol{\sigma} : V \to [q]$ of one of q possible values (often called "spins", or "states") to each vertex. The model assigns to each configuration $\boldsymbol{\sigma}$ a positive weight $w(\boldsymbol{\sigma})$ that depends on the local interactions, and the partition function is defined as the sum $Z_G = \sum_{\boldsymbol{\sigma}} w(\boldsymbol{\sigma})$.

To illustrate ideas, we discuss the classical Ising model, which was first introduced a century ago as a model for magnetic phase transitions by Lenz and Ising [Isi25], and has since become an important tool for the modeling of interacting systems. The Ising model, like many other spin systems inspired by statistical physics, is also studied as a graphical model (or Markov random field) in machine learning. There are $q = 2$ spins in the Ising model, so that a configuration is an assignment $\boldsymbol{\sigma} : V \to \{+, -\}$ of one of two possible values ($+$ or $-$) to each vertex. Let λ be the *vertex activity* (sometimes also called an "external field"), and let β be the *edge activity* that models the tendency of spins to agree with their neighbors. The model assigns a weight to each configuration $\boldsymbol{\sigma}$ as follows:

$$w(\boldsymbol{\sigma}) := \beta^{|\{\{u,v\}\in E \mid \boldsymbol{\sigma}(u)\neq\boldsymbol{\sigma}(v)\}|}\lambda^{|\{v \mid \boldsymbol{\sigma}(v)=+\}|} = \beta^{|E(S,\overline{S})|}\lambda^{|S|},$$

where $S = S(\boldsymbol{\sigma})$ is the set of vertices assigned spin $+$ in $\boldsymbol{\sigma}$ and $E\big(S,\overline{S}\big)$ is the set of edges in the cut $\big(S,\overline{S}\big)$ (i.e., the set of pairs of adjacent vertices assigned opposite spins). The probability of a configuration $\boldsymbol{\sigma}$ under the *Gibbs distribution* is then $\mu(\boldsymbol{\sigma}) := w(\boldsymbol{\sigma})/Z_G^{\beta}(\lambda)$, where the normalizing factor $Z_G^{\beta}(\lambda)$ is the *partition function* defined as

$$Z_G^{\beta}(\lambda) := \sum_{\boldsymbol{\sigma}:V\to\{+,-\}} w(\boldsymbol{\sigma}) = \sum_{S\subseteq V} \beta^{|E(S,\overline{S})|}\lambda^{|S|}. \tag{1.1}$$

Notice that the partition function may be interpreted combinatorially as a *cut generating polynomial* in the graph G. In particular, if $\beta > 1$ then the model is called *antiferromagnetic*, as the corresponding Gibbs distribution favors “disagreements” on edges, or in other words, larger cuts; while if $\beta < 1$, the model is *ferromagnetic* and the distribution favors “agreements” or smaller cuts. The parameter λ models an “external field”: if $\lambda = 1$, we say the model has “zero-field”, as the model is symmetric regarding the two spins; while if $\lambda > 1$, the distribution favors the spin $+$.

A natural generalization of the Ising model is the Potts model, which can be seen as a generating polynomial for graph colorings. Given a graph $G = (V, E)$, an *edge activity* w, and an integer $q \geq 2$, the partition function of the q-state Potts model is given by

$$Z_G(w) := \sum_{\sigma:V\to[q]} w^{|\{\{u,v\}\in E\,:\,\sigma(u)=\sigma(v)\}|}. \tag{1.2}$$

Here σ ranges over arbitrary (not necessarily proper) assignments of colors to vertices, and each such coloring has a weight $w^{m(\sigma)}$, where $m(\sigma)$ is the number of monochromatic edges in σ. Note that the number of proper q-colorings of G is $Z_G(0)$. Furthermore, since graph cuts are precisely 2-colorings, the zero-field Ising model is actually the special case of the Potts model when $q = 2$ (along with a change of variable from w to $\beta = \frac{1}{w}$, and rescaling by $w^{-|E|}$).

Given their ubiquitous role, a central question is the computation of partition functions. There has been much progress on *dichotomy theorems*, which attempt to completely classify these problems as being either #P-hard or computable (exactly) in FP (see, e.g., [CCL10, GGJT10]). Since the problems are in fact #P-hard in most cases, algorithmic interest has focused largely on *approximation*, motivated also by the general observation [JVV86] that approximate counting (approximating the partition function) is polynomial time equivalent to sampling (approximately) from the underlying Gibbs distribution. Unlike exact counting, much less is known about the complexity of approximating partition functions. Recent developments seem to suggest that one might be able to obtain a classification of the computational complexity of approximation, via the study of "phase transitions" in the Gibbs distribution. Such a connection has been established in various special cases, e.g., for the partition function of independent sets (also known as the *hard-core model*) over the real line [Sly10, SS14b, GGŠ+14], and for the antiferromagnetic Ising model over the real line [SS14b].

1.1 Algorithms, phase transitions, and zeros of polynomials

Historically, there have been two distinct (though closely related) mechanisms for defining and understanding phase transitions in statisical physics. The first is decay of long-range correlations in the Gibbs measure, which is familiar in theoretical computer science due to its extensive use in approximation algorithms and the analysis of spin systems and graphical models. Roughly speaking, "correlation decay" refers to the phenomenon whereby the effect on the spin value at a fixed vertex of spins at distant vertices decays to zero as the distance tends to infinity. The second mechanism, which is more classical and less familiar in computer science, is analyticity of an appropriate limit of the "free energy density" $\frac{1}{n}\log Z_n$ (where Z_n is the partition function, and n is a size parameter), as the size of the graphs under consideration increases to infinity. This second notion connects naturally to the stability theory of polynomials, and in particular to the study of the location of *complex* roots of the partition function Z, even when only real values of the parameters make physical sense in the model. The seminal work of Lee and Yang [LY52, YL52] was one of the first, and certainly the best known, to use this notion. It is interesting to note that the stability theory of polynomials has seen a recent surge of interest following the central role it has played in developments in a wide variety of areas ranging from mathematical physics to combinatorics and theoretical computer science: examples include the resolution of the Kadison-Singer conjecture [MSS15b], proofs of the existence of Ramanujan graphs [MSS15a], and progress on the traveling salesman problem and other algorithmic questions (see, e.g., [AG15, AG17, SV17]).

We now briefly describe the connection between the analyticity of the free energy and the location of complex zeros of the partition function. The first ingredient is that natural observables of the model (e.g., the *magnetization*) can be written as derivatives of the free energy with respect to an appropriate parameter of the model. Thus, analyticity of the free energy for a given range S of parameters implies that all such observables vary continuously (and have continuous derivatives) when the parameter value lies in S, which in turn implies that there is no phase transition in S. However, it is not hard to see that for any finite graph, the free energy is always analytic as a function of β when β lies on the positive real axis, suggesting a complete absence of phase transitions. Indeed, it turns out (see, e.g., [Sim93, Chapter 1]) that in order to see phase transitions one has to consider *infinite* graphs. For concreteness, we consider the case of the Ising model on the infinite 2-dimensional integer lattice $\mathbb{Z}^2$ [YL52]. The free energy density of such an infinite graph is defined as the limit of the free energy densities of a suitable increasing sequence of finite subgraphs (e.g., increasing rectangles in $\mathbb{Z}^2$). Lee and Yang [YL52] showed that, for infinite graphs of sub-exponential growth (including $\mathbb{Z}^2$), the free energy density obtained via this prescription is well defined and analytic for a range of parameters S provided that the partition functions of the finite graphs used in the limit definition, viewed as polynomials in the parameter, are zero-free in a *complex* neighborhood of S. Thus, proving zero-freeness of partition functions of such a sequence of finite graphs in a *fixed* (i.e., independent of the finite graphs in question) complex neighborhood of S implies the absence of phase transitions in S.

While the algorithmic consequences of phase transitions defined in terms of decay of correlations have been well studied, first in the context of Markov chain Monte Carlo algorithms (Glauber dynamics) and more recently in determinstic algorithms that directly exploit correlation decay (see, e.g. [Wei06, BG08]), algorithmic use of the information on complex roots of the partition function originated only recently in the work of Barvinok (see [Bar17] for a survey). This has led to increased interest in understanding the relationship between the above two notions of phase transitions. Such connections have been the focus of some recent work on the independent set (or "hard-core lattice gas") model [PR17b] (see also [PR18]), while related ideas can be traced back to early work of Shearer [She85], as later elucidated by Scott and Sokal [SS04]. The main focus of this thesis is to push the study of these ideas further, leading both to new algorithmic applications of roots of partition functions and to a deeper understanding of their connection with correlation decay.

1.2 Organization of thesis

In this thesis, we will focus on the following two general directions:

- identifying the boundaries of computational tractability via the study of phase transitions from statistical physics;
- conversely, locating phase transitions in statistical physics via the analysis of algorithms.

More specifically, in chapter 2 we will explain, in the context of the ferromagnetic Ising model, that absence of phase transitions (complex zeros) implies efficient approximate counting algorithms. Next, in chapter 3, we study the connection between the two notions of phase transitions, and show that in many models correlation decay implies the absence of zeros; interestingly, we will crucially exploit the algorithmic analysis of correlation decay to prove absence of zeros. Finally, in chapter 4, we will push this connection further, and obtain state-of-the-art results on both the Fisher zeros of the Potts model and deterministic approximate counting algorithms for graph colorings.

The results in this thesis are derived in collaboration with Piyush Srivastava and Alistair Sinclair, and some have already been published. In particular, chapter 2 is based mainly on [LSS19c]; chapter 3 is derived from ideas in [LSS19b]; and chapter 4 is based largely on [LSS19a].

1.3 Notation

Throughout this thesis, we use ι to denote the imaginary unit $\sqrt{-1}$, in order to avoid confusion with the symbol "i" used for other purposes. For a complex number $z = a + \iota b$ with $a, b \in \mathbb{R}$, we denote its real part a as $\Re z$, its imaginary part b as $\Im z$, its *length* $\sqrt{a^2 + b^2}$ as $|z|$, and, when $z \neq 0$, its *argument* $\sin^{-1}\left(\frac{b}{|z|}\right) \in (-\pi, \pi]$ as $\arg z$. We also generalize the notation $[x, y]$ used for closed real intervals to the case when $x, y \in \mathbb{C}$, and use it to denote the closed straight line segment joining x and y.

Chapter 2

An algorithmic Lee-Yang theorem for the Ising model

In this chapter, we describe the algorithmic paradigm of Barvinok, which shows how to approximately evaluate a partition function by exploiting information about the locations of its complex zeros. As a concrete example, we will prove an algorithmic version of the classical Lee-Yang theorem for the ferromagnetic Ising model, which provides the first deterministic approximation algorithm for the partition function for almost all parameter values. In section 2.2 we first describe Barvinok's paradigm using the Lee-Yang theorem as a black box. Then, combining with ideas from Patel and Regts, we will show how this can be turned into an efficient approximation algorithm in section 2.3. Finally, we will conclude the chapter with a generalization of the Lee-Yang theorem to hypergraphs, improving on a classical result of Suzuki and Fisher [SF71].

2.1 Statements of results and technical overview

We recall the definition of the partition function of the Ising model: given an undirected graph $G = (V, E)$, a *vertex activity* or *fugacity* λ, that models an "external field" and determines the propensity of a vertex to be in the spin $+$, and an *edge activity* $\beta \geq 0$ that models the tendency of vertices to agree with their neighbors, the Ising partition function is defined as

$$Z_G^{\beta}(\lambda) := \sum_{S \subseteq V} \beta^{|E(S,\overline{S})|} \lambda^{|S|}. \tag{2.1}$$

In this chapter, we focus on the classical *ferromagnetic* case in which $\beta \leq 1$, so that configurations in which a larger number of neighboring spins agree (small cuts) have higher probability. The *anti-ferromagnetic* regime $\beta > 1$ is qualitatively very different, and prefers configurations with disagreements between neighbors. We note also that all our results in this chapter hold in the more general setting where there is a different interaction β_e on each edge, provided that all the β_e satisfy whatever constraints we are putting on β. (So, e.g., in the ferromagnetic case $\beta_e \leq 1$ for all e.) In addition, our results allow β to be negative and λ to be complex; this will be discussed in more detail below.

Throughout this chapter, as indicated by the notation in eq. (2.1), we view the Ising partition function as a polynomial in λ for a fixed β. Then, the classical Lee-Yang theorem states that, for any graph G and any $|\beta| \leq 1$ (corresponding to the ferromagnetic regime), the complex zeros of $Z_G^{\beta}(\lambda)$ lie on the unit circle $|\lambda| = 1$.

As in almost all statistical physics and graphical models, the partition function captures the computational complexity of the Ising model. Partition functions are #P-hard[1] to compute exactly in virtually any interesting case (e.g., this is true for the Ising model except in the trivial cases $\lambda = 0$ or $\beta \in \{0, 1\}$), so attention is focused on approximation. An early result in the field due to Jerrum and Sinclair [JS93] establishes a *fully polynomial randomized approximation scheme* for the Ising partition function, valid for all graphs G and all values of the parameters (β, λ) in the ferromagnetic regime. Like many of the first results on approximating partition functions, this algorithm is based on random sampling and the Markov chain Monte Carlo method.

For several statistical physics models on bounded degree graphs (including the anti-ferromagnetic Ising model [SST14, LLY13] and the "hard core", or independent set model [Wei06]), fully-polynomial *deterministic* approximation schemes have since been developed, based on the decay of correlations property in those models: roughly speaking, one can estimate the local contribution to the partition function at a given vertex v by exhaustive enumeration in a neighborhood around v, using decay of correlations to truncate the neighborhood at logarithmic diameter. The range of applicability of these algorithms is of course limited to the regime in which decay of correlations holds, and indeed com-

[1] If a combinatorial counting problem, such as computing a partition function in a statistical physics model, is #P-hard, then it can be solved in polynomial time only if *all* counting problems belonging to a very rich class can be solved in polynomial time. Hence #P-hardness is widely regarded as compelling evidence of the intractibility of efficient exact computation. For a more detailed account of this phenomenon in the context of partition functions, see, e.g., [SS14a, Appendix A].

plementary results prove that the partition function is NP-hard to approximate outside this regime [SS14b, GGŠ+14]. Perhaps surprisingly, however, no deterministic approximation algorithm is known for the classical ferromagnetic Ising partition function that works over anything close to the full range of the randomized algorithm of [JS93]: to the best of our knowledge, the best deterministic algorithm, due to Zhang, Liang and Bai [ZLB11], is based on correlation decay and is applicable to graphs of maximum degree Δ only when $\beta > 1 - 2/\Delta$.

The restricted applicability of correlation decay based algorithms for the ferromagnetic Ising model arises from two related reasons: the first is that this model does not exhibit correlation decay for β sufficiently close to 0 (for any given value of the external field), so any straightforward approach based only on this property cannot be expected to work for all β. Secondly, there is a regime of parameters for which, even though decay of correlation holds, there is evidence to believe that it cannot be exploited to give an algorithm using the usual techniques: see [SST14, Appendix 2] for a more detailed discussion of this point.

The first goal of this chapter is an algorithmic Lee-Yang theorem for graphs, which supplies the first deterministic algorithm that covers almost the entire range of parameters of the model except for the "zero-field" case $|\lambda| = 1$:

Theorem 2.1.1. *Fix any $\Delta > 0$. There is a fully polynomial time approximation scheme (FPTAS)[2] for the Ising partition function $Z_G^\beta(\lambda)$ in all graphs G of maximum degree Δ for all edge activities $-1 \leq \beta \leq 1$ and all (possibly complex) vertex activities λ with $|\lambda| \neq 1$.*

Remarks 1. (i) For fixed Δ and λ such that $|\lambda| < 1$, the running time of the FPTAS for producing a $(1 \pm \varepsilon)$-factor approximation on n-vertex graphs of degree at most Δ is $(n/\varepsilon)^{O\left(\frac{\log \Delta}{|1-|\lambda||}\right)}$. (The running times of the algorithms in Theorems 2.1.3 and 2.1.4 below have a similar dependence on λ and Δ.) Such dependence on the "distance to the critical boundary" (in this case, the circle $|\lambda| = 1$) of the degree of the polynomial bounding the running time of the FPTAS appears to be a common feature of algorithms based on correlation decay [Wei06, SSŠY16, LLY13] as well as our present analytic continuation approach. In contrast, approximate counting algorithms based on Markov chain Monte Carlo (e.g., [JS89, EHŠ+16, LV97]) often have the desirable feature that they are, in a sense, "fixed

[2] An FPTAS takes as input an n-vertex (hyper)graph G, model parameters β, λ, and an accuracy parameter $\varepsilon \in (0, 1)$ and outputs a value that approximates $Z_G^\beta(\lambda)$ within a factor $1 \pm \varepsilon$ (see also eq. (2.3)). The running time of the algorithm is polynomial in n and $1/\varepsilon$.

parameter tractable": even as the fixed parameters of the problem are moved close to the boundary up to which the algorithm is applicable, the degree of the polynomial bounding its running time does not increase; it is only the constant factors which increase to infinity. A similar phenomenon occurs in the case of the dependence of the exponent of the running time on the maximum degree of the graph: MCMC methods typically have no dependence, while both correlation decay and the methods used here have an exponent linear in $\log \Delta$. In the present case, this dependence seems to be inevitable since a crucial step in the algorithm is the enumeration of all connected sub-graphs of size roughly $\Theta(\log n)$, and the number of such sub-graphs may grow as $n^{\Theta(\log \Delta)}$ (see Section 2.3 and, in particular, the proof of Lemma 2.3.9). (ii) Note that although λ, β are positive in the "physically relevant" range in most applications of the Ising model, the above theorem also applies more generally to $\beta \in [-1, 1]$ and complex valued λ not on the unit circle. Moreover, we can allow edge-dependent activities β_e provided all of them lie in $[-1, 1]$. (iii) A result of Goldberg and Jerrum [GJ08, Lemmas 7 and 16] shows that if one can approximate the partition function of the Ising model at $\lambda = 1$ with edge-dependent activities $\beta_e \in [-1, 1]$ in polynomial time, then there is a deterministic polynomial time algorithm for approximately counting perfect matchings in general graphs. This leads to the following tantalizing possibility: extending Theorem 2.1.1 to the case $\lambda = 1$ (which lies at the boundary of the current range of applicability of the theorem) will lead to a deterministic FPTAS for counting perfect matchings in graphs, a problem that continues to remain wide open. (Note that if Theorem 2.1.1 applied to the case $\lambda = 1$ with edge dependent activities $\beta_e \in [-1, 1]$, then it would apply to unbounded degree graphs as well. This is because in this case, a high degree vertex can be replaced by a "comb" in which each edge has activity 0.)

The above theorem is actually a special case of a more general theorem for the hypergraph version of the Ising model (Theorem 2.1.3 below). We now illustrate our approach to proving these theorems, which will also allow us to introduce and motivate our further results in this chapter.

In contrast to previous algorithms based on correlation decay, our algorithm is based on isolating the complex zeros of the partition function $Z := Z_G^\beta(\lambda)$ (viewed as a polynomial in λ for a fixed value of β). This approach was introduced recently by Barvinok [Bar15a, Bar15b] (see also the recent monograph [Bar17] for a discussion of the approach in a more general context). We start with the observation that the $1 \pm \varepsilon$ multiplicative approximation of Z required for an FPTAS corresponds to a $O(\varepsilon)$ additive approximation of $\log Z$. Barvinok's approach considers a Taylor expansion of $\log Z$ around a point λ_0 such that $Z(\lambda_0)$ is easy to evaluate. (For the Ising model, $\lambda_0 = 0$ is such a choice.) It then seeks to evaluate the function at other points by carrying out an explicit analytic continuation. More concretely, suppose it can be shown that there are no zeros of Z in the open disk $D(\lambda_0, r)$ of radius r around λ_0. From standard results in complex analysis, it then follows that the Taylor expansion around λ_0 of $\log Z$ is absolutely convergent in $D(\lambda_0, r)$ and further, that the first m terms of this expansion evaluated at a point $\lambda \in D(\lambda_0, r)$ provide a $O\left(\frac{n\alpha^m}{1-\alpha}\right)$ additive approximation of $\log Z(\lambda)$, where $\alpha = |\lambda - \lambda_0|/r < 1$, and n is the degree of Z as a polynomial in λ. We note that Barvinok's approach may be seen as an algorithmic counterpart of the traditional view of phase transitions in statistical physics in terms of analyticity of $\log Z$ [YL52].

To apply this approach in the case of the ferromagnetic Ising model, we may appeal to the famous Lee-Yang theorem of the 1950s [LY52], which establishes that the zeros of $Z(\lambda)$ all lie on the unit circle in the complex plane. This allows us to take $\lambda_0 = 0$ and $r = 1$ in the previous paragraph, and thus approximate $Z(\lambda)$ at any point λ satisfying $|\lambda| < 1$. This extends to all λ with $|\lambda| \neq 1$ via the fact that $Z(\lambda) = \lambda^n Z(\frac{1}{\lambda})$.

Remark 2. We note that the case $|\lambda| = 1$ is likely to require additional ideas because it is known that there exist bounded degree graphs (namely, Δ-ary trees) for which the partition function $Z_G^\beta(\lambda)$ has complex zeros within distance $O(1/n)$ of $\lambda = 1$, where n is the size of the graph. In fact, the zeros are even known to become dense on the whole unit circle as n increases to infinity [BG01, BM97]. This precludes the possibility of efficiently carrying out the analytic continuation procedure for $\log Z$ to arbitrary points on the unit circle, and to the point $\lambda = 1$ in particular.

Converting the above approach into an algorithm requires computing the first k coefficients in the Taylor expansion of $\log Z$ around λ_0. Barvinok showed that this computation can in turn be reduced to computing the $O(k)$ lowest-degree coefficients of the partition function itself. In the case of the Ising model, computing k such coefficients is roughly analogous to computing k-wise correlations between the vertex spins, and doing so naively on a graph of n vertices requires time $\Omega(n^k)$. Until recently, no better ways to perform this calculation were known and hence approximation algorithms using this approach typically required quasi-polynomial time[3] in order to achieve a $(1 \pm 1/\operatorname{poly}(n))$-factor multiplicative approximation of Z (equivalently, a $1/\operatorname{poly}(n)$ additive approximation of $\log Z$), since this requires the Taylor series for $\log Z$ to be evaluated to $k = \Omega(\log n)$ terms [BS16a, BS16b, Bar15b].

Recently, Patel and Regts [PR17a] proposed a way to get around this barrier for several classes of partition functions. Their method is based on writing the coefficients in the Taylor series of $\log Z$ as linear combinations of counts of connected induced subgraphs of size up to $\Theta(\log n)$. This is already promising, since the number of connected induced subgraphs of size $O(\log n)$ of a graph G of maximum degree Δ is polynomial in the size of G when Δ is a fixed constant. Further, the count of induced copies of such a subgraph can also be computed in time polynomial in the size of G [BCKL13]. Patel and Regts built on these tools to show a way to compute the $\Theta(\log n)$ Taylor coefficients of $\log Z$ needed in Barvinok's approach for several families of partition functions, for some of which correlation decay based algorithms are still not known.

Unfortunately, for the case of the Ising model, it is not clear how to write the Taylor coefficients in terms of induced subgraph counts. Indeed, in their paper [PR17a, Theorem 1.4], Patel and Regts apply their method to the Ising model viewed as a polynomial in β rather than λ, which allows them to use subgraph counts. However, this prevents them from using the Lee-Yang theorem, and they are consequently able to establish only a small zero-free region. As a result, they can handle only the zero-field "high-temperature" regime, specifically the regime $|\beta - 1| \leq 0.34/\Delta$ and $\lambda = 1$. (Note that in fact the correlation decay property also holds in this regime.)

In this chapter, we instead propose a generalization of the framework of Patel and Regts to labelled hypergraphs via objects that we call *insects*. In the special case of graphs, an insect can be seen as a graph that includes edges to additional boundary vertices: we refer to Section 2.3.1 for precise definitions. Using the appropriate notions for counting induced sub-insects, we are then able to write the coefficients arising in the Taylor expansion of $\log Z$ for the Ising model in terms of induced sub-insect counts, and derive from there algorithms for computing $\Omega(\log n)$ such coefficients in polynomial time in graphs of bounded degree. This establishes Theorem 2.1.1. We note that if one is only interested in deriving Theorem 2.1.1, then this can also be done using the notion of *fragments*, developed

[3] A *quasi-polynomial time* algorithm is one which runs in time $\exp\{O((\log n)^c)\}$ for some constant $c > 1$.

by Patel and Regts [PR17a] in the different context of edge coloring models, which turns out to be a special case of our notion of insects.

Our framework of insects, however, also allows us to extend the above approach to edge-dependent activities and, more significantly, to the much more general setting where G is a hypergraph, as studied, for example, in the classical work of Suzuki and Fisher [SF71], and also more recently in the literature on approximate counting [GG16, LYZ16, SYZ16]. In a hypergraph of edge size $k \geq 3$, the pairwise interactions in the standard Ising model are replaced by higher-order interactions of order k. We note that the Jerrum-Sinclair MCMC approach [JS93] apparently does not extend to hypergraphs, and there is currently no known polynomial time approximation algorithm (even randomized) for a wide range of β in this setting. For a hypergraph $H = (V, E)$, configurations are again assignments of spins to the vertices V and the partition function $Z_H^\beta(\lambda)$ is defined exactly as in (2.1), where the cut $E(S, \overline{S})$ now consists of those hyperedges with at least one vertex in each of S and $\overline{S}$. The computation of coefficients via insects carries through as before, but the missing ingredient is an extension of the Lee-Yang theorem to hypergraphs. Suzuki and Fisher [SF71] prove a weak version of the Lee-Yang theorem for hypergraphs (see Theorem 2.4.3 in section 2.4), which we are able to strengthen to obtain the following optimal statement, which is of independent interest:

Theorem 2.1.2. *Let $H = (V, E)$ be a hypergraph with maximum hyperedge size $k \geq 3$. Then all the zeros of the Ising model partition function $Z_H^\beta(\lambda)$ lie on the unit circle if the edge activity β lies in the range $-\frac{1}{2^{k-1}-1} \leq \beta \leq \frac{1}{2^{k-1}\cos^{k-1}\left(\frac{\pi}{k-1}\right)+1}$. Further, when $\beta \neq 1$ does not lie in this range, there exists a hypergraph H with maximum hypergraph edge size at most k such that the zeros of the Ising partition function $Z_H^\beta(\lambda)$ of H do not lie on the unit circle.*

Remark 3. Once again, we can allow edge-dependent activities β_e provided all of them lie in the range stipulated above. This extension also applies to Theorem 2.1.3 below.

Note that the classical Lee-Yang theorem (for the graph case $k = 2$) establishes this property for $0 \leq \beta \leq 1$ (the ferromagnetic regime). Our proof of Theorem 2.1.2 follows along the lines of Asano's inductive proof of the Lee-Yang theorem [Asa70], but we need to carefully analyze the base case (where H consists of a single hyperedge) in order to obtain the above bounds on β. The optimality of the range of β in our result follows essentially from the fact that our analysis of the base case is tight. For a detailed comparison with the Suzuki-Fisher theorem, see the Remark following Corollary 2.4.5.

Combining Theorem 2.1.2 with our earlier algorithmic approach immediately yields the following generalization of Theorem 2.1.1 to hypergraphs.

Theorem 2.1.3. *Fix any $\Delta > 0$ and $k \geq 3$. There is an FPTAS for the Ising partition function $Z_H^\beta(\lambda)$ in all hypergraphs H of maximum degree Δ and maximum edge size k, for all edge activities β in the range of Theorem 2.1.2 and all vertex activities $|\lambda| \neq 1$.*

Finally, we extend our results to general ferromagnetic two-spin systems on hypergraphs, again as studied in [SF71]. A *two-spin system* on a hypergraph $H = (V, E)$ is specified by hyperedge activities $\varphi_e : \{+,-\}^{|e|} \to \mathbb{C}$ for $e \in E$, and a vertex activity $\psi : \{+,-\} \to \mathbb{C}$. (Note that we treat each hyperedge e as a set of vertices.) Then the partition function is defined as:

$$Z_H^{\boldsymbol{\varphi},\boldsymbol{\psi}} := \sum_{\boldsymbol{\sigma}:V\to\{+,-\}} \prod_{e\in E} \varphi_e\big(\boldsymbol{\sigma}\big|_e\big) \prod_{v\in V} \psi(\boldsymbol{\sigma}(v)).$$

Without loss of generality, we will henceforth assume that $\varphi_e(-,\cdots,-) = 1$, and that $\psi(-) = 1$, $\psi(+) = \lambda$. We can then write the partition function as

$$Z_H^{\boldsymbol{\varphi}}(\lambda) = \sum_{\boldsymbol{\sigma}:V\to\{+,-\}} \prod_{e\in E} \varphi_e\big(\boldsymbol{\sigma}\big|_e\big)\, \lambda^{|\{v:\boldsymbol{\sigma}(v)=+\}|}. \tag{2.2}$$

We call a hypergraph two-spin system *symmetric* if $\varphi_e(\boldsymbol{\sigma}) = \overline{\varphi_e(-\boldsymbol{\sigma})}$. Suzuki and Fisher [SF71] proved a Lee-Yang theorem for symmetric hypergraph two-spin systems (which is weaker than our Theorem 2.1.2 above when specialized to the Ising model). Combining this with our general algorithmic approach yields our final result of this chapter:

Theorem 2.1.4. *Fix any $\Delta > 0$ and $k \geq 2$ and a family of symmetric edge activities $\boldsymbol{\varphi} = \{\varphi_e\}$ satisfying $|\varphi_e(+,\cdots,+)| \geq \frac{1}{4}\sum_{\boldsymbol{\sigma}\in\{+,-\}^V} |\varphi_e(\boldsymbol{\sigma})|$. Then there exists an FPTAS for the partition function $Z_H^{\boldsymbol{\varphi}}(\lambda)$ of the corresponding symmetric hypergraph two-spin system in all hypergraphs H of maximum degree Δ and maximum edge size k for all vertex activities $\lambda \in \mathbb{C}$ such that $|\lambda| \neq 1$.*

2.2 Approximation of the log-partition function by Taylor series

In this section we present an approach due to Barvinok [Bar15b] for approximating the partition function of a physical system by truncating the Taylor series of its logarithm, as discussed above. We will work in our most general setting of symmetric two-spin systems on hypergraphs, which of course includes the Ising model (on graphs or hypergraphs) as a special case. As in (2.2), such a system has partition function

$$Z_H^{\boldsymbol{\varphi}}(\lambda) = \sum_{\boldsymbol{\sigma}:V\to\{+,-\}} \prod_{e\in E} \varphi_e\big(\boldsymbol{\sigma}\big|_e\big)\, \lambda^{|\{v:\boldsymbol{\sigma}(v)=+\}|}.$$

Our goal is an FPTAS for $Z_H^{\boldsymbol{\varphi}}(\lambda)$, i.e., a deterministic algorithm that, given as input H, $\{\varphi_e\}$, λ with $|\lambda| \neq 1$ and $\varepsilon \in (0,1]$, runs in time polynomial in $n = |H|$ and ε^{-1} and outputs a $(1 \pm \varepsilon)$-multiplicative approximation of $Z_H^{\boldsymbol{\varphi}}(\lambda)$, i.e., a number $\hat{Z}$ satisfying

$$|\hat{Z} - Z_H^{\boldsymbol{\varphi}}(\lambda)| \leq \varepsilon |Z_H^{\boldsymbol{\varphi}}(\lambda)|. \tag{2.3}$$

(Note that in our setting $\hat{Z}$ and $Z_H^{\boldsymbol{\varphi}}(\lambda)$ may be complex numbers.) By the symmetry $\varphi_e(\boldsymbol{\sigma}) = \overline{\varphi_e(-\boldsymbol{\sigma})}$, we also have $Z^{\boldsymbol{\varphi}}(\lambda) = \lambda^n Z^{\overline{\boldsymbol{\varphi}}}(\frac{1}{\lambda})$, so that without loss of generality we may assume $|\lambda| < 1$.

For fixed H and (hyper)edge activities $\boldsymbol{\varphi}$, we will write $Z(\lambda) = Z_H^{\boldsymbol{\varphi}}(\lambda)$ for short. Letting $f(\lambda) = \log Z(\lambda)$, using the Taylor expansion around $\lambda = 0$ we get

$$f(\lambda) = \sum_{j=0}^{\infty} f^{(j)}(0) \cdot \frac{\lambda^j}{j!}, \tag{2.4}$$

where $f(0) = \log Z(0) = 0$. Note that $Z = \exp(f)$, and thus an additive error in f translates to a multiplicative error in Z. More precisely, given $\varepsilon \leq 1/4$, and $f, \widetilde{f} \in \mathbb{C}$ such that $|f - \widetilde{f}| \leq \varepsilon$, we have

$$|\exp(\widetilde{f}) - \exp(f)| = |\exp(\widetilde{f} - f) - 1| \times |\exp(f)| \leq 4\varepsilon |\exp(f)|,$$

where the last inequality, valid for $\varepsilon \leq 1/4$, follows by elementary complex analysis. In other words, to achieve a multiplicative approximation of Z within a factor $1 \pm \varepsilon$, as required by an FPTAS, it suffices to obtain an additive approximation of f within $\varepsilon/4$.

To get an additive approximation of f, we use the first m terms in the Taylor expansion. Specifically, we compute $f_m(\lambda) := \sum_{j=0}^{m} f^{(j)}(0) \cdot \frac{\lambda^j}{j!}$. We show next how to compute the derivatives $f^{(j)}(0)$ from the derivatives of Z itself (which are more readily accessible).

To compute $f^{(j)}(0)$, note that $f'(\lambda) = \frac{1}{Z(\lambda)}\frac{\mathrm{d}Z(\lambda)}{\mathrm{d}\lambda}$, or $\frac{\mathrm{d}Z(\lambda)}{\mathrm{d}\lambda} = f'(\lambda)Z(\lambda)$. Thus for any $m \geq 1$,

$$\frac{\mathrm{d}^m}{\mathrm{d}\lambda^m}Z(\lambda) = \sum_{j=0}^{m-1}\binom{m-1}{j}\frac{\mathrm{d}^j}{\mathrm{d}\lambda^j}Z(\lambda)\cdot\frac{\mathrm{d}^{m-j}}{\mathrm{d}\lambda^{m-j}}f(\lambda). \tag{2.5}$$

Given $\frac{\mathrm{d}^j}{\mathrm{d}\lambda^j}Z(\lambda)\big|_{\lambda=0}$ for $j = 0, \cdots, m$, eq. (2.5) is a triangular system of linear equations in $\{f^{(j)}(0)\}_{j=1}^m$ of representation length poly(m), and is non-degenerate since $Z(0) = 1$; hence it can be solved in poly(m) time.

We can now specify the algorithm: first compute $\left\{\frac{\mathrm{d}^j}{\mathrm{d}\lambda^j}Z(\lambda)\big|_{\lambda=0}\right\}_{j=0}^m$; next, use the system in eq. (2.5) to solve for $\{f^{(j)}(0)\}_{j=1}^m$; and finally, compute and ouput the approximation $f_m(\lambda)$.

To quantify the approximation error in this algorithm, we need to study the locations of the complex roots $r_1, \cdots, r_n$ of Z. Throughout this chapter, we will be using (some variant of) the Lee-Yang theorem to argue that, for the range of interactions $\boldsymbol{\varphi}$ we are interested in, the roots r_i all lie on the unit circle in the complex plane, i.e., $|r_i| = 1$ for all i. Note that since we are assuming that $\varphi_e(-, \cdots, -) = 1$, the constant term $\prod_{i=1}^n(-r_i)$ of $Z(\lambda)$ is 1, and hence we have $Z(\lambda) = \prod_i(1 - \frac{\lambda}{r_i})$. The log partition function can then be written as

$$f(\lambda) = \log Z(\lambda) = \sum_{i=1}^n \log\left(1 - \frac{\lambda}{r_i}\right) = -\sum_{i=1}^n\sum_{j=1}^\infty \frac{1}{j}\left(\frac{\lambda}{r_i}\right)^j. \tag{2.6}$$

Note that due to the uniqueness of the Taylor expansion of meromorphic functions, the two power series expansions of $f(\lambda)$ in eqs. (2.4) and (2.6) are identical in the domain of their convergence. Denoting the first m terms of the above expansion by $f_m(\lambda) = -\sum_{i=1}^n\sum_{j=1}^m \frac{1}{j}(\frac{\lambda}{r_i})^j$, the error due to truncation is bounded by

$$|f(\lambda) - f_m(\lambda)| \leq n\sum_{j=m+1}^\infty \frac{|\lambda|^j}{j} \leq \frac{n\,|\lambda|^{m+1}}{(m+1)(1-|\lambda|)},$$

recalling that by symmetry we are assuming $|\lambda| < 1$. Thus to get within $\varepsilon/4$ additive error, it suffices to take $m \geq \frac{1}{\log(1/|\lambda|)}\left(\log(\frac{4n}{\varepsilon}) + \log(\frac{1}{1-|\lambda|})\right)$. The following result summarizes our discussion so far.

Lemma 2.2.1. *Given* $\varepsilon \in (0,1)$, $m \geq \frac{1}{\log(1/|\lambda|)}\left(\log(\frac{4n}{\varepsilon}) + \log(\frac{1}{1-|\lambda|})\right)$, *and the values of the first* m *derivatives* $\left\{\frac{\mathrm{d}^j}{\mathrm{d}\lambda^j} Z(\lambda)\big|_{\lambda=0}\right\}_{j=0}^{m}$, $f_m(\lambda)$ *can be computed in time* $\mathrm{poly}(n/\varepsilon)$. *Moreover, if the Lee-Yang theorem holds for the partition function* $Z(\lambda)$, *then* $|f_m(\lambda) - f(\lambda)| < \varepsilon/4$, *and thus* $\exp(f_m(\lambda))$ *approximates* $Z(\lambda)$ *within a multiplicative factor* $1 \pm \varepsilon$.

The missing ingredient in turning Lemma 2.2.1 into an FPTAS is the computation of the derivatives $\frac{\mathrm{d}^j}{\mathrm{d}\lambda^j} Z(\lambda)\big|_{\lambda=0}$ for $1 \leq j \leq m$, which themselves are just multiples of the $m+1$ lowest-degree coefficients of Z. Computing these values naively using the definition of $Z(\lambda)$ requires $n^{\Omega(m)}$ time. Since m is required to be of order $\Omega(\log(n/\varepsilon))$, this results in only a quasi-polynomial time algorithm. In the next section, we show how to compute these values in polynomial time when H is a hypergraph of bounded degree and bounded hyperedge size, which when combined with Lemma 2.2.1 gives an FPTAS.

2.3 Computing coefficients via insects

As discussed in section 2.1, Patel and Regts [PR17a] recently introduced a technique for efficiently computing the low-degree coefficients of a partition function using induced subgraph counts. In this section we introduce the notion of *sub-insect counts*, and show how it allows the Patel-Regts framework to be adapted to any hypergraph two-spin system with vertex activities (including the Ising model with vertex activities as a special case). We will align our notation with [PR17a] as much as possible. From now on, unless otherwise stated, we will use G to denote a hypergraph. Recall from section 2.1 the partition function of a two-spin system on a hypergraph $G = (V, E)$:

$$Z_G^{\boldsymbol{\varphi}}(\lambda) = \sum_{\boldsymbol{\sigma}: V \to \{+,-\}} \prod_{e \in E} \varphi_e\left(\boldsymbol{\sigma}\big|_e\right) \lambda^{|\{v : \boldsymbol{\sigma}(v) = +\}|}. \tag{2.7}$$

Due to the normalization $\varphi_e(-, \cdots, -) = 1$, each term in the summation depends only on the set $S = \{v : \boldsymbol{\sigma}(v) = +\}$ and the labelled induced sub-hypergraph $\left(S \cup \partial S, E[S] \cup E(S, \overline{S})\right)$, where $E[S]$ is the set of edges within S, ∂S is the boundary of S defined as $\partial S := \bigcup_{v \in S} N_G(v) \setminus S$, and $N_G(v)$ is the set of vertices adjacent to the vertex v in G. This fact motivates the induced sub-structures we will consider.

Let $\boldsymbol{\sigma}^S$ be the configuration where the set of vertices assigned $+$-spins is S, that is, $\boldsymbol{\sigma}^S(v) = +$ for $v \in S$ and $\boldsymbol{\sigma}^S(v) = -$ otherwise. We will also write $\varphi_e(S) := \varphi_e(\boldsymbol{\sigma}^S\big|_e)$

for simplicity. Thus the partition function can be written

$$Z_G^{\varphi}(\lambda) = \sum_{S\subseteq V} \prod_{e:e\cap S\neq\emptyset} \varphi_e(S)\lambda^{|S|}.$$

We start with the standard factorization of the partition function in terms of its complex zeros $r_1, \ldots, r_n$, where $n = |V|$. As explained in the paragraph preceding eq. (2.6), the assumption $\varphi_e(-, \cdots, -) = 1$ allows one to write the partition function as

$$Z_G^{\varphi}(\lambda) = \prod_{j=1}^{n}(1 - \lambda/r_j) = \sum_{i=0}^{n}(-1)^i e_i(G)\lambda^i,$$

where $e_i(G)$ is the elementary symmetric polynomial of degree i evaluated at $(\frac{1}{r_1}, \cdots, \frac{1}{r_n})$.

On the other hand, we can also express the coefficients $e_i(G)$ combinatorially using the definition of the partition function:

$$e_i(G) = (-1)^i \sum_{\substack{S\subseteq V\\|S|=i}} \prod_{e:e\cap S\neq\emptyset} \varphi_e(S). \tag{2.8}$$

Once we have computed the first m coefficients of Z (i.e., the values $e_i(G)$ for $i = 1, \cdots, m$), where $m = \Omega\Big(\frac{\log(n/\varepsilon)-\log(1-|\lambda|)}{\log(1/|\lambda|)}\Big)$, we can use Lemma 2.2.1 to obtain an FPTAS as claimed in Theorems 2.1.1, 2.1.3 and 2.1.4. The main result in this section will be an algorithm for computing these coefficients $e_i(G)$:

Theorem 2.3.1. *Fix $k, \Delta \in \mathbb{N}$ and $C > 0$. There exists a deterministic* $\mathrm{poly}(n/\varepsilon)$*-time algorithm that, given any n-vertex hypergraph G of maximum degree Δ and maximum hyperedge size k, and any $\varepsilon \in (0, 1)$, computes the coefficients $e_i(G)$ for $i = 1, \cdots, m$, where $m = \lceil C \log(n/\varepsilon)\rceil$.*

2.3.1 Insects in a hypergraph

To take advantage of the fact that each term in eq. (2.7) only depends on the set S and the induced sub-hypergraph $\big(S \cup \partial S, E[S] \cup E(S, \overline{S})\big)$, we define the following structure.

Definition 2.3.2. Given a vertex set S and a set E of hyperedges, $H = (S, E)$ is called an *insect* if for all $e \in E$, $e \cap S \neq \emptyset$. The set S is called the *label set* of the insect H and the set $B(H) := \big(\bigcup_{e\in E} e\big) \setminus S$ is called the *boundary set.*

Given an insect H, we use the notation $\underline{V}(H)$ for its label set. The *size* $|H|$ of the insect H is defined to be $|\underline{V}(H)|$. An insect $H = (S, E)$ is said to be *connected* if the hypergraph $(S, \{e \cap S \mid e \in E\})$ is connected. It is said to be *disconnected* otherwise. In the latter case, there exists a partition of S into non-empty sets S_1, S_2, and a partition of E into sets E_1 and E_2, such that (S_i, E_i) are insects for $i = 1, 2$, and the sets $S_2 \cap B(H_1)$ and $S_1 \cap B(H_2)$ are empty. In this case, we write $H = H_1 \uplus H_2$, and say that the insects H_1 and H_2 are *disjoint*. (Note that disjoint insects may share boundary vertices.)

Remark 4. Note that a hypergraph $G = (V, E)$ can itself be viewed as an insect. However, as is clear from the definition, not all insects are hypergraphs.

In order to exploit the structure of the terms in eq. (2.7) alluded to above, we now define the notion of an *induced sub-insect* of an insect. Given an insect $H = (S, E)$ and a subset S' of S, we define the induced sub-insect $H^+[S']$ as $(S', \{e \in E \mid e \cap S' \neq \emptyset\})$. Further, we say that an insect H is an induced sub-insect of an insect G, denoted $H \hookrightarrow G$, if there is a set $S \subseteq \underline{V}(G)$ such that $G^+[S] = H$.

2.3.2 Weighted sub-insect counts

Just as graph invariants may be expressed as sums over induced subgraph counts, we will consider weighted sub-insect counts of the form $f(G) = \sum_{S \subseteq \underline{V}(G)} a_{G^+[S]}$ and the functions f expressible in this way. Here G is any insect, and the coefficients a_H depend only on H, not on G.

Let $\mathcal{G}_t^{\Delta,k}$ be the set of insects up to size t, with maximum degree Δ and maximum hyperedge size k. Note that since insects are labelled, this is an infinite set. We will fix Δ and k throughout, and write $\mathcal{G} := \bigcup_{t \geq 1} \mathcal{G}_t^{\Delta,k}$. Let $\mathbf{1}[H \hookrightarrow G]$ be the indicator that H is an induced sub-insect of G, that is,

$$\mathbf{1}[H \hookrightarrow G] = 1 \text{ if there is a set } S \subseteq \underline{V}(G) \text{ such that } G^+[S] = H, \text{ and } 0 \text{ otherwise.}$$

A weighted sub-insect count $f(G)$ of the form considered above can then also be written as $f(G) = \sum_{H \in \mathcal{G}} a_H \cdot \mathbf{1}[H \hookrightarrow G]$. This alternative notation helps simplify the presentation of some of the combinatorial arguments below. Note that even though $\mathcal{G}$ is infinite, the above sum has only finitely many non-zero terms for any finite insect G. Further, as insects are labelled, $f(G)$ may also depend on the labelling of G, unlike a graph invariant where isomorphic copies of a graph yield the same value.

A weighted sub-insect count f is said to be *additive* if, given any two disjoint insects G_1 and G_2, $f(G_1 \uplus G_2) = f(G_1) + f(G_2)$. An argument due to Csikvári and Frenkel [CF16], also employed in the case of graph invariants by Patel and Regts [PR17a], can then be adapted to give the following:

Lemma 2.3.3. *Let f be a weighted sub-insect count, so that f may be written as*

$$f(G) = \sum_{H \in \mathcal{G}} a_H \cdot \mathbf{1}[H \hookrightarrow G]\,.$$

Then f is additive if and only if $a_H = 0$ for all insects H that are disconnected.

Proof. When H is connected, we have $\mathbf{1}[H \hookrightarrow G_1 \uplus G_2] = \mathbf{1}[H \hookrightarrow G_1] + \mathbf{1}[H \hookrightarrow G_2]$; thus f given in the above form is additive if $a_{H'} = 0$ for all H' that are not connected.

Conversely, suppose f is additive. By the last paragraph, we can assume without loss of generality that the sequence a_H is supported on disconnected insects (by subtracting the component of f supported on connected H). We now show that for such an f, a_H must be 0 for all disconnected H as well.

For if not, let H be a (necessarily disconnected) insect of smallest size for which $a_H \neq 0$. Since $a_J = 0$ for all insects J with $|J| < |H|$, we must have $f(J) = 0$ for all such insects. Also, since H is disconnected, there exist non-empty insects H_1 and H_2 such that $H = H_1 \uplus H_2$. By additivity, we then have $f(H) = f(H_1) + f(H_2) = 0$, where the last equality follows since both $|H_1|, |H_2|$ are strictly smaller than $|H|$. On the other hand, since H is an insect with the smallest possible number of vertices such that $a_H \neq 0$, we also have $f(H) = a_H \mathbf{1}[H \hookrightarrow H] = a_H$. This implies $a_H = 0$, which is a contradiction. Hence we must have $a_H = 0$ for all disconnected H. □

The next lemma implies that the product of weighted sub-insect counts can also be expressed as a weighted sub-insect count. We begin with a definition.

Definition 2.3.4. An insect $H_1 = (S_1, E_1)$ is *compatible* with another insect $H_2 = (S_2, E_2)$ if the insect $H := (S_1 \cup S_2, E_1 \cup E_2)$ satisfies $H^+[S_1] = H_1$ and $H^+[S_2] = H_2$.

Lemma 2.3.5. *Let $H_1 = (S_1, E_1), H_2 = (S_2, E_2)$ be arbitrary insects.*

(i) If H_1 and H_2 are not compatible, then there is no insect G such that $H_1 \hookrightarrow G$ and $H_2 \hookrightarrow G$. In other words, for every insect G,

$$\mathbf{1}[H_1 \hookrightarrow G]\,\mathbf{1}[H_2 \hookrightarrow G] = 0.$$

(ii) If H_1 and H_2 are compatible, then for every insect G,

$$\mathbf{1}[H_1 \hookrightarrow G]\,\mathbf{1}[H_2 \hookrightarrow G] = \mathbf{1}[H \hookrightarrow G]\,,$$

where H is the insect $(S_1 \cup S_2, E_1 \cup E_2)$, and satisfies $H^+[S_i] = H_i$ for $i = 1, 2$.

Proof. We start by making two observations. First, if $G^+[S_1] = H_1$ and $G^+[S_2] = H_2$ then $G^+[S_1 \cup S_2] = H = (S_1 \cup S_2, E_1 \cup E_2)$. Second, if $T \subseteq S \subseteq V$ and $G_1 := G^+[S]$ then $G_1^+[T] = G^+[T]$.

Suppose first that H_1 and H_2 are not compatible. Suppose, for the sake of contradiction, that there exists an insect G such that $G^+[S_i] = H_i$ for $i = 1, 2$. Then, from the first observation above we have $G^+[S_1 \cup S_2] = H = (S_1 \cup S_2, E_1 \cup E_2)$, while from the second observation we have $H^+[S_i] = G^+[S_i] = H_i$ for $i = 1, 2$. This contradicts the assumption that H_1 and H_2 are incompatible. Thus, we must have $\mathbf{1}[H_1 \hookrightarrow G]\,\mathbf{1}[H_2 \hookrightarrow G] = 0$ for every G, proving part (i).

Now suppose that H_1 and H_2 are compatible. As seen above, $G^+[S_i] = H_i$ for $i = 1, 2$ implies that $G^+[S_1 \cup S_2] = H$. On the other hand, if $G^+[S_1 \cup S_2] = H$, then by the compatibility of H_1 and H_2, and the second observation above, $G^+[S_i] = H^+[S_i] = H_i$ for $i = 1, 2$. This proves part (ii) of the lemma. □

An immediate corollary of the above lemma is that a product of weighted sub-insect counts is also a sub-insect count supported on slightly larger insects.

Corollary 2.3.6. *If $f_1(G) = \sum_H a_H \cdot \mathbf{1}[H \hookrightarrow G]$ and $f_2(G) = \sum_H b_H \cdot \mathbf{1}[H \hookrightarrow G]$ are weighted sub-insect counts, then so is $g(G) := f_1(G) f_2(G)$. Moreover, if f_1, f_2 are supported on sub-insects of sizes $\leq t_1, t_2$ respectively (i.e., if $a_H = 0$ when $|H| > t_1$ and $b_H = 0$ when $|H| > t_2$), then g is supported on sub-insects of size $\leq t_1 + t_2$.*

Proof. For compatible insects $H_i = (S_i, E_i)$ we denote by $H_1 \cup H_2$ the insect $(S_1 \cup S_2, E_1 \cup E_2)$. Now, for any insect G we have,

$$\begin{aligned} g(G) &= \sum_{H_1, H_2} a_{H_1} b_{H_2} \cdot \mathbf{1}[H_1 \hookrightarrow G] \cdot \mathbf{1}[H_2 \hookrightarrow G] \\ &= \sum_{H_1, H_2 \text{ compatible}} a_{H_1} b_{H_2} \cdot \mathbf{1}[H_1 \cup H_2 \hookrightarrow G] \\ &= \sum_{H} c_H \cdot \mathbf{1}[H \hookrightarrow G]\,, \end{aligned}$$

where in the second line we have used Lemma 2.3.5, and where

$$c_H := \sum_{\substack{H_1,H_2 \text{ compatible} \\ H=H_1\cup H_2}} a_{H_1} b_{H_2}. \tag{2.9}$$

Note that the number of non-zero terms in the definition of each c_H is finite, and that $|H_1 \cup H_2| \leq |H_1| + |H_2|$. This completes the proof. □

2.3.3 Enumerating connected sub-insects

We observe next that $e_i(G)$, as defined in eq. (2.8), can be written as a weighted sub-insect count. Accordingly, we generalize eq. (2.8) to arbitrary insects G of maximum degree Δ and hyperedge size k as follows:

$$e_i(G) = (-1)^i \sum_{\substack{S\subseteq \underline{V}(G) \\ |S|=i}} \prod_{e:e\cap S\neq\emptyset} \varphi_e(S) = \sum_{H\in\mathcal{G}_i^{\Delta,k}} \mu_H \cdot \mathbf{1}[H \hookrightarrow G], \tag{2.10}$$

where $\mu_H := (-1)^i \prod_{e:e\cap\underline{V}(H)\neq\emptyset} \varphi_e(\underline{V}(H))$. Note that this definition coincides with eq. (2.8) when G is a hypergraph, and also extends the definition of the partition function from hypergraphs to insects via the equation $Z_G(\lambda) = \sum_{i=0}^{|G|}(-1)^i e_i(G)\lambda^i$; when $G = (S, E)$ this definition is equivalent to that of the partition function on the hypergraph $(S \cup B(G), E)$, with the vertices in $B(G)$ set to '$-$'. This latter observation implies that when the insect G is disconnected and $G = G_1 \uplus G_2$, we have $Z_G(\lambda) = Z_{G_1}(\lambda)Z_{G_2}(\lambda)$.

We now consider the computational properties of the above expansion. Note that each coefficient μ_H is readily computable in time poly($|H|$); however, as discussed in section 2.1, the number of $H \in \mathcal{G}_i^{\Delta,k}$ such that $\mathbf{1}[H \hookrightarrow G] \neq 0$ is $\Omega(n^i)$, so that a naive computation of $e_i(G)$ using eq. (2.10) would be inefficient. To prove Theorem 2.3.1, we consider the set of *connected* insects, denoted by $\mathcal{C}_i^{\Delta,k}$, rather than $\mathcal{G}_i^{\Delta,k}$. We will show in this subsection that $\mathcal{C}_i^{\Delta,k}$ can be efficiently enumerated, and then in the following subsection reduce the above summation over $\mathcal{G}_i^{\Delta,k}$ to enumerations of $\mathcal{C}_i^{\Delta,k}$.

As in [PR17a], we use the following calculation of Borgs et al. [BCKL13, Lemma 2.1 (c)].

Lemma 2.3.7. *Let G be a multigraph with maximum degree Δ (counting edge multiplicity) and let v be a vertex of G. Then the number of subtrees of G with t vertices containing the vertex v is at most $\frac{(e\Delta)^{t-1}}{2}$.*

Proof. Consider the infinite rooted Δ-ary tree T_Δ. The number of subtrees with t vertices starting from the root is $\frac{1}{t}\binom{t\Delta}{t-1} < \frac{(e\Delta)^{t-1}}{2}$. (See also [SF99, Theorem 5.3.10].) The proof is completed by observing that the set of t-vertex subtrees of G containing vertex v can be mapped injectively into subtrees of T_Δ containing the root. $\square$

Corollary 2.3.8. *Let G be a hypergraph with maximum degree Δ and maximum hyperedge size k, and let $v \in V(G)$. Then the number of connected induced sub-insects of G of size t whose label set contains the vertex v is at most $\frac{(e\Delta k)^{t-1}}{2}$.*

Proof. Consider the multigraph H obtained by replacing every hyperedge of size r in G by an r-clique. For any connected induced sub-insect A of G, the label set $\underline{V}(A)$ is connected in H. Now, for any two distinct connected induced sub-insects A and B, let S_A and S_B be the sets of trees in H that span the label sets $\underline{V}(A)$ and $\underline{V}(B)$ of A and B respectively. Since the label sets of A and B are different, we must have $S_A \cap S_B = \emptyset$. Thus the number of connected subtrees on t vertices in H which contain the vertex v is an upper bound on the number of connected induced sub-insects in G whose label set contains v.

Finally, in the multigraph H the maximum degree is Δk, so by Lemma 2.3.7 the number of such subtrees is at most $\frac{(e\Delta k)^{t-1}}{2}$. $\square$

As a consequence we can efficiently enumerate all connected induced sub-insects of logarithmic size in a bounded degree graph. This follows from a similar reduction to a multigraph, applying [PR17a, Lemma 3.4]. However, for the sake of completeness we also include a direct proof.

Lemma 2.3.9. *For a hypergraph G of maximum degree Δ and maximum hyperedge size k, there exists an algorithm that enumerates all connected induced sub-insects of size at most t in G and runs in time $\widetilde{O}(nt^3(e\Delta k)^{t+1})$. Here $\widetilde{O}$ hides factors of the form* $\mathrm{polylog}(n)$, $\mathrm{polylog}(\Delta k)$ *and* $\mathrm{polylog}(t)$.

Proof. Let $\mathcal{T}_t$ be the set of $S \subseteq V(G)$ such that $|S| \leq t$ and $G^+[S]$ is connected. Note that given $S \in \mathcal{T}_t$, $G^+[S]$ will be a sub-insect of size t, and this clearly enumerates all of them. Also, by Corollary 2.3.8, $|\mathcal{T}_t| \leq O(n(e\Delta k)^t)$. Thus it remains to give an algorithm to construct $\mathcal{T}_t$ in about the same amount of time.

We construct $\mathcal{T}_t$ inductively. For $t = 1$, $\mathcal{T}_1 := V(G)$. Then given $\mathcal{T}_{t-1}$, define the multiset

$$\mathcal{T}_t^* := \mathcal{T}_{t-1} \cup \{S \cup \{v\} : S \in \mathcal{T}_{t-1} \text{ and } v \in N_G(S) \setminus S\}.$$

Since $|N_G(S)| < t\Delta k$, we can compute the set $N_G(S) \setminus S$ in time $O(t\Delta k)$, and construct $\mathcal{T}_t^*$ in time $\tilde{O}(|\mathcal{T}_{t-1}|\, t^2\Delta k) = \tilde{O}(nt^2(e\Delta k)^t)$. Finally, we remove duplicates in $\mathcal{T}_t^*$ to get $\mathcal{T}_t$

(e.g., by sorting the sets $S \in \mathcal{T}_t^*$, where each is represented as a string of length $\widetilde{O}(t)$), in time $\widetilde{O}(nt^3(e\Delta k)^t)$.

Starting from $\mathcal{T}_1$, inductively we perform t iterations to construct $\mathcal{T}_t$. Thus the overall running time is $\sum_{i=1}^{t} \widetilde{O}(ni^3(e\Delta k)^i) = \widetilde{O}(nt^3(e\Delta k)^{t+1})$. □

2.3.4 Proof of Theorem 2.3.1

The results in the previous subsection allow us to efficiently enumerate connected sub-insects. To prove Theorem 2.3.1, it remains to reduce the sum over all (possibly disconnected) H in eq. (2.10) to a sum over *connected* H. We now show that the method of doing so using Newton's identities and the multiplicativity of the partition function developed by Patel and Regts [PR17a] for graphs extends to the case of insects. Let G be any insect of size n and consider the t-th power sum of the inverses of the roots r_i, $1 \le i \le n$, of $Z_G(\lambda)$ (extended to insects G as in the paragraph following eq. (2.10)):

$$p_t(G) = \sum_{i=1}^{n} \frac{1}{r_i^t}.$$

Now by Newton's identities (which relate power sums to elementary symmetric polynomials), we have

$$p_t = \sum_{i=1}^{t-1} (-1)^{i-1} p_{t-i} e_i + (-1)^{t-1} t e_t. \tag{2.11}$$

Recall from eq. (2.10) that e_i is a weighted sub-insect count supported on insects of size $\le i$, and also from Corollary 2.3.6 that the product of two weighted sub-insect counts supported on insects of size $\le t_i, t_2$ respectively is a weighted sub-insect count supported on insects of size $\le t_1 + t_2$. Therefore, by eq. (2.11) and induction, each p_t is also a weighted sub-insect count supported on insects of size $\le t$. Thus, for any insect G, we may write

$$p_t(G) = \sum_{H \in \mathcal{G}_t^{\Delta,k}} a_H^{(t)} \cdot \mathbf{1}[H \hookrightarrow G] \tag{2.12}$$

for some coefficients $a_H^{(t)}$ to be determined. (The superscript (t) reflects the fact that a given H will in general have different coefficients for different p_t.)

Recall now that if G is disconnected with $G = G_1 \uplus G_2$ then $Z_G(\lambda) = Z_{G_1}(\lambda) \cdot Z_{G_2}(\lambda)$. Thus, the polynomials $Z_G(\lambda)$ are multiplicative over G, and hence sums of powers of their roots, such as $p_t(G)$ are additive: $p_t(G_1 \uplus G_2) = p_t(G_1) + p_t(G_2)$. Hence by

Lemma 2.3.3, the coefficients of p_t are supported on *connected* insects, and we may write eq. (2.12) as

$$p_t(G) = \sum_{H \in \mathcal{C}_t^{\Delta,k}} a_H^{(t)} \cdot \mathbf{1}[H \hookrightarrow G]\,. \tag{2.13}$$

Notice that by Corollary 2.3.8, there are at most $n(e\Delta k)^t$ non-zero terms in this sum.

Lemma 2.3.10. *There is a* $\mathrm{poly}(n/\varepsilon)$*-time algorithm to compute all the coefficients* $a_H^{(t)}$ *in eq.* (2.13)*, for* $t \le O(\log(n/\varepsilon))$*.*

Proof. By Lemma 2.3.9, we compute $\mathcal{T}_t$, consisting of all $S \subseteq V(G)$ such that $|S| \le t$ and $G^+[S]$ is connected. As we have removed duplicates, this is exactly $\mathcal{C}_t^{\Delta,k}$. We then use dynamic programming to compute the coefficients $a_H^{(t)}$.

By eq. (2.11), for $t = 1$ we have $p_1 = e_1$, so by eq. (2.10) we can read off the coefficients $a_H^{(1)}$ from $e_1(G)$. Next suppose we have computed $a_{H'}^{(t')}$ for $|H'| \le t' < t$, and we want to compute $a_H^{(t)}$ for some fixed connected $H \in \mathcal{C}_t^{\Delta,k}$ such that $\mathbf{1}[H \hookrightarrow G]$. Again by eq. (2.11), it suffices to compute the coefficient corresponding to H in $p_{t-i}e_i$ for each $1 \le i \le k-1$ (since the contribution of the last term in eq. (2.11) is simply $(-1)^{t-1}t\mu_H$ if $|H| = t$ and 0 otherwise). By eqs. (2.9) and (2.13), this coefficient is given by

$$\sum_{\substack{H_1 \in \mathcal{G}_i^{\Delta,k},\, H_2 \in \mathcal{C}_{(t-i)}^{\Delta,k} \\ H_1 \text{ compatible with } H_2 \\ H_1 \cup H_2 = H}} a_{H_2}^{(t-i)} \mu_{H_1}^{(i)} = \sum_{\substack{(S_1,S_2) \\ S_1 \cup S_2 = \underline{V}(H) \\ H^+[S_1] \in \mathcal{G}_i^{\Delta,k},\, H^+[S_2] \in \mathcal{C}_{(t-i)}^{\Delta,k}}} a_{H^+[S_2]}^{(t-i)} \mu_{H^+[S_1]}\,. \tag{2.14}$$

Since $t \le O(\log(n/\varepsilon))$, the second sum involves at most $4^t = \mathrm{poly}(n/\varepsilon)$ terms. Moreover, due to Corollary 2.3.8, there are at most $nt(e\Delta k)^t = \mathrm{poly}(n/\varepsilon)$ previously computed $a_{H'}^{(t')}$, where H' is a connected sub-insect of G and $|H'| \le t' < t$. In order to look up $a_{H^+[S]}^{(t-i)}$, one can do a linear scan, which also takes time $\mathrm{poly}(n/\varepsilon)$ for $t \le O(\log(n/\varepsilon))$. The coefficients $\mu_{H^+[S]}$ can simply be read off from their definition in eq. (2.10).

To conclude, because $t \le O(\log(n/\varepsilon))$, eq. (2.13) only contains $\mathrm{poly}(n/\varepsilon)$ terms. And for each term, $a_H^{(t)}$ can be computed using the above dynamic programming scheme in $\mathrm{poly}(n/\varepsilon)$ time. $\square$

Finally, now that we can compute $a_{H,t}$ efficiently, by eq. (2.13) we can compute p_k using the sub-insect enumerator in Lemma 2.3.9, and we can then compute e_k using Newton's identities as in eq. (2.11), which completes the proof of Theorem 2.3.1.

2.3.5 Proofs of main theorems

Our first main result in section 2.1, the FPTAS for the Ising model on graphs throughout the ferromagnetic regime with non-zero field stated in Theorem 2.1.1, now follows by combining Theorem 2.3.1 with Lemma 2.2.1 and the Lee-Yang theorem [LY52] (also stated as Theorem 2.4.2 in the next section). Recall from section 2.1 that the Lee-Yang theorem ensures that the partition function has no zeros inside the unit disk.

Similarly, Theorem 2.1.4, the FPTAS for two-spin systems on hypergraphs, follows by combining Theorem 2.3.1 with Lemma 2.2.1 and the Suzuki-Fisher theorem [SF71] (also stated as Theorem 2.4.3 in the next section). Again, the Suzuki-Fisher theorem ensures that there are no zeros inside the unit disk, under the condition on the hyperedge activities stated in Theorem 2.1.4.

To establish our final main algorithmic result, Theorem 2.1.3, we first need to prove a new Lee-Yang theorem for the hypergraph Ising model as stated in Theorem 2.1.2 in section 2.1. This will be the content of the next section. Once we have that, Theorem 2.1.3 follows immediately by the same route as above.

2.4 A Lee-Yang Theorem for Hypergraphs

In this section we prove a tight Lee-Yang theorem for the hypergraph Ising model (Theorem 2.1.2 in section 2.1). We start by extending the definition of the hypergraph Ising model to the multivariate setting, where each vertex and each hyperedge is allowed to have a different activity. As before, we have an underlying hypergraph $G = (V, E)$ with $|V| = n$ vertices. Given vertex activities $\lambda_1, \lambda_2, \ldots, \lambda_n$ and hyperedge activities $\boldsymbol{\beta} = (\beta_e)$, we define

$$Z_G^{\boldsymbol{\beta}}(\lambda_1, \cdots, \lambda_n) = \sum_{S \subseteq V} \prod_{e \in E\left(S, \overline{S}\right)} \beta_e \prod_{i \in S} \lambda_i \,,$$

where for a subset $S \subseteq V$, $E(S, \overline{S})$ is the set of hyperedges with at least one vertex in each of S and $\overline{S}$. Note that

$$Z_G^{\boldsymbol{\beta}}(\lambda_1, \cdots, \lambda_n) = \prod_{i=1}^{n} \lambda_i \cdot Z_G^{\boldsymbol{\beta}}\left(\frac{1}{\lambda_1}, \cdots, \frac{1}{\lambda_n}\right). \tag{2.15}$$

We use the following definition of the Lee-Yang property. This definition is based on the results of Asano [Asa70] and Suzuki and Fisher [SF71], and somewhat stricter than the definition used by Ruelle [Rue10].

Definition 2.4.1 (Lee-Yang property). Let $P(z_1, z_2, \ldots, z_n)$ be a multilinear polynomial. P is said to have the *Lee-Yang property* (sometimes written as "P is LY") if for any complex numbers $\lambda_1, \cdots, \lambda_n$ such that $|\lambda_1| \geq 1, \cdots, |\lambda_n| \geq 1$, and $|\lambda_i| > 1$ for some i, it holds that $P(\lambda_1, \cdots, \lambda_n) \neq 0$.

Then the seminal Lee-Yang theorem [LY52] can be stated as follows:

Theorem 2.4.2. *Let G be a connected undirected graph, and suppose $0 < \beta < 1$. Then the Ising partition function $Z_G^\beta(\lambda_1, \cdots, \lambda_n)$ has the Lee-Yang property.*

The following extension of the Lee-Yang theorem to general symmetric two-spin systems on hypergraphs is due to Suzuki and Fisher [SF71]. Again the theorem is stated in the multivariate setting, where in the two-spin partition function in eq. (2.7) each vertex i has a distinct activity λ_i.

Theorem 2.4.3. *Consider any symmetric hypergraph two-spin system, with a connected hypergraph G and edge activities $\{\varphi_e\}$. Then the partition function $Z_G^{\boldsymbol{\varphi}}(\lambda_1, \cdots, \lambda_n)$ has the Lee-Yang property if $|\varphi_e(+, \cdots, +)| \geq \frac{1}{4} \sum_{\boldsymbol{\sigma} \in \{+,-\}^V} |\varphi_e(\boldsymbol{\sigma})|$ for every hyperedge e.*

Theorem 2.4.3 is not tight for the important special case of the Ising model on hypergraphs. Our goal in this section is to prove a tight analog of the original Lee-Yang theorem for this case. Specifically, we will prove the following:

Theorem 2.4.4. *Let $G = (V, E)$ be a connected hypergraph, and $\boldsymbol{\beta} = (\beta_e)_{e \in E}$ be a vector of real valued hyperedge activities so that the activity of edge $e \in E$ is β_e. Then $Z_G^{\boldsymbol{\beta}}$ has the Lee-Yang property if the following condition holds for every hyperedge e, where $k \geq 2$ is the size of e:*

- *if $k = 2$, then $-1 < \beta_e < 1$;*
- *if $k \geq 3$, then $-\frac{1}{2^{k-1}-1} < \beta_e < \frac{1}{2^{k-1}\cos^{k-1}\left(\frac{\pi}{k-1}\right)+1}$.*

Further, if the above condition is not satisfied for a given real edge activity β and integer $k \geq 2$, then there exists a k-uniform hypergraph H with edge activity β such that Z_H^β does not have the Lee-Yang property.

Note that the case $k = 2$ is just the original Lee-Yang theorem (Theorem 2.4.2).

The following corollary for the univariate polynomial $Z_G^{\boldsymbol{\beta}}(\lambda)$ follows immediately via eq. (2.15) and the fact that, by Hurwitz's theorem, the zeros of $Z_G^{\boldsymbol{\beta}}(\lambda)$ are continuous functions of $\boldsymbol{\beta}$ and thus remain on the unit circle after taking the limit in the range of each β_e.

Corollary 2.4.5. *Let $G = (V, E)$ be a connected hypergraph, and $\boldsymbol{\beta} = (\beta_e)_{e \in E}$ be the vector of real valued hyperedge activities so that the activity of edge $e \in E$ is β_e. Then, all complex zeros of the univariate partition function $Z_G^{\boldsymbol{\beta}}(\lambda)$ lie on the unit circle if the following condition holds for every hyperedge e, where $k \geq 2$ is the size of e:*

- *if $k = 2$, then $-1 \leq \beta_e \leq 1$;*
- *if $k \geq 3$, then $-\frac{1}{2^{k-1}-1} \leq \beta_e \leq \frac{1}{2^{k-1}\cos^{k-1}\left(\frac{\pi}{k-1}\right)+1}$.*

The corollary establishes the first part of Theorem 2.1.2 in section 2.1, and hence also Theorem 2.1.3 as explained at the end of the previous section. The second part of Theorem 2.1.2, which asserts that the range of edge activities under which the theorem holds is optimal, is proven in Section 2.4.1. (Note that the optimality for the univariate case claimed in Theorem 2.1.2 does not directly follow from the optimality for the multivariate case guaranteed by Theorem 2.4.4 above.)

Remark 5. As a comparison, the result of Suzuki and Fisher, which we restated in Theorem 2.4.3, implies that a sufficient condition for the Lee-Yang property of $Z_G^{\boldsymbol{\beta}}(\lambda)$ is

$$-\frac{1}{2^{k-1}-1} \leq \beta_e \leq \frac{1}{2^{k-1}-1}.$$

Note that while the lower bound on β_e is the same as ours, our (tight) upper bound is always better, and significantly so for the more interesting case of small k. For example, for $k = 3$ our result gives the optimal range $-\frac{1}{3} \leq \beta_e \leq 1$, while the Suzuki-Fisher theorem gives $-\frac{1}{3} \leq \beta_e \leq \frac{1}{3}$. Similarly, for $k = 4$ the respective ranges are $[-1/7, 1/2]$ (for ours) and $[-1/7, 1/7]$ (for Suzuki-Fisher). We note here that there is a combinatorial explanation for the fact that for positive β one gets the same range for $k = 3$ as that for the case of graphs ($k = 2$): a hyperedge of size three with activity β^2 is equivalent to a clique on three vertices in which each edge has activity β. Such constructions however do not work for $k \geq 4$: the special nature of $k = 3$ comes from the fact that in any configuration of a hyperedge on three vertices, at least two vertices have the same spin.

We turn now to the proof of Theorem 2.4.4. The main technical step in our proof is to derive conditions under which the Ising partition function of a hypergraph consisting of a *single hyperedge* has the Lee-Yang property. This "base case" turns out to be more difficult than in the case of the original Lee-Yang theorem for graphs. However, as in the graph case, it will turn out that the base case still determines the range of β in which the theorem can be claimed to be valid; we show this latter claim, which implies the second part of Theorem 2.4.4 in Section 2.4.1.

We begin with the following two lemmas which, taken together, give a partial characterization of the Lee-Yang property. While similar in spirit to the results of Ruelle [Rue10], these lemmas do not follow directly from those results since, as noted above, the version of the Lee-Yang property used here imposes a stricter condition on the polynomial than does the definition used in [Rue10].

Lemma 2.4.6. *Given a multilinear polynomial $P(z_1, z_2, \ldots, z_n)$ with real coefficients define, for each $1 \leq j \leq n$, multilinear polynomials A_j and B_j in the variables $z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_n$ such that*

$$P = A_j z_j + B_j.$$

If P has the Lee-Yang property then, for every j such that the variable z_j has positive degree in P, it holds that $A_j(z_1, \ldots, z_{j-1}, z_{j+1}, \ldots, z_n) \neq 0$ when $|z_i| \geq 1$ for all $i \neq j$. In particular, A_j itself is LY.

Proof. Without loss of generality, we assume that $j = 1$. Note that since z_1 has positive degree in P, A_1 is a non-zero polynomial. Suppose that, in contradiction to the claim of the lemma, there exist complex numbers $\lambda_2, \ldots, \lambda_n$ satisfying $|\lambda_i| \geq 1$ such that $A_1(\lambda_2, \ldots, \lambda_n) = 0$. Since P is LY, it follows that $B_1(\lambda_2, \ldots, \lambda_n) \neq 0$ (for otherwise, we get a contradiction to the Lee-Yang property by choosing z_1 to be an arbitrary value outside the closed unit disk).

By continuity, this implies that $|B_1|$ is positive in any small enough neighborhood of $(\lambda_2, \ldots, \lambda_n)$ in $\mathbb{C}^{n-1}$. In particular, let S_ε be the open set

$$S_\varepsilon := \{(y_2, \ldots, y_n) \mid |y_i - \lambda_i| < \varepsilon \text{ and } |y_i| > 1 \text{ for } 2 \leq i \leq n \}.$$

Then there exist positive δ_0 and ε_0 such that $|B_1|$ is at least δ_0 in the open set S_ε when $\varepsilon < \varepsilon_0$.

Now, since A_1 is a non-zero multilinear polynomial, it cannot vanish identically on any open set. In particular, it cannot vanish identically in S_ε for any $\varepsilon > 0$. On the other hand, since A_1 vanishes at $(\lambda_2, \ldots, \lambda_n)$ it follows from continuity that for $\varepsilon < \varepsilon_0$ small enough, $|A_1| \le \delta_0/2$ in S_ε. Since A_1 does not vanish identically on S_ε, there must exist a point $(y_2, \ldots, y_n)$ in S_ε such that $0 < |A_1(y_2, \ldots, y_n)| < \delta_0/2$. Since $|B_1(y_2, \ldots, y_n)| \ge \delta_0$ by the choice of ε_0, it follows that if we define $y_1 = -B_1(y_2, \ldots, y_n)/A_1(y_2, \ldots, y_n)$ then $2 < |y_1| < \infty$. However, we then have $P(y_1, y_2, \ldots, y_n) = 0$ even though $|y_1| > 1$ and $|y_i| \ge 1$ for all i. This contradicts the Lee-Yang property of P. $\square$

By iterating the above lemma, we get the following corollary.

Corollary 2.4.7. *Let $P(z_1, z_2, \ldots, z_n)$ be a multilinear polynomial with non-zero real coefficients, i.e.,*

$$P(z_1, \ldots, z_n) = \sum_{S \subseteq [n]} p_S \prod_{i \in S} z_i,$$

where $p_S \in \mathbb{R}$ are non-zero for all $S \subseteq [n]$, and assume that P is LY. Then, for every subset S of $[n]$, the polynomial A_S defined by the equation

$$P(z_1, \ldots, z_n) = A_S((z_i)_{i \notin S}) \prod_{i \in S} z_i + \sum_{T:\, S \not\subseteq T} p_T \prod_{i \in T} z_i$$

has the property that $A_S((z_i)_{i \notin S}) \neq 0$ when $|z_i| \ge 1$ for all $i \notin S$. In particular, A_S is LY.

We next show that Lemma 2.4.6 has a partial converse for symmetric multilinear functions.

Lemma 2.4.8. *Let $P(z_1, z_2, \ldots, z_n)$ be a symmetric multilinear polynomial with non-zero real coefficients, i.e.,*

$$P(z_1, \ldots, z_n) = \sum_{S \subseteq [n]} p_S \prod_{i \in S} z_i,$$

where $p_S \neq 0$ for all $S \subseteq [n]$ and $p_S = p_{\overline{S}}$. Assume further that the polynomials A_j as defined in Lemma 2.4.6 all have the property that they are non-zero when all their arguments z_i satisfy $|z_i| \ge 1$. Then P is LY.

Proof. We first show that, under our assumptions, if all but one of the z_j lie on the unit circle, then P can only vanish if the remaining z_j is also on the unit circle. Without loss of generality we set $j = 1$, that is, we will show that if $|z_i| = 1$ for $i \ge 2$, then any root $z_1 = \zeta_1$ of the equation $A_1 z_1 + B_1 = 0$ satisfies $|\zeta_1| = 1$. (Here A_1 and B_1 are in the notation of Lemma 2.4.6.)

Since by assumption $A_1 = \sum_{S\subseteq[2,n]} p_{S\cup\{1\}} \prod_{i\in S} z_i$ does not vanish with this setting of the z_i, we have

$$|\zeta_1| = \left|\frac{B_1}{A_1}\right| = \left|\frac{\sum_{S\subseteq[2,n]} p_S \prod_{i\in S} z_i}{\sum_{S\subseteq[2,n]} p_{S\cup\{1\}} \prod_{i\in S} z_i}\right| = \left|\left(\prod_{i\in[2,n]} z_i\right) \frac{\sum_{S\subseteq[2,n]} p_S \prod_{\substack{i\notin S\\ i\neq 1}}(1/z_i)}{\sum_{S\subseteq[2,n]} p_{S\cup\{1\}} \prod_{i\in S} z_i}\right|$$
$$\overset{(\star)}{=} \left|\frac{\sum_{S\subseteq[2,n]} p_S \prod_{\substack{i\notin S\\ i\neq 1}} \overline{z_i}}{\sum_{S\subseteq[2,n]} p_{S\cup\{1\}} \prod_{i\in S} z_i}\right| \overset{(\dagger)}{=} \left|\frac{\sum_{S\subseteq[2,n]} p_{S\cup\{1\}} \prod_{i\in S} \overline{z_i}}{\sum_{S\subseteq[2,n]} p_{S\cup\{1\}} \prod_{i\in S} z_i}\right| = 1. \tag{2.16}$$

Here $(\star)$ uses the fact that $|z_i| = 1$ for $i \geq 2$ and $(\dagger)$ uses the symmetry of P. We have thus shown that if $(z_1, z_2, \ldots, z_n)$ is a zero of P such that $|z_i| \geq 1$ for all i then it is impossible for only one z_i to lie outside the closed unit disk.

We now show that if there are $k \geq 2$ values of i for which z_i lies outside the closed unit disk, then we can find another zero $(\zeta_1, \zeta_2, \zeta_3, \ldots, \zeta_n)$ of P such that $|\zeta_i| \geq 1$ for all i, and exactly $k-1$ of the ζ_i lie outside the closed unit disk. We can then iterate this process to reduce k to 1, in which case the observation in the previous paragraph leads to a contradiction.

By re-numbering the indices if needed, we can assume that $|z_1|, |z_2| > 1$ and $|z_i| \geq 1$ for $i \geq 3$. We can then write

$$P(z_1, \ldots, z_n) = \alpha_{12} z_1 z_2 + \alpha_1 z_1 + \alpha_2 z_2 + \alpha_\emptyset,$$

where $\alpha_{12}, \alpha_1, \alpha_2$ and $\alpha_\emptyset$ are non-zero polynomials in $z_3, \ldots, z_n$. Further, the hypotheses of the lemma imply that $A_1 = \alpha_{12} z_2 + \alpha_1$ and $A_2 = \alpha_{12} z_1 + \alpha_2$ both have the Lee-Yang property. Thus, by Lemma 2.4.6, $\alpha_{12}(z_3, \ldots, z_n) \neq 0$ when $|z_i| \geq 1$ for $i \geq 3$. Now, again by hypothesis, $A_2 \neq 0$ when $|z_1|$ and $|z_3|, \ldots, |z_n|$ are at least 1, while $z_1 = -\frac{\alpha_2(z_3,\ldots,z_n)}{\alpha_{12}(z_3,\ldots,z_n)}$ gives $A_2 = 0$. Thus, we must have that

$$\frac{|\alpha_2(z_3, \ldots, z_n)|}{|\alpha_{12}(z_3, \ldots, z_n)|} < 1 \text{ when } |z_i| \geq 1 \text{ for } i \geq 3. \tag{2.17}$$

We now set $\zeta_i = z_i$ for $i \geq 3$, and consider z_1 as a function of z_2. The equality $P(z_1, z_2, \zeta_3, \ldots, \zeta_n) = 0$ is then equivalent to

$$z_1 = -\frac{\alpha_2 z_2 + \alpha_\emptyset}{\alpha_{12} z_2 + \alpha_1}, \tag{2.18}$$

where the hypotheses of the lemma imply that the denominator (which is equal to $A_1(z_2, \zeta_3, \ldots, \zeta_n)$) is non-zero when $|z_2| \geq 1$. We thus see that

$$\lim_{z_2\to\infty} |z_1| = \frac{|\alpha_2|}{|\alpha_{12}|} < 1. \tag{2.19}$$

Initially, both z_1 and z_2 lie outside the closed unit disk. Thus, by eq. (2.19) and continuity, we can take z_2 large enough in absolute value such that z_1 as defined in eq. (2.18) lies on the unit circle. We now choose ζ_1 and ζ_2 to be these values of z_1 and z_2, respectively, so that we have $P(\zeta_1, \ldots, \zeta_n) = 0$ and the number of the ζ_i lying on the unit circle is exactly one less than the number of the z_i lying on the unit circle, as required. □

Along with the above general facts about LY polynomials, we also need the following technical lemma.

Lemma 2.4.9. *Let m be any integer, and k a positive integer such that $2\,|m| \leq k$. Consider the maximization problem*

$$\begin{aligned} & \max \prod_{i=1}^{k} \cos\theta_i \\ \textit{subject to} \quad & -\frac{\pi}{2} \leq \theta_i \leq \frac{\pi}{2}, \\ & \sum_{i=1}^{k} \theta_i = m\pi. \end{aligned}$$

The maximum is $\cos^k\left(\frac{m\pi}{k}\right)$, and is attained when $\theta_i = \frac{m\pi}{k}$ for all i.

Proof. We may assume without loss of generality that $\theta_i \in (-\pi/2, \pi/2)$ at any maximum (for otherwise the objective value is 0). Now, consider the function $f(x) = \log\cos x$ defined on the interval $(-\pi/2, \pi/2)$. Since $f'(x) = -\tan x$ is a decreasing function, $f(x)$ is concave for $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Thus by Jensen's inequality,

$$\log\prod_{i=1}^{k}\cos\theta_i = \sum_{i=1}^{k} f(\theta_i) \leq kf\left(\frac{\sum_{i=1}^{k}\theta_i}{k}\right) \leq k\log\cos\left(\frac{m\pi}{k}\right),$$

and equality holds when $\theta_i = \frac{m\pi}{k}$ for all i. Note that these θ_i are in $(-\pi/2, \pi/2)$ since $2\,|m| \leq k$. □

We are now ready to tackle the case of a single hyperedge.

Lemma 2.4.10. *Fix an integer $k \geq 2$ and a hyperedge activity $\beta \in \mathbb{R}$. Let $G = (V = \{v_1, v_2, \ldots, v_k\}, E = \{\{v_1, v_2, \ldots, v_k\}\})$ be a hypergraph consisting of a single hyperedge of size k and activity β. If $k = 2$ and $\beta \in (-1, 1)$, or $k \geq 3$ and β satisfies*

$$-\frac{1}{2^{k-1}-1} < \beta < \frac{1}{2^{k-1}\cos^{k-1}\left(\frac{\pi}{k-1}\right)+1},$$

then the partition function Z_G^β has the Lee-Yang property.

Remark 6. Note that the condition on β imposed above is monotone in k: i.e., if β is such that the partition function of a hyperedge of size $k \geq 2$ is LY, then for the same β the partition function of a hyperedge of size $k' < k$ is also LY.

Proof. For $k = 2$, the lemma is a special case of the Lee-Yang theorem [LY52] (although it also follows by specializing the argument below). We therefore assume $k \geq 3$.

Since the Ising partition function is symmetric and all terms in the polynomial appear with positive coefficients, Lemma 2.4.8 applies and it suffices to verify that the polynomials A_j do not vanish when $|z_i| \geq 1$ for $i \neq j$. Without loss of generality we fix $j = 1$. We then have

$$A_1 = \beta \prod_{i=2}^{k}(1 + z_i) + (1 - \beta) \prod_{i=2}^{k} z_i.$$

Thus $A_1 = 0$ is equivalent to

$$\frac{1}{\beta} = 1 - \prod_{i=2}^{k}\left(1 + \frac{1}{z_i}\right). \tag{2.20}$$

To establish the lemma, we therefore only need to show that for the claimed values of β, eq. (2.20) has no solutions when $|z_i| \geq 1$ for all $i \geq 2$. We now proceed to establish this by analyzing the product on the right hand side of eq. (2.20).

The map $z \mapsto 1 + 1/z$ is a bijection from the complement of the open unit disk to the closed disk D of radius 1 centered at 1. Any $y \in D$ can be written as $y = r\exp(\iota\theta)$ for $\theta \in [-\pi/2, \pi/2]$ and $0 \leq r \leq 2\cos\theta$. Consider now the set $\mathbb{R} \cap \left\{\prod_{i=2}^{k} y_i \mid y_i \in D \text{ for } 2 \leq i \leq k\right\}$. We show that, for $k \geq 3$, this set is exactly the interval $[-\tau_0, \tau_1]$ where $\tau_0 = 2^{k-1}\cos^{k-1}(\pi/(k-1))$ and $\tau_1 = 2^{k-1}$. The claim of the lemma then follows since for the given values of β, $1 - 1/\beta$ lies outside $[-\tau_0, \tau_1]$ and hence eq. (2.20) cannot hold.

Recalling that each $y \in D$ can be written in the form $r\exp(\iota\theta)$ where $\theta \in [-\pi/2, \pi/2]$ and $0 \leq r \leq 2\cos\theta$, we find that the values τ_0 and τ_1 are defined by the following optimization problems (both of which are feasible since $k \geq 3$):

$$\begin{aligned}\tau_0 = 2^{k-1} \max \prod_{i=2}^{k} \cos\theta_i \\ \text{subject to} \quad -\frac{\pi}{2} \le \theta_i \le \frac{\pi}{2}, \\ \sum_{i=2}^{k} \theta_i = (2n+1)\pi \\ \text{for some } n \in \mathbb{Z} \\ \text{s.t. } |2n+1| \le (k-1)/2.\end{aligned} \qquad \begin{aligned}\tau_1 = 2^{k-1} \max \prod_{i=2}^{k} \cos\theta_i \\ \text{subject to} \quad -\frac{\pi}{2} \le \theta_i \le \frac{\pi}{2}, \\ \sum_{i=2}^{k} \theta_i = 2n\pi \\ \text{for some } n \in \mathbb{Z} \\ \text{s.t. } |n| \le (k-1)/4.\end{aligned}$$

Using Lemma 2.4.9, we then see that $\tau_0 = 2^{k-1}\cos^{k-1}(\pi/(k-1))$ and $\tau_1 = 2^{k-1}$, as required. □

We now proceed to an inductive proof of Theorem 2.4.4, using Lemma 2.4.10 as the base case.

Proof of Theorem 2.4.4. The case $k = 2$ is a special case of the Lee-Yang theorem [LY52] (though, as with the proof of Lemma 2.4.10, the argument below can again be specialized to directly establish this). We assume therefore that $k \ge 3$.

The proof uses the inductive method of Asano [Asa70]. When the hypergraph consists of a single hyperedge of size $k' \le k$, it follows from Lemma 2.4.10 and the remark immediately after it that the partition function is LY for the claimed values of the edge activity β. For the induction, we use the fact that the Lee-Yang property of the partition function is preserved under the following two operations:

1. **Adding a hyperedge**: In this operation, a new hyperedge e of size $k' \le k$ and activity β_e as claimed in the statement of the theorem, is added to a connected hypergraph in such a way that exactly one of its k' vertices already exists in the starting hypergraph, while the other $k' - 1$ vertices are new. Note that this operation keeps the hypergraph connected. We assume that the partition functions of both the original hypergraph as well as the newly added edge separately have the Lee-Yang property: this follows from the induction hypothesis (for the hypergraph) and Lemma 2.4.10 (for the new hyperedge).

2. **Asano contraction**: In this operation, two vertices u', u'' in a connected hypergraph that are not both included in any one hyperedge are merged so that the new merged

vertex u is incident on all the hyperedges incident on u' or u'' in the original graph. Note that this operation keeps the hypergraph connected and does not change the size of any of the hyperedges.

Any connected non-empty hypergraph G can be constructed by starting with any arbitrary hyperedge present in G and performing a finite sequence of the above two operations: to add a new hyperedge e with activity β_e, one first uses operation 1 to add a hyperedge which has the same activity β_e and has new copies of all but one of the incident vertices of e, and then uses operation 2 to merge these new copies with their counterparts, if any, in the previous hypergraph. Note that in this process, a hyperedge e can be added only when at least one of its vertices is already included in the current hypergraph. However, since G is assumed to be connected, its hyperedges can be ordered so that all of them are added by the above process. Thus, assuming that the above two operations preserve the Lee-Yang property, it follows by induction that the partition functions of all connected hypergraphs of hyperedge size at most k, and edge activities β_e as claimed in the theorem, have the Lee-Yang property.

Given Corollary 2.4.7, it can be proved, by adapting an argument first developed by Asano [Asa70], that these two operations preserve the Lee-Yang property. Asano's method has by now become standard (see, e.g., [SF71, Propositions 1, 2]), but we include the details here for completeness.

Consider first operation 1. Let G be the original hypergraph and H the new hyperedge (with $k' \leq k$ vertices) being added, and assume, by renumbering vertices if required, that the single shared vertex is v_1 in G and u_1 in H, respectively. Let $P(z_1, z_2, \ldots, z_n) = A(z_2, \ldots, z_n)z_1 + B(z_2, \ldots, z_n)$ and $Q(y_1, y_2, \ldots, y_{k'}) = C(y_2, \ldots, y_{k'})y_1 + D(y_2, \ldots, y_{k'})$ be the Ising partition functions of G and H, respectively, where z_1 and y_1 are the variables corresponding to v_1 and u_1, respectively. Both P and Q are LY by the hypothesis of the operation. The partition function R of the new graph can be written as

$$R(z, z_2, \ldots, z_n, y_2, \ldots, y_{k'}) = A(z_2, \ldots, z_n)C(y_2, \ldots, y_{k'})z + B(z_2, \ldots, z_n)D(y_2, \ldots, y_{k'}),$$

where z is a new variable corresponding to the new vertex created by the merger of u_1 and v_1. Let $\lambda_2, \ldots, \lambda_n$, and $\mu_2, \ldots, \mu_{k'}$ be complex numbers lying outside the open unit disk. In order to prove that R is LY, we need to show that (i) $R(z, \lambda_2, \ldots, \lambda_n, \mu_2, \ldots, \mu_{k'}) = 0$ implies that $|z| \leq 1$; and (ii) when at least one of these complex numbers lies strictly outside the *closed* unit disk then $R(z, \lambda_2, \ldots, \lambda_n, \mu_2, \ldots, \mu_{k'}) = 0$ implies that $|z| < 1$. Now, since P and Q are assumed to be LY, Lemma 2.4.6 implies that $A = A(\lambda_2, \ldots, \lambda_n)$ and $C = C(\mu_2, \ldots, \mu_{k'})$ are both non-zero. Thus, $R = 0$ implies that

$$|z| = |B/A| \cdot |D/C|, \tag{2.21}$$

where $B = B(\lambda_2, \ldots \lambda_n)$ and $D = D(\mu_2, \ldots, \mu_{k'})$. Since all the λ_i and μ_i lie outside the open unit disk and P and Q are LY, $|B/A|, |D/C| \leq 1$, so that from eq. (2.21), $|z| \leq 1$. This establishes condition (i). Further, when at least one of the λ_i lies strictly outside the closed unit disk, then again, since P is LY, $|B/A| < 1$. Similarly,$|D/C| < 1$ when one of the μ_i lies outside the closed unit disk. Thus, when at least one of the λ_i and the μ_i lies outside the closed unit disk, it follows from eq. (2.21) that $|z| < 1$, thus establishing condition (ii) and concluding the argument that R is LY.

We now consider operation 2. By renumbering vertices if necessary, let v_1 and v_2 be the vertices to be merged. The partition function P of the original graph (where v_1 and v_2 are not merged) can be written as

$$P(z_1, z_2, z_3, \ldots, z_n) = A(z_3, \ldots, z_n)z_1z_2 + B(z_3, \ldots, z_n)z_1 + C(z_3, \ldots, z_n)z_2 + D,$$

and is LY by the hypothesis of the operation. The partition function R after the merger is then given by

$$R(z, z_3, \ldots, z_n) = A(z_3, \ldots, z_n)z + D,$$

where z is a new variable corresponding to the vertex created by the merger of v_1 and v_2. Now, let $\lambda_3, \ldots, \lambda_n$ be complex numbers lying outside the open unit disk. Corollary 2.4.7 implies that $A = A(\lambda_3, \ldots, \lambda_n) \neq 0$. Thus, $R(z, \lambda_3, \ldots, \lambda_n) = 0$ implies that

$$|z| = |D(\lambda_3, \ldots, \lambda_n)/A(\lambda_3, \ldots, \lambda_n)| = |D/A|. \tag{2.22}$$

Now, since P is LY, both zeros of the quadratic equation $P(x, x, \lambda_3, \ldots, \lambda_n) = 0$ satisfy $|x| \leq 1$, and indeed, $|x| < 1$ when at least one of the λ_i lies strictly outside the closed unit disk. Thus, the product D/A of its zeros also satisfies $|D/A| \leq 1$, and further satisfies the stronger inequality $|D/A| < 1$ in case at least one of the λ_i lies strictly outside the closed unit disk. Eq. (2.22) then implies that $|z| \leq 1$ in the first case and $|z| < 1$ in the second case, which establishes that R is LY.

This concludes the proof of the first part of Theorem 2.4.4. We now prove the optimality of the conditions imposed on the edge parameters. In the case $k = 2$, this follows by considering the partition function $z_1z_2 + \beta z_1 + \beta z_2 + 1$ of a single edge. When $\beta > 1$ (respectively, when $\beta < -1$), $z_1 = z_2 = -\beta - \sqrt{\beta^2 - 1}$ (respectively, $z_1 = z_2 = -\beta + \sqrt{\beta^2 - 1}$) is a zero of the partition function satisfying $|z_1|, |z_2| > 1$ and hence contradicting the Lee-Yang property. Similarly $z_1 = -1, z_2 = 2$ when $\beta = 1$, and $z_1 = 1, z_2 = 2$ when $\beta = -1$, are zeros which contradict the Lee-Yang property.

We now consider the case $k \geq 3$. In this case, we take our example to be the single hyperedge of size k and consider its partition function

$$P(z_1, z_2, \ldots, z_k) := \beta \prod_{i=1}^{k}(1 + z_i) + (1 - \beta)\left(1 + \prod_{i=1}^{k} z_i\right). \tag{2.23}$$

Our strategy is to show that when

$$\beta \notin \left(-\frac{1}{2^{-k-1} - 1}, \frac{1}{2^{k-1}\cos^{k-1}(\frac{\pi}{k-1}) + 1}\right), \tag{2.24}$$

the polynomial $A_1(z_2, z_2, \ldots, z_k)$, which is the coefficient of z_1 in P as defined in Lemma 2.4.6, vanishes at a point with $|z_i| \geq 1$ for $i \geq 2$. It then follows from Lemma 2.4.6 that P cannot have the Lee-Yang property.

To carry out the strategy, we reuse some of the notation and calculations from the proof of Lemma 2.4.10. Let D be the closed disk of radius 1 centered at 1, as defined in the proof of that lemma. Eq. (2.20), taken together with the discussion following it, implies that finding a zero of $A_1(z_2, \ldots, z_k)$ with $|z_i| \geq 1$, $2 \leq i \leq k$, is equivalent to finding $y_i \in D$, $y_i \neq 1$ such that $1 - \frac{1}{\beta} = \prod_{i=2}^{k} y_i$. We can in fact choose all the y_i to be equal, so that using the same representation of D as in the proof of Lemma 2.4.10, our task reduces to finding an angle $\theta \in [-\pi/2, \pi/2]$ and $0 \leq r \leq 2\cos\theta$ such that $y_i = re^{\iota\theta}$, $2 \leq i \leq k$, and

$$1 - \frac{1}{\beta} = \left(re^{\iota\theta}\right)^{k-1}. \tag{2.25}$$

Let $\gamma := 1 - \frac{1}{\beta}$. We now partition the condition on β in (2.24) into three different cases. Suppose first that $\beta \leq -\frac{1}{2^{k-1}-1}$. This is equivalent to $1 < \gamma \leq 2^{k-1}$. In this case $\theta = 0$, $r = \gamma^{\frac{1}{k-1}} \in (1, 2]$ gives a desired solution to (2.25) (note that we have $y_i \in (1, 2]$ in this case). The same solution for θ and r also works when $\beta > 1$ (in which case $0 < \gamma < 1$ and $y_i \in (0, 1)$) . The remaining case is $1 \geq \beta \geq \frac{1}{2^{k-1}\cos^{k-1}(\frac{\pi}{k-1})+1}$, which in turn is equivalent to $-2^{k-1}\cos^{k-1}(\frac{\pi}{k-1}) \leq \gamma \leq 0$, and $\theta = \frac{\pi}{k-1}$, $r = |\gamma|^{\frac{\pi}{k-1}} \leq 2\cos\theta$ gives a solution in this case. $\square$

2.4.1 Optimality of the univariate hypergraph Lee-Yang theorem

We now prove the second part of the univariate hypergraph Lee-Yang theorem, Theorem 2.1.2, i.e., that the range of edge activities under which the first part of that theorem holds is optimal. The tight example for the case $k = 2$ is a single edge, and as observed above, the roots of the univariate partition function of the edge when $|\beta| > 1$ are $-\beta \pm \sqrt{\beta^2 - 1}$, which do not lie on the unit circle.

We now consider the case $k \geq 3$. The tight example is again a hyperedge of size $k' \leq k$. The partition function $P_{k'}(z)$ of this graph is

$$P_{k'}(z) := \beta(1+z)^{k'} + (1-\beta)(1+z^{k'}),$$

and we will show that it has at least one root outside the unit circle when $\beta \neq 1$ satisfies

$$\beta \notin \left[-\frac{1}{2^{-k-1}-1}, \frac{1}{2^{k-1}\cos^{k-1}\left(\frac{\pi}{k-1}\right)+1} \right]. \tag{2.26}$$

We consider three exhaustive cases under which (2.26) holds.

Case 1: $\beta > 1$. In this case our tight example is a hyperedge of size $k' = 2 \leq k$, and the result follows from that of the case $k = 2$.

Case 2: $\beta < -\frac{1}{2^{k-1}-1}$. In this case, our example is a hyperedge of size k. We note then that $P_k(0) = 1$ and $P_k(1) = 2\beta(2^{k-1}-1)+2 < 0$. Thus, there exists a z in the interval $(0, 1)$ for which $P_k(z) = 0$, and hence P_k has a zero that is not on the unit circle.

Case 3: $\frac{1}{2^{k-1}\cos^{k-1}\left(\frac{\pi}{k-1}\right)+1} < \beta < 1$. Our tight example is again a hyperedge of size k. We will show that the degree k polynomial P_k has at most $k-3$ zeros (counting with multiplicities) on the unit circle C, and hence must have at least one zero outside it.

We first consider the point $z = -1$. Note that since $\beta \neq 1$, $P_k(-1) = 0$ if and only if k is odd, and in this case $P_k'(-1) = k(1-\beta) \neq 0$. Therefore, -1 is a zero of multiplicity 1 of P_k when k is odd, and is not a zero of P_k when k is even.

We now consider zeros of P_k in $C \setminus \{-1\}$. Let $\tau := 2^{k-1}\frac{\beta}{\beta-1}$ and $g(z) := \frac{1+z^k}{(1+z)^k}$. Note that any $z \in C \setminus \{-1\}$ is a zero of multiplicity l of P_k if and only if it is a zero of the same multiplicity l of the meromorphic function $g(z) - \tau/2^{k-1}$. In particular, at such a z, the order of the first non-zero derivative of P_k is the same as the order of the first non-zero derivative of g, and this number is the multiplicity of z as a zero of P (or equivalently, as a root of $g(z) = \tau/2^{k-1}$). Note also that $g(z)$ maps $C \setminus \{-1\}$ into the real line: in fact, for $z = e^{2\iota\theta}$, $\theta \in (-\pi/2, \pi/2)$, we have

$$2^{k-1}g(z) = 2^{k-1} \cdot \frac{1 + \cos 2k\theta + \iota \sin 2k\theta}{(1 + \cos 2\theta + \iota \sin 2\theta)^k} = \frac{2^k \cos k\theta}{(2\cos\theta)^k} \cdot \frac{e^{\iota k\theta}}{e^{\iota k\theta}} = \frac{\cos k\theta}{\cos^k \theta} =: h(\theta),$$

and further $h'(\theta) = 2^k \iota z g'(z)$, so that $h'(\theta) = 0$ if and only if $g'(z) = 0$. Indeed, by computing further derivatives, one sees that the multiplicity of any root of $h(\theta) = \tau$ in $(-\pi/2, \pi/2)$ (i.e., the order of the first non-zero derivative of h at the root) is the same as the multiplicity of the corresponding root $z = e^{2\iota\theta}$ of $g(z) = \tau/2^{k-1}$.

Thus, in order to establish our claim that $P_k(z)$ has at most $k-3$ zeros (counting with multiplicities and also accounting for the possible zero at -1 considered above) on the unit circle C, we only need to show that the number of roots of the equation $h(\theta) = \tau$ on $(-\pi/2, \pi/2)$ (counted with multiplicities) is at most $k - 4$. We now proceed to establish this property of h. Note that for the range of β being considered, we have $\tau < -\sec^{k-1}\left(\frac{\pi}{k-1}\right)$.

Since $h(\theta) = h(-\theta)$, we consider its behavior only in the interval $I = [0, -\pi/2)$. We have $h'(\theta) = -\frac{k\sin(k-1)\theta}{\cos^{k+1}\theta}$, so that the zeros of h' in I are given by $\rho_i = i\pi/(k-1)$, where $0 \le i < \lfloor k/2 \rfloor$ is an integer. Note that all these zeros of h' are in fact simple: $h''(\rho_i) \neq 0$. Thus, any root of $h(\theta) = \tau$ is of multiplicity at most 2. Now, define $\rho_{\lfloor k/2 \rfloor} = \pi/2$, and let I_i be the interval $[\rho_i, \rho_{i+1})$ for $0 \le i \le \lfloor k/2 \rfloor - 1$. We note the following facts (see Figure 2.1 for an example):

1. In the interval I_i, h is strictly decreasing when i is even and strictly increasing when i is odd.
2. For $i < \lfloor k/2 \rfloor$, $h(\rho_i) = (-1)^i \sec^{k-1}\left(\frac{i\pi}{k-1}\right)$, so that $h(\rho_i)$ is strictly positive when i is even and strictly negative when i is odd. Further, $h(\rho_1) = -\sec^{k-1}\left(\frac{\pi}{k-1}\right) > \tau$.

From these observations we can now deduce that when $-\sec^{k-1}\left(\frac{\pi}{k-1}\right) > \tau$, $h(\theta) = \tau$ has

1. no roots in I_0 and I_1,
2. at most two roots in $I_i \cup I_{i+1}$, counting multiplicities, when i is a positive even integer strictly less than $\lfloor k/2 \rfloor - 1$. The two roots can arise in only the following two ways: there can be one root each, with multiplicity 1, in each of the two intervals I_i and I_{i+1}, or else there can be a root of multiplicity 2 at ρ_{i+1}.

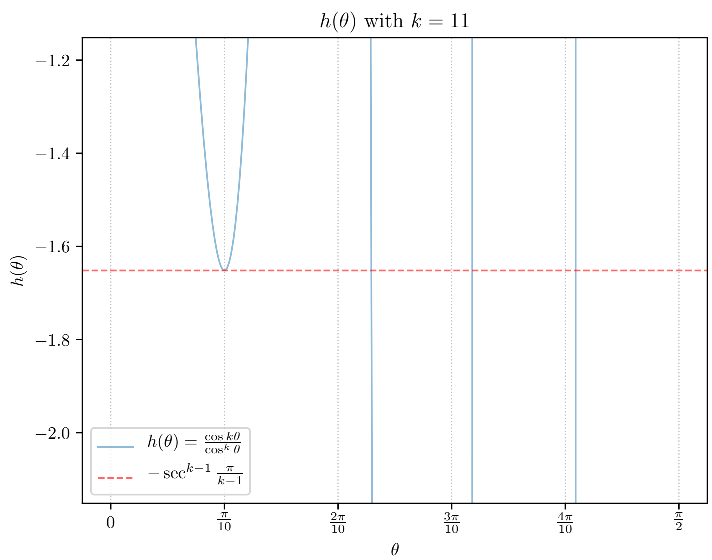

Figure 2.1. The function $h(\theta)$

3. at most one additional root in $I_{\lfloor k/2 \rfloor - 1}$, and this additional root can arise only when $\lfloor k/2 \rfloor - 1$ is even.

Together, the above three items imply that when $\tau < -\sec^{k-1}\left(\frac{\pi}{k-1}\right)$, the number of roots of $h(\theta) = \tau$ in $I = [0, -\pi/2)$, counted with their multiplicities, is at most $\lfloor k/2 \rfloor - 2$. Using the symmetry of h around 0 pointed out above, we thus see that the number of roots of $h(\theta) - \tau$ in $(-\pi/2, \pi/2)$ is at most $k - 4$, so that P_k has at most $k - 3$ zeros (accounting for the possible simple zero at -1 when k is odd) on the unit circle for such β. This implies that at least one zero of the degree k polynomial P_k must lie outside the unit circle, as required.

2.5 Related work

The problem of computing partition functions has been widely studied, not only in statistical physics but also in combinatorics, because the partition function is often a generating function for combinatorial objects (cuts, in the case of the Ising model). There has been much progress on *dichotomy theorems*, which attempt to completely classify these problems as being either #P-hard or computable (exactly) in polynomial time (see, e.g., [CCL10, GGJT10]).

Since the problems are in fact #P-hard in most cases, algorithmic interest has focused largely on *approximation*, motivated also by the general observation that approximating the partition function is polynomial time equivalent to sampling (approximately) from the underlying Gibbs distribution [JVV86]. In fact, most early approximation algorithms exploited this connection, and gave *fully-polynomial randomized approximation schemes (FPRAS)* for the partition function using Markov chain Monte Carlo (MCMC) samplers for the Gibbs distribution. In particular, for the ferromagnetic Ising model, the MCMC-based algorithm of Jerrum and Sinclair [JS93] is valid for all positive real values of λ and for all graphs, irrespective of their vertex degrees. (For the connection with random sampling in this case, see [RW99].) This was later extended to ferromagnetic two-spin systems by Goldberg, Jerrum and Paterson [GJP03]. Similar techniques have been applied recently to the related random-cluster model by Guo and Jerrum [GJ17].

Much detailed work has been done on MCMC for Ising spin configurations for several important classes of graphs, including two-dimensional lattices (e.g., [MO94a, MO94b, LS12]), random graphs and graphs of bounded degree (e.g., [MS13]), the complete graph (e.g., [LNNP14]) and trees (e.g., [BKMP05, MSW04]); we do not attempt to give a comprehensive summary of this line of work here.

In the *anti-ferromagnetic* regime ($\beta > 1$), *deterministic* approximation algorithms based on correlation decay have been remarkably successful for graphs of bounded degree. Specifically, for any fixed integer $\Delta \geq 3$, techniques of Weitz [Wei06] give a deterministic FPTAS for the anti-ferromagnetic Ising partition function on graphs of maximum degree Δ throughout a region R_Δ in the (β, λ) plane (corresponding to the regime of uniqueness of the Gibbs measure on the Δ-regular tree) [SST14, LLY13]. A complementary result of Sly and Sun [SS14b] (see also [GGŠ+14]) shows that the problem is NP-hard outside R_Δ. In contrast, no MCMC based algorithms are known to provide an FPRAS for the anti-ferromagnetic Ising partition function throughout R_Δ. More recently, correlation decay techniques have been extended to obtain deterministic approximation algorithms for the anti-ferromagnetic Ising partition function on hypergraphs over a range of parameters [LYZ16], as well as to several other problems not related to the Ising model. In the ferromagnetic setting, however, for reasons mentioned earlier, correlation decay techniques have had more limited success: Zhang, Liang and Bai [ZLB11] handle only the "high-temperature" regime of the Ising model, while the recent results for ferromagnetic two-spin systems of Guo and Lu [GL16] do not apply to the case of the Ising model.

In a parallel line of work, Barvinok initiated the study of Taylor approximation of the logarithm of the partition function, which led to quasipolynomial time approximation algorithms for a variety of counting problems [Bar15b, Bar15a, BS16a, BS16b]. More recently, Patel and Regts [PR17a] showed that for several models that can be written as induced subgraph sums, one can actually obtain an FPTAS from this approach. In particular, for problems such as counting independent sets with negative (or, more generally, complex valued) activities on bounded degree graphs, they were able to match the range of applicability of existing algorithms based on correlation decay, and were also able to extend the approach to Tutte polynomials and edge-coloring models (also known as Holant problems) where little is known about correlation decay.

The Lee-Yang program was initiated by Lee and Yang [YL52] in connection with the analysis of phase transitions. By proving the famous Lee-Yang circle theorem for the ferromagnetic Ising model [LY52], they were able to conclude that there can be at most one phase transition for the model. Asano [Asa70] extended the Lee-Yang theorem to the Heisenberg model, and provided a simpler proof. Asano's work was generalized further by Suzuki and Fisher [SF71], while Sinclair and Srivastava [SS14a] studied the multiplicity of Lee-Yang zeros. A complete characterization of Lee-Yang polynomials that are independent of the "temperature" of the model was recently obtained by Ruelle [Rue10]. The study of Lee-Yang type theorems for other statistical physics models has also generated beautiful connections with other areas of mathematics. For example, Shearer [She85] and Scott and Sokal [SS04] established the close connection between the location of the zeros of the independence polynomial and the Lovász Local Lemma, while the study of the zeros of generalizations of the matching polynomial was used in the recent celebrated work of Marcus, Spielman and Srivastava on the existence of Ramanujan graphs [MSS15a]. Such Lee-Yang type theorems are exemplars of the more general stability theory of polynomials [BB09a, BB09b], a field of study that has had numerous recent applications to theoretical computer science and combinatorics (see, e.g., [BBL09, MSS15a, MSS15b, AG15, AG17, SS14a, SV17]).

Chapter 3

Correlation decay implies absence of zeros

In this chapter, we study the connection between the notions of correlation decay phase transition and analyticity phase transition. In particular, we will see how one can study the latter more classical notion of phase transition using the former one. We will begin by formulating a general paradigm in the context of the Ising model (without external field), and will then apply this paradigm to two other notable examples: the antiferromagnetic Ising model with external field and the hard-core model, where the correlation decay phase transitions have been well studied. Our paradigm is robust: to establish the absence of zeros in all three cases, we are able to simply lift the existing correlation decay analysis in an identical fashion. In doing so, we establish a formal connection between the two notions of phase transition.

3.1 Ising model

In this section, we consider *Fisher zeros* of the Ising model, and show that there are no Fisher zeros in a complex neighborhood around the *correlation decay* interval. Recall that given a graph G, an edge activity β and a vertex activity λ, the Ising partition function is defined as $Z_G(\beta, \lambda) = \beta^{|E(S,\overline{S})|}\lambda^{|S|}$. Formally, we view this partition function as a polynomial in β for a fixed λ, and study the complex zeros in β. This is in contrast to the previous chapter, where we viewed the Ising partition function as a polynomial in λ for a fixed β, and focused on the complex zeros in terms of λ. As discussed in the last chapter, the study of complex zeros in λ was famously pioneered by Lee and Yang [LY52], and has given rise to a well developed theory; in contrast, very little is known about the zeros in β, which were first studied in the classical 1965 paper of Fisher [Fis65] and are thus known as "Fisher zeros".

More precisely, we fix $\lambda = 1$, and hence we will simply write $Z_G(\beta) := Z_G(\beta, 1)$ for the rest of the section. The correlation decay interval for the Ising model has been well studied: let Δ be the maximum degree, and let $d = \Delta - 1$; then the correlation decay interval for β is the interval $(\frac{\Delta-2}{\Delta}, \frac{\Delta}{\Delta-2})$. When β lies in this interval, the Gibbs distribution of the Ising model on a Δ-regular tree exhibits decay of long-range correlations. The main result of this section will be Corollary 3.1.7, which says that there are no Fisher zeros in a complex neighborhood of the correlation decay interval. This provides a formal link between the "decay of correlations" and "analyticity of free energy density" views of phase transitions. Last but not least, we remark that this zero-freeness result also implies the existence of efficient approximation algorithms for the partition function $Z_G(\beta)$ via Barvinok's paradigm discussed in section 2.2. The main difference here is that, instead of a disk, the zero-free region is now a strip containing the interval of interest; but one can apply the same analysis as in section 2.2 after mapping the zero-free region to a disk. A more detailed discussion of this is deferred to Corollary 3.4.1 and section 4.2.3.

We proceed to give an overview of our approach. Let G be any graph of maximum degree Δ. Our starting point is a recursive criterion that guarantees that the partition function $Z_G(\beta) \neq 0$. For any non-isolated vertex v of G, let $Z^+_{G,v}(\beta)$ (respectively, $Z^-_{G,v}(\beta)$) be the contribution to $Z_G(\beta)$ from configurations with $\sigma(v) = +$ (respectively, $\sigma(v) = -$), so that $Z_G(\beta) = Z^+_{G,v}(\beta) + Z^-_{G,v}(\beta)$. Define also the ratio $R_{G,v}(\beta) := \frac{Z^+_{G,v}(\beta)}{Z^-_{G,v}(\beta)}$. Now note that $Z^+_{G,v}(\beta)$ and $Z^-_{G,v}(\beta)$ can be seen as Ising partition functions defined on the same graph G with the vertex v *pinned* to the appropriate spin; i.e., they are partition functions defined on a graph with one less unpinned vertex. Without loss of generality, we assume that every pinned vertex has degree exactly one. [1] We will prove, inductively on the number of unpinned vertices, that neither $Z^+_{G,v}(\beta)$ nor $Z^-_{G,v}(\beta)$ vanishes. Under this inductive hypothesis, the condition $Z_G(\beta) \neq 0$ is equivalent to $R_{G,v}(\beta) \neq -1$. As we will see, for $\beta \in \mathbb{R}$, $R_{G,v}(\beta) > 0$. Thus it suffices to show that for complex β sufficiently close to the

[1] Suppose that a vertex v of degree k is pinned in a graph G, and consider the graph G' obtained by replacing v with k copies of itself, each pinned to the same spin and connected to exactly one of the original neighbors of v. Then $Z_G(\beta) = Z_{G'}(\beta)$ for all β.

correlation decay interval on the real line, $R_{G,v}(\beta) \approx R_{G,v}(\Re\beta)$.

Now we are ready to give the first technical ingredient: a formal recurrence, due to Weitz [Wei06], for computing ratios such as $R_{G,v}(\beta)$ in two-state spin systems. We start with some notation and definitions. For a vertex u in a graph G, if u has s^+ neighbors pinned to the spin $+$, and s^- neighbors pinned to the spin $-$, then we say that u has $(s^- - s^+)$ *signed pinned neighbors.*

Definition 3.1.1 (The graphs G_i). Given a graph G and an unpinned vertex u in G, let $v_1, \cdots, v_k$ be the unpinned neighbors of u. We define G_i (the vertex u will be understood from the context) to be the graph obtained from G as follows:

- first, replace vertex u with $u_1, \cdots, u_k$, and connect u_1 to v_1, u_2 to v_2, and so on;
- next, pin vertices $u_1, \cdots, u_{i-1}$ to spin $+$, and vertices $u_{i+1}, \cdots, u_k$ to spin $-$;
- finally, remove vertex u_i.

Note that the graph G_i has one fewer unpinned vertex than G. Moreover, the number of unpinned neighbors of v_i in G_i is at most $d = \Delta - 1$.

Lemma 3.1.2. *Let ω be a formal variable. Given a graph G and an unpinned vertex u, let k be the number of unpinned neighbors of u, and s be the number of signed pinned neighbors of u. Denoting $h_\omega(x) := \frac{\omega+x}{\omega x+1}$, we have*

$$R_{G,u}(\omega) = \omega^s \prod_{i=1}^{k} h_\omega(R_{G_i,v_i}(\omega)) .$$

Remark 7. Note that when a numerical value $\beta \in \mathbb{C}$ is substituted for ω in the above formal equalities, they remain valid numerical equalities as long as $\beta x_i + 1 \neq 0$ for any $\boldsymbol{x}$ appearing in the computation, and $Z^-_{G,v}(\beta) \neq 0$.

Moreover, as the number of unpinned neighbors of v_i in G_i is at most $d = \Delta - 1$, the tree recurrence will be applied with $k \leq d$ except at the root.

Proof. Let $v_1, v_2, \cdots, v_k$ be the unpinned neighbors of u, and $v_{k+1}, \cdots, v_{\deg_G(u)}$ be its pinned neighbors. For $0 \leq i \leq \deg_G(u)$, let H_i be the graph obtained from G as follows:

- replace vertex u with $u_1, \cdots, u_{\deg_G(u)}$, and connect u_1 to v_1, u_2 to v_2, and so on;

- pin vertices $u_1, \cdots, u_i$ to spin $+$, and vertices $u_{i+1}, \cdots, u_{\deg_G(u)}$ to spin $-$.

Note that H_i is the same as G_i, except that the last step of the construction of G_i is skipped, i.e, the vertex u_i is not removed, and, further, u_i is pinned to spin $+$. We can now write

$$R_{G,u}(\omega) = \frac{Z^+_{G,u}(\omega)}{Z^-_{G,u}(\omega)} = \frac{Z_{H_{\deg_G(u)}}(\omega)}{Z_{H_0}(\omega)} = \prod_{i=1}^{\deg_G(u)} \frac{Z_{H_i}(\omega)}{Z_{H_{i-1}}(\omega)} = \omega^s \cdot \prod_{i=1}^{k} \frac{Z_{H_i}(\omega)}{Z_{H_{i-1}}(\omega)}.$$

We observe that

$$\begin{aligned} Z_{H_i}(\omega) &= Z^+_{G_i,v_i} + \omega \cdot Z^-_{G_i,v_i}, \\ Z_{H_{i-1}}(\omega) &= \omega \cdot Z^+_{G_i,v_i} + Z^-_{G_i,v_i}. \end{aligned}$$

Therefore we have

$$R_{G,u}(\omega) = \omega^s \cdot \prod_{i=1}^{k} \frac{Z^+_{G_i,v_i} + \omega \cdot Z^-_{G_i,v_i}}{\omega \cdot Z^+_{G_i,v_i} + Z^-_{G_i,v_i}} = \omega^s \cdot \prod_{i=1}^{k} \frac{\frac{Z^+_{G_i,v_i}}{Z^-_{G_i,v_i}} + \omega}{\omega \cdot \frac{Z^+_{G_i,v_i}}{Z^-_{G_i,v_i}} + 1} = \omega^s \prod_{i=1}^{k} h_\omega(R_{G_i,v_i}(\omega)) .$$

This completes the proof. □

Given Lemma 3.1.2, we consider the following recurrence relation on the ratios:

$$F_{\beta,k,s}(\boldsymbol{x}) := \beta^s \prod_{i=1}^{k} h_\beta(x_i), \tag{3.1}$$

where as before $h_\beta(x) := \frac{\beta+x}{\beta x+1}$. This recurrence has been studied extensively in the literature on the Ising model on trees. It has also been found useful to re-parameterize the recurrence in terms of logarithms of likelihood ratios as follows (see, e.g., [Lyo89]). Let $\varphi(x) := \log x$ and define

$$F^\varphi_{\beta,k,s}(\boldsymbol{x}) := \left(\varphi \circ F_{\beta,k,s} \circ \varphi^{-1}\right)(\boldsymbol{x}) = s \log \beta + \sum_{i=1}^{k} \log h_\beta(e^{x_i}). \tag{3.2}$$

One then has the following "step-wise" version of correlation decay [Lyo89, ZLB11].

Proposition 3.1.3. *Fix a degree $\Delta = d+1 \geq 3$ and integers $k \geq 0$ and s. If $\frac{\Delta-2}{\Delta} < \beta < \frac{\Delta}{\Delta-2}$ then there exists an $\eta > 0$ (depending upon β and d) such that $\|\nabla F^\varphi_{\beta,k,s}(\boldsymbol{x})\|_1 \leq \frac{k}{d}(1-\eta)$ for every $\boldsymbol{x} \in \mathbb{R}^k$.*

Proof. By direct calculation, one has

$$\|\nabla F^{\varphi}_{\beta,k,s}(\boldsymbol{x})\|_1 = \sum_{i=1}^{k} \frac{\left|1-\beta^2\right|}{\beta^2+1+\beta(e^{x_i}+e^{-x_i})}.$$

By the AM-GM inequality, $e^x + e^{-x} \geq 2$ for every real x, and the right hand side is therefore at most $k \times \frac{|1-\beta|}{1+\beta}$. Now the condition on β implies that $\frac{|1-\beta|}{1+\beta} \leq \frac{1-\eta}{d}$ for some fixed $\eta > 0$. Therefore, we have $\|\nabla F^{\varphi}_{\beta,k,s}(\boldsymbol{x})\|_1 \leq k \times \frac{|1-\beta|}{1+\beta} \leq \frac{k}{d}(1-\eta)$. □

Next we give a bound on $R_{G,u}(\beta)$ for real-valued β. For any integers $k \geq 0$ and s, and a positive real β, we have $\beta^{k+|s|} \leq F_{\beta,k,s}(\boldsymbol{x}) \leq \frac{1}{\beta^{k+|s|}}$ when $\beta \leq 1$, and $\frac{1}{\beta^{k+|s|}} \leq F_{\beta,k,s}(\boldsymbol{x}) \leq \beta^{k+|s|}$ for $\beta \geq 1$, for all non-negative $\boldsymbol{x} \in \mathbb{R}^k_+$. Taking the logarithm of these bounds motivates the definition of the intervals $I_0(\beta, d)$ as follows:

$$I_0 = I_0(\beta, d) := \left[-d\left|\log\beta\right|, d\left|\log\beta\right|\right]. \tag{3.3}$$

Recalling Lemma 3.1.2, we see that the ratios $R_{G,u}(\beta)$ can be obtained by recursively applying the recurrence $F_{\beta,k,s}(\boldsymbol{x})$. Therefore, for $\beta \in \mathbb{R}$, any graph G and unpinned vertex u, we have $\log R_{G,u}(\beta) \in I_0(\beta, d)$. Next we state a corollary of Proposition 3.1.3 in the complex plane.

Corollary 3.1.4. *Fix a degree $\Delta = d+1 \geq 3$ and integers $k \geq 0$ and s. If $\frac{\Delta-2}{\Delta} < \beta < \frac{\Delta}{\Delta-2}$ then there exist positive constants $\eta, \varepsilon, \delta$ (depending upon β and d) such that the following is true. Let $D := D(\beta, d)$ be the set of points within distance ε of $I_0(\beta, d)$ in $\mathbb{C}$. Then $\|\nabla F^{\varphi}_{\beta,k,s}(\boldsymbol{x})\|_1 \leq \frac{k}{d}(1-\eta/2)$ for every $\boldsymbol{x} \in D^k$. Moreover, there is a finite constant M (depending upon β and d) such that $F^{\varphi}_{\beta,k,s}$ is M-Lipschitz in a complex neighborhood around β, i.e.,*

$$\sup_{\boldsymbol{x}\in D^k, \beta'\in\mathbb{C}:|\beta'-\beta|<\delta} \left|F^{\varphi}_{\beta,k,s}(\boldsymbol{x}) - F^{\varphi}_{\beta',k,s}(\boldsymbol{x})\right| < M\left|\beta-\beta'\right|.$$

Proof. Observe that $\|\nabla F^{\varphi}_{\beta,k,s}(\boldsymbol{x})\|_1 = \sum_{i=1}^{k} \frac{\left|1-\beta^2\right|}{\beta^2+1+\beta(e^{x_i}+e^{-x_i})}$ is a continuous function in x_i for every i. Since it is uniformly upper bounded by $\frac{k}{d}(1-\eta)$, for small enough ε, the expression can be bounded by $\frac{k}{d}(1-\eta/2)$ for all $x \in D^k$.

Finally, the existence of M follows from the analyticity of $F^{\varphi}_{\beta,k,s}$ around the respective point. □

In order to prove, inductively, that $R_{G,u}(\beta) \approx R_{G,u}(\Re\beta)$, we use a consequence of the mean value theorem for complex functions, tailored to our needs.

Lemma 3.1.5. *Let $F(\boldsymbol{x})$ be a holomorphic function on a complex poly-region D^k. For any $\boldsymbol{x}, \boldsymbol{x'} \in D^k$, we have*

$$\left|F(\boldsymbol{x}) - F(\boldsymbol{x'})\right| \leq \sup_{\boldsymbol{\xi} \in D^k} \|\nabla F(\boldsymbol{\xi})\|_1 \cdot \|\boldsymbol{x} - \boldsymbol{x'}\|_\infty.$$

Proof. Consider $g(t) := F(\boldsymbol{x} + t(\boldsymbol{x'} - \boldsymbol{x}))$. Observe that

$$g'(t) = \nabla F\big(\boldsymbol{x} + t(\boldsymbol{x'} - \boldsymbol{x})\big)^\intercal \left(\boldsymbol{x} - \boldsymbol{x'}\right).$$

Thus, for any $\boldsymbol{x}, \boldsymbol{x'} \in D^k$, we have

$$\begin{aligned}\left|F(\boldsymbol{x}) - F(\boldsymbol{x'})\right| = |g(1) - g(0)| = \left|\int_0^1 g'(t)dt\right| \\ \leq \sup_{t\in[0,1]} \left|g'(t)\right| \leq \sup_{\boldsymbol{\xi} \in D^k} \|\nabla F(\boldsymbol{\xi})\|_1 \cdot \|\boldsymbol{x} - \boldsymbol{x'}\|_\infty.\end{aligned}$$

□

Now we are ready to spell out the induction on the number of unpinned vertices.

Theorem 3.1.6. *Fix a degree $\Delta = d + 1 \geq 3$, and let $\beta \in (\frac{\Delta-2}{\Delta}, \frac{\Delta}{\Delta-2})$. There exist positive constants δ, ε (both depending on β and Δ) such that, for any graph G of maximum degree Δ, any unpinned vertex u in G with k unpinned neighbors, and any β' with $|\beta' - \beta| < \delta$, the following are true:*

1. $\left|Z^+_{G,u}(\beta')\right| > 0, \left|Z^-_{G,u}(\beta')\right| > 0.$
2. $|\varphi(R_{G,u}(\beta)) - \varphi(R_{G,u}(\beta'))| < \varepsilon \cdot \max\left\{\frac{k}{d}, 1\right\}.$

We will also refer to the two items above as the "induction hypothesis". We remark that $\beta \in (\frac{\Delta-2}{\Delta}, \frac{\Delta}{\Delta-2})$ is only needed so that we may appeal to correlation decay (in the form of Corollary 3.1.4).

Proof. We use induction on the number of unpinned vertices in G. For the base case, if u is the only unpinned vertex in G, with s^+ neighbors pinned to spin $+$ and s^- neighbors pinned to spin $-$, then $Z^+_{G,u}(\beta') = (\beta')^{s^-}, Z^-_{G,u}(\beta') = (\beta')^{s^+}$, and $\varphi(R_{G,u}(\beta)) = s \log \beta, \varphi(R_{G,u}(\beta')) = s \log \beta'$. Thus it suffices to choose $\delta < \frac{\varepsilon}{6d}$ (where ε is to be determined later), and the base case is satisfied. From now on, we consider a graph G with at least two unpinned vertices.

For the first item, let G' be the graph where we pin vertex u to spin $+$. Note that by definition, $Z^{+}_{G,u}(\beta') = Z_{G'}(\beta')$. Let v be any unpinned vertex in G'. Since G' has one less unpinned vertex than G, by the induction hypothesis we have $\left|Z^{-}_{G',v}(\beta')\right| > 0$, and $\left|\varphi\big(R_{G',v}(\beta)\big) - \varphi\big(R_{G',v}(\beta')\big)\right| < \varepsilon$. Recalling from eq. (3.3) that the range of $R_{G',v}(\beta)$ is such that φ is analytic, we see that $\left|R_{G',v}(\beta) - R_{G',v}(\beta')\right| = O(\varepsilon)$. Next, we write

$$\begin{aligned}\left|Z_{G'}(\beta')\right| = \left|Z^{+}_{G',v}(\beta') + Z^{-}_{G',v}(\beta')\right| &= \left|Z^{-}_{G',v}(\beta')\right| \cdot \left|1 + R_{G',v}(\beta')\right| \\ &= \left|Z^{-}_{G',v}(\beta')\right| \cdot \left|1 + R_{G',v}(\beta) + \xi\right|,\end{aligned}$$

for some $\xi \in \mathbb{C}$ with $|\xi| = O(\varepsilon)$. Thus, for sufficiently small ε, we have $\Re\big(R_{G',v}(\beta')\big) = \Re\big(R_{G',v}(\beta) + \xi\big)$, which is certainly at least $-\frac{1}{2}$ since $R_{G',v}(\beta) > 0$ as β is real. This means that $\left|1 + R_{G',v}(\beta) + \xi\right| \geq 1/2$, which implies that

$$\left|Z^{+}_{G,u}(\beta')\right| = \left|Z_{G'}(\beta')\right| \geq 0.5 \cdot \left|Z^{-}_{G',v}(\beta')\right| > 0.$$

An identical argument also proves that $\left|Z^{-}_{G,u}(\beta')\right| > 0$, completing the verification of item 1 of the induction hypothesis.

For the second item, let $v_1, \cdots, v_k$ be the unpinned neighbors of u. We denote $A(\beta) := \varphi(R_{G,u}(\beta))$, $B_i(\beta) := \varphi(R_{G_i,v_i}(\beta))$, $\boldsymbol{B}(\beta) := \{B_1(\beta), B_2(\beta), \cdots, B_k(\beta)\}$, and $H_\beta(x_1, x_2, \cdots, x_k) := F^{\varphi}_{\beta,k,s}(x_1, x_2, \cdots, x_k)$. Then by Lemma 3.1.2 we have

$$\begin{aligned}A(\beta) &= H_\beta(\boldsymbol{B}(\beta)), \\ A(\beta') &= H_{\beta'}\big(\boldsymbol{B}(\beta')\big).\end{aligned}$$

Let $\eta, \varepsilon, \delta, M$ be the constants (depending on β and d) whose existence is guaranteed by Corollary 3.1.4. By the triangle inequality, we have

$$\begin{aligned}\left|A(\beta) - A(\beta')\right| &\leq \left|H_\beta(\boldsymbol{B}(\beta)) - H_\beta\big(\boldsymbol{B}(\beta')\big)\right| + \left|H_\beta\big(\boldsymbol{B}(\beta')\big) - H_{\beta'}\big(\boldsymbol{B}(\beta')\big)\right| \\ &\leq \sup \|\nabla F^{\varphi}_{\beta,k,s}\|_1 \cdot \max_i \left|B_i(\beta) - B_i(\beta')\right| + M\left|\beta - \beta'\right| \\ &\leq (1 - \eta/2)\frac{k}{d}\varepsilon + M\delta,\end{aligned}$$

where the second line follows from Lemma 3.1.5, and the third line from Corollary 3.1.4 and the induction hypothesis applied to G_i.

Finally, replacing δ by $\min\left\{\delta, \frac{\varepsilon\eta}{2M}\right\}$ ensures that this last expression is bounded by $\max\left\{\frac{k}{d}, 1\right\}\varepsilon$, thus concluding the proof of the second item in the induction hypothesis. □

The main result of this section now follows immediately, by noting that the proof of the first item in the induction hypothesis remains valid even if u has $\Delta = d+1$ unpinned neighbors, as it only required the analyticity of $F^{\varphi}_{\beta,k,s}$, and that throughout the rest of the induction, v_i has at most $d = \Delta - 1$ unpinned neighbors in G_i.

Corollary 3.1.7. *Fix a degree $\Delta = d+1 \geq 3$, and let $\beta \in (\frac{\Delta-2}{\Delta}, \frac{\Delta}{\Delta-2})$. There exist positive constants δ, ε (both depending on β and Δ) such that, for any graph G of maximum degree Δ, and any β' with $|\beta' - \beta| < \delta$, we have $Z_G(\beta') \neq 0$.*

3.2 Antiferromagnetic Ising model

In this section, we consider the antiferromagnetic Ising model. For any $\beta < 1$, there is a *critical activity* $\lambda_c(\beta, \Delta)$ such that the Gibbs measure on the Δ-regular tree is unique if and only if $|\log \lambda| > \log \lambda_c(\beta, \Delta)$ (see, e.g., [Geo11]). We will refer to this as the *correlation decay region* for the antiferromagnetic Ising model. Fix any β, λ in the correlation decay region. We will show that there exists $\delta > 0$ such that for any β' with $|\beta' - \beta| < \delta$, the partition function $Z_G(\beta', \lambda) \neq 0$.

As before, for a fixed vertex v, we write $Z_G(\beta, \lambda) = Z^+_{G,v}(\beta, \lambda) + Z^-_{G,v}(\beta, \lambda)$, and let $R_{G,v}(\beta, \lambda) := \frac{Z^+_G(\beta,\lambda)}{Z^-_G(\beta,\lambda)}$. Then, there is a formal recurrence relation analogous to Lemma 3.1.2 as follows.

Lemma 3.2.1. *Let $\omega_\beta, \omega_\lambda$ be formal variables. Given a graph G and an unpinned vertex u, let k be the number of unpinned neighbors of u, and s be the number of signed pinned neighbors of u. Denoting $h_\omega(x) := \frac{\omega+x}{\omega x+1}$, we have*

$$R_{G,u}(\omega_\beta, \omega_\lambda) = \lambda \omega^s \prod_{i=1}^{k} h_\omega(R_{G_i,v_i}(\omega_\beta, \omega_\lambda)),$$

where the graphs G_i are defined as in Definition 3.1.1.

Given integers k and s, let $F_{\beta,\lambda}(\boldsymbol{x}) := \lambda \beta^s \prod_{i=1}^{k} h_\beta(x_i)$. This recurrence has been studied before in the literature [LLY12, SST14]. It has been found useful to reparameterize $F_{\beta,\lambda}$ with a "potential function" φ as follows: $F^{\varphi}_{\beta,\lambda} := \varphi \circ F_{\beta,\lambda} \circ \varphi^{-1}$. In [SST14], $\varphi(x) := \log \frac{x+D}{1-x+D}$ was chosen, where $D > 0$ is a constant depending on β and d. (Other choice of φ can be found in, e.g., [LLY12, LLY13].) For this choice of φ, the following step-wise correlation decay in the 1-norm is established:

Proposition 3.2.2. *Fix a degree $\Delta = d+1 \geq 3$ and integers $k \leq d$ and s. If (β, λ) is in the correlation decay region, then there exists an $\eta > 0$ (depending upon β, λ and d) such that $\|\nabla F^{\varphi}_{\beta,\lambda}(\boldsymbol{x})\|_1 < 1 - \eta$ for every $\boldsymbol{x} \in \mathbb{R}^k$.*

We also note that an analog of eq. (3.3) gives the bound $\frac{\lambda}{\beta^d} \leq R_{G,u}(\beta, \lambda) \leq \lambda\beta^d$. Thus we define

$$I_0(\beta, \lambda, d) := \left[\varphi\left(\frac{\lambda}{\beta^d}\right), \varphi\left(\lambda\beta^d\right)\right]. \tag{3.4}$$

Since φ is analytic, it has finite and continuous derivative. Thus the following corollary is immediate.

Corollary 3.2.3. *Fix a degree $\Delta = d+1 \geq 3$ and integers $k \leq d$ and s. If (β, λ) is in the correlation decay region, then there exist positive constants $\eta, \varepsilon, \delta$ (depending upon β, λ and d) such that the following is true. Let $D := D(\beta, \lambda, d)$ be the set of points within distance ε of $I_0(\beta, \lambda, d)$ in $\mathbb{C}$. Then $\|\nabla F^{\varphi}_{\beta,\lambda}(\boldsymbol{x})\|_1 < 1 - \eta/2$ for every $\boldsymbol{x} \in D^k$. Moreover, there is a finite constant M (depending upon β, λ and d) such that $F^{\varphi}_{\beta,\lambda}$ is M-Lipschitz in a complex neighborhood around β, i.e.,*

$$\sup_{\boldsymbol{x} \in D^k, \beta' \in \mathbb{C}: |\beta' - \beta| < \delta} \left|F^{\varphi}_{\beta,\lambda}(\boldsymbol{x}) - F^{\varphi}_{\beta,\lambda}(\boldsymbol{x})\right| < M\left|\beta - \beta'\right|.$$

Finally, we see that given Lemma 3.2.1 and Corollary 3.2.3, an identical argument to that in the proof of Theorem 3.1.6 proves the following:

Theorem 3.2.4. *Fix a degree $\Delta = d+1 \geq 3$, and let (β, λ) be in the correlation decay region. There exist positive constants δ, ε (both depending on β, λ and Δ) such that, for any graph G of maximum degree Δ, any unpinned vertex u in G with at most d unpinned neighbors, and any β' with $|\beta' - \beta| < \delta$, the following are true:*

1. $\left|Z^{+}_{G,u}(\beta', \lambda)\right| > 0, \left|Z^{-}_{G,u}(\beta', \lambda)\right| > 0.$
2. $|\varphi(R_{G,u}(\beta, \lambda)) - \varphi(R_{G,u}(\beta', \lambda))| < \varepsilon.$

The main result of this section now follows in an identical fashion to Corollary 3.1.7.

Corollary 3.2.5. *Fix a degree $\Delta = d+1 \geq 3$, and let (β, λ) be in the correlation decay interval. There exist positive constants δ, ε (both depending on β, λ and Δ) such that, for any graph G of maximum degree Δ, and any β' with $|\beta' - \beta| < \delta$, we have $Z_G(\beta', \lambda) \neq 0$.*

3.3 Hard-core model

In this section, we consider the *independence polynomial*, which is the partition function of the hard-core model. Formally, given a graph $G = (V, E)$ and a *vertex activity* λ, we let $\mathcal{I}(G)$ be the set of independent sets in G. Then the independence polynomial is given by

$$Z_G(\lambda) = \sum_{I \in \mathcal{I}(G)} \lambda^{|I|}.$$

The hardcore model is a simple model of the "excluded volume" phenomenon: vertices in the independent set I correspond to particles, each of which prevents neighboring sites from being occupied The parameter λ specifies the density of particles in the system.

In two seminal papers, Weitz [Wei06] and Sly [Sly10] (see also [GGŠ+14]) showed that there is a critical activity $\lambda_c(\Delta)$ such that when $\lambda < \lambda_c(\Delta)$, the partition function can be approximated efficiently for graphs of maximum degree Δ, while for $\lambda > \lambda_c(\Delta)$ close to the threshold, it becomes NP-hard to approximate the partition function. Sly and Sun [SS14b] (see also [GŠV15]) later extended the NP-hardness to the entire range of $\lambda > \lambda_c(\Delta)$. We will refer to $\lambda < \lambda_c(\Delta)$ as the *correlation decay interval* for the hard-core model. In this section, we view $Z_G(\lambda)$ as a polynomial in λ, and study the complex zeros in λ. The main result of this section will again be that there are no zeros in a complex neighborhood of the correlation decay interval $(0, \lambda_c(\Delta))$.

In similar fashion to the Ising model, for a fixed vertex v, we write $Z_G(\lambda) = Z_{G \setminus v}(\lambda) + \lambda \cdot Z_{G \setminus N_G[v]}(\lambda)$, and let $R_{G,v}(\lambda) := \frac{Z_{G \setminus N_G[v]}(\lambda)}{Z_{G \setminus v}(\lambda)}$. It is worth noting that $Z_{G \setminus v}(\lambda)$ is the same as pinning v to be "unoccupied" (not in the independent set) in G, while $Z_{G \setminus N_G[v]}(\lambda)$ is the same as pinning v to be "occupied" (in the independent set) in G. By analogy with Lemmas 3.2.1 and 3.1.2, we have the following formal recurrence relation for $R_{G,u}$ [Wei06], which is easily verified.

Lemma 3.3.1. *Let ω be a formal variable. Given a graph G and an unpinned vertex u, let k be the number of unpinned neighbors of u. We then have*

$$R_{G,u}(\omega) = \lambda \prod_{i=1}^{k} \frac{1}{1 + R_{G_i, v_i}(\omega)},$$

where the graphs G_i are defined analogous to Definition 3.1.1: specifically, $G_i := G \setminus \{u, v_1, \cdots, v_{i-1}\}$.

Given integers k and s, let $F_\lambda(\boldsymbol{x}) := \lambda\beta^s \prod_{i=1}^{k} \frac{1}{1+x_i}$. This recurrence has been studied before in the literature. As with the Ising model examples above, it has been found useful to reparameterize F_λ with a "potential function" φ as follows: $F_\lambda^\varphi := \varphi \circ F_\lambda \circ \varphi^{-1}$. In [LLY13], $\varphi(x) = \sinh^{-1}(\sqrt{x})$ was chosen, leading to the following step-wise correlation decay in the 1-norm:

Proposition 3.3.2. *Fix a degree $\Delta = d+1 \geq 3$ and integers $k \leq d$. If λ is in the correlation decay interval, then there exists an $\eta > 0$ (depending upon λ and d) such that $\|\nabla F_\lambda^\varphi(\boldsymbol{x})\|_1 < 1 - \eta$ for every $\boldsymbol{x} \in \mathbb{R}^k$.*

We also note that an analog of eq. (3.3) gives the bound $0 \leq R_{G,u}(\beta, \lambda) \leq \lambda$. Thus we define

$$I_0(\lambda, d) := [\varphi(0)\,, \varphi(\lambda)\,]\,. \tag{3.5}$$

Since φ is analytic, it has finite and continuous derivative. Thus the following corollary is immediate.

Corollary 3.3.3. *Fix a degree $\Delta = d+1 \geq 3$ and integers $k \leq d$ and s. If λ is in the correlation decay interval, then there exist positive constants $\eta, \varepsilon, \delta$ (depending on λ and d) such that the following is true. Let $D := D(\lambda, d)$ be the set of points within distance ε of $I_0(\lambda, d)$ in $\mathbb{C}$. Then, $\|\nabla F_\lambda^\varphi(\boldsymbol{x})\|_1 < 1 - \eta/2$ for every $\boldsymbol{x} \in D^k$. Moreover, there is a finite constant M (depending on λ and d) such that F_λ^φ is M-Lipschitz in a complex neighborhood around λ:*

$$\sup_{\boldsymbol{x} \in D^k, \lambda' \in \mathbb{C}: |\lambda' - \lambda| < \delta} \left| F_\lambda^\varphi(\boldsymbol{x}) - F_\lambda^\varphi(\boldsymbol{x}) \right| < M \left| \lambda - \lambda' \right|$$

Finally, we see that given Lemma 3.3.1 and Corollary 3.3.3, an identical argument to that in the proof of Theorem 3.1.6 proves the following:

Theorem 3.3.4. *Fix a degree $\Delta = d+1 \geq 3$, and let λ be in the correlation decay interval. There exist positive constants δ, ε (both depending on λ and Δ) such that, for any graph G of maximum degree Δ, any unpinned vertex u in G, and any λ' with $|\lambda' - \lambda| < \delta$, the following are true:*

1. *$|Z_G(\lambda')| > 0$.*

2. *$|\varphi(R_{G,u}(\lambda')) - \varphi(R_{G,u}(\lambda))| < \varepsilon$.*

The main result of this section now follows in an identical fashion to Corollary 3.1.7.

Corollary 3.3.5. *Fix a degree $\Delta = d+1 \geq 3$, and let λ be in the correlation decay interval. There exist positive constants δ, ε (both depending on λ and Δ) such that, for any graph G of maximum degree Δ, and any λ' with $|\lambda' - \lambda| < \delta$, we have $Z_G(\lambda') \neq 0$.*

3.4 Related work and discussion

The main highlight of this chapter is to go beyond the well studied Lee-Yang zeros for the Ising model, and obtain new results on Fisher zeros. While there are some results in the literature on Fisher zeros in the case of specific regular lattices (see, e.g., [LW01] and [KHK08]), to the best of our knowledge, the previous best general result on Fisher zeros appears in the work of Barvinok [Bar17] (see also Barvinok and Soberón [BS16a]), who showed that $Z_G(\beta)$ is non-zero if $|\beta - 1| < c/\Delta$, where Δ is the maximum degree of G and c can be chosen to be 0.34 (and as large as 1.12 if Δ is large enough). While this result provides a disk around 1 in which there are no Fisher zeros, it cannot guarantee the absence of Fisher zeros in a neighborhood of the correlation decay region B (which would require at least that $c \geq 2 - o_\Delta(1)$). Our Corollary 3.1.7 therefore strengthens this result to a neighborhood of the entire correlation decay region B.[2]

Our main theorem on Fisher zeros can also be combined with the techniques of Barvinok [Bar17] and Patel and Regts [PR17a] to give a new deterministic polynomial time approximation algorithm for the partition function of the ferromagnetic Ising model with zero field on graphs of degree at most Δ when $\beta \in (\frac{\Delta-2}{\Delta}, \frac{\Delta}{\Delta-2})$. In particular, combining Corollary 3.1.7 with [Bar17, Lemmas 2.2.1 and 2.2.3] (see also the discussion at the bottom of page 27 therein) and the proof of Theorem 6.1 of [PR17a], we obtain the following corollary:

Corollary 3.4.1. *Fix an integer $\Delta \geq 3$ and $\delta_1 > 0$. There exist positive constants $\delta > 0$ and c such that for any complex β with $\Re(\beta) \in \left[\frac{\Delta-2}{\Delta} + \delta_1, \frac{\Delta}{\Delta-2} - \delta_1\right]$ and $|\Im(\beta)| \leq \delta$, the following is true. There exists an algorithm which, on input a graph G of degree at most*

[2] Technically the results are incomparable in the sense that, while our results cover a much larger portion of the real line than that in [BS16a], the diameter of the disk centered around 1 in the region of [BS16a] may be larger than the radius guaranteed by our result.

Δ on n vertices, and an accuracy parameter $\epsilon > 0$, runs in time $O(n/\epsilon)^c$ and outputs $\hat{Z}$ satisfying $\left|\hat{Z} - Z_G(\beta)\right| \leq \epsilon\left|Z_G(\beta)\right|$.

For real β in the same range, a deterministic algorithm with the above properties, based on correlation decay, was already analyzed in [ZLB11]. However, our extension to complex values of the parameter is of independent algorithmic interest in light of the fact that algorithms for approximating the Ising partition function at complex values of the parameters have applications to the classical simulation of restricted models of quantum computation [MB18]. Analogous algorithmic results to that in Corollary 3.4.1 for the antiferromagnetic Ising model and the hard-core model follow in similar fashion from Corollary 3.2.5 and Corollary 3.3.5, respectively.

In contrast to most other recent applications of Barvinok's method (e.g., [PR17a, BS16b, BS16a, Bar15b, LSS19c]), where the required results on the location of the roots of the associated partition function are derived without reference to correlation decay, the algorithmic version of correlation decay is crucial to our proof. Indeed, implicit in our proof is an analysis of Weitz's celebrated correlation decay algorithm [Wei06], which was proposed originally for the independent set, or hard-core model. For the "zero field" Ising model, Weitz's algorithm was first analyzed by Zhang, Liang and Bai [ZLB11]; for the antiferromagnetic Ising model, it was first analyzed in [SST14, LLY12]; and for the hard-core model, the first "step-wise correlation decay" analysis can be found in [LLY13].[3] We are able to lift all these existing analyses and show that, in each case, there is a zero-free region of constant width that contains the entire correlation decay interval. Thus, as mentioned earlier, our work shows that Weitz's algorithm can be viewed as a bridge between the "decay of correlations" and "analyticity of free energy density" views of phase transitions.

We note that our work is close in spirit to recent work of Peters and Regts [PR17b] (see also [BC18]), who employ correlation decay in the hard-core model to prove stability results for the hard-core partition function. However, we note that the arguments of Peters and Regts crucially require an ad-hoc choice of "potential function" in the correlation decay analysis, so that certain geometric properties are satisfied. As a result, they were not able to exploit the existing correlation decay analysis for the hard-core model. In contrast, in our approach described in section 3.3, we can work with *any* correlation decay analysis based on the tree recurrence. We mention also that, in more recent work posted after a preprint

[3] [LLY13] worked with the more general two-state antiferromagnetic spin system, which includes the hard-core model as a special case.

of the present results appeared, Peters and Regts [PR18] also applied a combination of techniques from complex dynamics and correlation decay to study the location (on the unit circle) of Lee-Yang zeros of bounded degree graphs.

Chapter 4

Graph colorings and Fisher zeros of the Potts model

In the last chapter we saw that one can prove absence of zeros provided that a certain tree recurrence has the correlation decay property. However, there are other forms of correlation decay results, one notable example of which is the "strong spatial mixing" result for list-colorings by Gamarnik, Katz and Misra [GKM15]. Perhaps more interestingly, the argument in [GKM15] is not algorithmic. In the following, we will show that with a more sophisticated argument, the strong spatial mixing arguments of [GKM15] can also be exploited to prove absence of zeros. As a result, we are able to obtain efficient approximation algorithms in the same regime, which go well beyond the range of applicability of all previous deterministic algorithms for coloring.

4.1 Statements of results and technical overview

Counting colorings of a bounded degree graph is a benchmark problem in approximate counting, due both to its importance in combinatorics and statistical physics, as well as to the fact that it has repeatedly challenged existing algorithmic techniques and stimulated the development of new ones.

Given a finite graph $G = (V, E)$ of maximum degree Δ, and a positive integer q, the goal is to count the number of (proper) vertex colorings of G with q colors. It is well known [Bro41] that a greedy coloring exists if $q \geq \Delta + 1$. While counting colorings exactly is #P-complete, a long-standing conjecture asserts that approximately counting colorings is possible in polynomial time provided $q \geq \Delta + 1$. It is known that when $q \leq \Delta$, even approximate counting is NP-hard [GŠV15].

This question has led to numerous algorithmic developments over the past 25 years. The first approach was via Markov chain Monte Carlo (MCMC), based on the fact that approximate counting can be reduced to sampling a coloring (almost) uniformly at random. Sampling can be achieved by simulating a natural local Markov chain (or Glauber dynamics) that randomly flips colors on vertices: provided the chain is rapidly mixing, this leads to an efficient algorithm (a *fully polynomial randomized approximation scheme*, or *FPRAS*).

Jerrum's 1995 result [Jer95] that the Glauber dynamics is rapidly mixing for $q \geq 2\Delta + 1$ gave the first non-trivial randomized approximation algorithm for colorings and led to a plethora of follow-up work on MCMC (see, e.g., [DF03, DFHV13, FV06, GMP04, Hay03, HV03, HV05, Mol04, Vig00] and [FV08] for a survey), focusing on reducing the constant 2 in front of Δ. The best constant known for general graphs remains essentially $\frac{11}{6}$, obtained by Vigoda [Vig00] using a more sophisticated Markov chain, though this was very recently reduced to $\frac{11}{6} - \varepsilon$ for a very small ε by Chen et al. [CDM+19]. The constant can be substantially improved if additional restrictions are placed on the graph: e.g., Dyer et al. [DFHV13] achieve roughly $q \geq 1.49\Delta$ provided the girth is at least 6 and the degree is a large enough constant, while Hayes and Vigoda improve this to $q \geq (1+\varepsilon)\Delta$ for girth at least 11 and degree $\Delta = \Omega(\log n)$, where n is the number of vertices.

A significant recent development in approximate counting is the emergence of *deterministic* approximation algorithms that in some cases match, or even improve upon, the best known MCMC algorithms.[1] These algorithms have made use of one of two main techniques: *decay of correlations*, which exploits decreasing influence of the spins (colors) on distant vertices on the spin at a given vertex; and *polynomial interpolation*, which uses the absence of zeros of the partition function in a suitable region of the complex plane. Early examples of the decay of correlations approach include [Wei06, BG08, BGK+07], while for early examples of the polynomial interpolation method, we refer to the monograph of Barvinok [Bar17] (see also, e.g., [BR17, HPR19, PR17a, JKP19, GLLZ19, LSS19c, EM18]

[1] In this case, the notion of an FPRAS is replaced by that of a *fully polynomial time approximation scheme*, or *FPTAS*. An FPTAS for q-colorings of graphs of maximum degree at most Δ is an algorithm that given the graph G and an error parameter δ on the input, produces a $(1 \pm \delta)$-factor multiplicative approximation to the number of q-colorings of G in time $\mathrm{poly}(|G|, 1/\delta)$ (the degree of the polynomial is allowed to depend upon the constants q and Δ).

for more recent examples). Unfortunately, however, in the case of colorings on general bounded degree graphs, these techniques have so far lagged well behind the MCMC algorithms mentioned above. One obstacle to getting correlation decay to work is the lack of a higher-dimensional analog of Weitz's beautiful algorithmic framework [Wei06], which allows correlation decay to be fully exploited via strong spatial mixing in the case of spin systems with just two spins (as opposed to the q colors present in coloring). For polynomial interpolation, the obstacle has been a lack of precise information about the location of the zeros of associated partition functions (see below for a definition of the partition function in the context of colorings).

So far, the best algorithmic condition for colorings obtained via correlation decay is $q \geq 2.58\Delta + 1$, due to Lu and Yin [LY13], and this remains the best available condition for any deterministic algorithm. This improved on an earlier bound of roughly $q \geq 2.78\Delta$ (proved only for triangle-free graphs), due to Gamarnik and Katz [GK12]. For the special case $\Delta = 3$, Lu et al. [LYZZ17] give a correlation decay algorithm for counting 4-colorings. Furthermore, Gamarnik, Katz and Misra [GKM15] establish the related property of "strong spatial mixing" under the weaker condition $q \geq \alpha\Delta + \beta$ for any constant $\alpha > \alpha^\star$, where $\alpha^\star \approx 1.7633$ is the unique solution to $xe^{-1/x} = 1$ and β is a constant depending on α, and under the assumption that G is triangle-free (see also [GŠ11, GMP04] for similar results on restricted classes of graphs). However, as discussed in [GKM15], this strong spatial mixing result unfortunately does not lead to a deterministic algorithm.[2]

The newer technique of polynomial interpolation, pioneered by Barvinok [Bar17], has also recently been brought to bear on counting colorings. In a recent paper, Bencs et al. [BDPR18] use this technique to derive a FPTAS for counting colorings provided $q \geq e\Delta+1$. This result is of independent interest because it uses a different algorithmic approach, and because it establishes a new zero-free region for the associated partition function in the complex plane (see below), but it is weaker than those obtained via correlation decay.

In this chapter, we push the polynomial interpolation method further and obtain a FPTAS for counting colorings under the condition $q \geq 2\Delta$:

Theorem 4.1.1. *Fix positive integers q and Δ such that $q \geq 2\Delta$. Then there exists a fully*

[2]The strong spatial mixing condition does imply fast mixing of the Glauber dynamics, and hence an FPRAS, but only when the graph family being considered is "amenable", i.e., if the size of the ℓ-neighborhood of any vertex does not grow exponentially in ℓ. This restriction is satisfied by regular lattices, but fails, e.g., for random regular graphs.

polynomial time deterministic approximation scheme (FPTAS) for counting q-colorings in any graph of maximum degree Δ.

This is the first deterministic algorithm (of *any* kind) that for all Δ matches the "natural" bound for MCMC, first obtained by Jerrum [Jer95]. Indeed, $q \geq 2\Delta + 1$ remains the best bound known for rapid mixing of the basic Glauber dynamics that does not require either additional assumptions on the graph or a spectral comparison with another Markov chain: all the improvements mentioned above require either lower bounds on the girth and/or maximum degree, or (in the case of Vigoda's result [Vig00]) analysis of a more sophisticated Markov chain. This is for good reason, since the bound $q \geq 2\Delta + 1$ coincides with the closely related Dobrushin uniqueness condition from statistical physics [SS97], which in turn is closely related [Wei05] to the path coupling method of Bubley and Dyer [BD97] that provides the simplest currently known proof of the $q \geq 2\Delta + 1$ bound for the Glauber dynamics.

We therefore view our result as a promising starting point for deterministic coloring algorithms to finally compete with their randomized counterparts. In fact, as discussed later in section 4.1.2, our technique is capable of directly harnessing strong spatial mixing arguments used in the analysis of Markov chains for certain classes of graphs. As an example, we can exploit such an argument of Gamarnik, Katz and Misra [GKM15] to improve the bound on q in Theorem 4.1.1 when the graph is triangle-free, for all but small values of Δ. (Recall that $\alpha^\star \approx 1.7633$ is the unique positive solution of the equation $xe^{-1/x} = 1$.)

Theorem 4.1.2. *For every $\alpha > \alpha^\star$, there exists a $\beta = \beta(\alpha)$ such that the following is true. For all integers q and Δ such that $q \geq \alpha\Delta + \beta$, there exists a fully polynomial time deterministic approximation scheme (FPTAS) for counting q-colorings in any triangle-free graph of maximum degree Δ.*

We mention also that our technique applies without further effort to the more general setting of *list* colorings, where each vertex has a list of allowed colors of size q, under the same conditions as above on q. Indeed, our proofs are written to handle this more general situation.

In the next subsection we describe our algorithm in more detail.

4.1.1 Our approach

Let $G = (V, E)$ be an n-vertex graph of maximum degree Δ, and $[q] := \{1, \ldots, q\}$ a set of colors. Define the polynomial

$$Z_G(w) := \sum_{\sigma: V \to [q]} w^{|\{\{u,v\} \in E \,:\, \sigma(u) = \sigma(v)\}|}. \tag{4.1}$$

Here σ ranges over arbitrary (not necessarily proper) assignments of colors to vertices, and each such coloring has a weight $w^{m(\sigma)}$, where $m(\sigma)$ is the number of monochromatic edges in σ. Note that the number of proper q-colorings of G is just $Z_G(0)$.

The polynomial $Z_G(w)$ is the partition function of the Potts model of statistical physics, and implicitly defines a probability distribution on colorings σ according to their weights in (4.1). The parameter w measures the strength of nearest-neighbor interactions. The value $w = 1$ corresponds to the trivial setting where there is no constraint on the colors of neighboring vertices, while $w = 0$ imposes the hard constraint that no neighboring vertices receive the same color. For intermediate values $w \in [0, 1]$, neighbors with the same color are penalized by a factor of w. Theorems 4.1.1 and 4.1.2 are in fact special cases of the following more general theorem.

Theorem 4.1.3. *Suppose that the hypotheses of either Theorem 4.1.1 or Theorem 4.1.2 are satisfied, and fix $w \in [0, 1]$. Then there exists an FPTAS for the partition function $Z_G(w)$.*

Theorem 4.1.3 of course subsumes Theorems 4.1.1 and 4.1.2, but the extension to other values of w is of independent interest as the computation of partition functions is a very active area of study in statistical physics and combinatorics.

To prove Theorem 4.1.3, we view $Z_G(w)$ as a polynomial in the *complex* variable w and identify a region in the complex plane in which $Z_G(w)$ is guaranteed to have no zeros. Specifically, we will show that this holds for the open connected set $\mathcal{D}_\Delta \subset \mathbb{C}$ obtained by augmenting the real interval $[0, 1]$ with a ball of radius τ_Δ around each point, where τ_Δ is a (small) constant depending only on Δ.

Theorem 4.1.4. *Fix a positive integer Δ. Then there exists a $\tau_\Delta > 0$ and a region $\mathcal{D}_\Delta$ of the above form containing the interval $[0, 1]$ such that the following is true. For any graph G of maximum degree Δ and integer q satisfying the hypotheses of either Theorem 4.1.1 or Theorem 4.1.2, $Z_G(w) \neq 0$ when $w \in \mathcal{D}_\Delta$.*

We remark that this theorem is also of independent interest, as the location of zeros of partition functions has a long and noble history going back to the Lee-Yang theorem of the 1950s [LY52, YL52]. In the case of the Potts model, Sokal [Sok01, Sok05] proved (in the language of the Tutte polynomial) that the partition function has no zeros in w in the entire unit disk centered at 0, under the strong condition $q \geq 7.964\Delta$; the constant was later improved to 6.907 by Fernández and Procacci [FP08] (see also [JPS13]). Much more recently, the work of Bencs et al. [BDPR18] referred to above gives a zero-free region analogous to that in Theorem 4.1.4 above, but under the stronger condition $q \geq e\Delta + 1$. We note also that Barvinok and Soberón [BS16a] (see also [Bar17] for an improved version) established a zero-free region in a disk centered at $w = 1$.

Theorem 4.1.4 immediately gives our algorithmic result, Theorem 4.1.3, by appealing to the recent algorithmic paradigm of Barvinok [Bar17]. The paradigm (see Lemma 2.2.3 of [Bar17]) states that, for a partition function Z of degree m, if one can identify a simply connected, zero-free region $\mathcal{D}$ for Z in the complex plane that contains a τ-neighborhood of the interval $[0, 1]$, and a point on that interval where the evaluation of Z is easy (in our setting this is the point $w = 1$), then using the first $O\big(e^{\Theta(1/\tau)} \log(m/\varepsilon)\big)$ coefficients of Z, one can obtain a $1 \pm \varepsilon$ multiplicative approximation of $Z(x)$ at any point $x \in \mathcal{D}$. Barvinok's framework is based on exploiting the fact that the zero-freeness of Z in D is equivalent to $\log Z$ being analytic in D, and then using a carefully chosen transformation to deform D into a disk (with the easy point at the center) in order to perform a convergent Taylor expansion. The coefficients of Z are used to compute the coefficients of this Taylor expansion.

Barvinok's framework in general leads to a quasi-polynomial time algorithm as the computation of the $O\big(e^{\Theta(1/\tau)} \log(m/\varepsilon)\big)$ terms of the expansion may take quasiploynomial time $O\Big((m/\varepsilon)^{e^{\Theta(1/\tau)} \log m}\Big)$ for the partition functions considered here. However, additional insights provided by Patel and Regts [PR17a] (see, e.g., the proof of Theorem 6.2 in [PR17a]) show how to reduce this computation time to $O\Big((m/\varepsilon)^{e^{\Theta(1/\tau)} \log \Delta}\Big)$ for many models on bounded degree graphs of degree at most Δ, including the Potts model with a bounded number of colors at each vertex. Hence we obtain an FPTAS. This (by now standard) reduction is the same path as that followed by Bencs et al. [BDPR18, Corollary 1.2]; for completeness, we provide a sketch in Appendix 4.2.3. We note that for each fixed Δ and q the running time of our final algorithm is polynomial in n (the size of G) and ε^{-1}, as required for an FPTAS. However, as is typical of deterministic algorithms for approximate counting, the exponent in the polynomial depends on Δ (through the quantity τ_Δ in Theorem 4.1.4, which in the case where all lists are subsets of $[q]$, is inverse polynomial in q).

We end this section by sketching our approach to proving Theorem 4.1.4.

4.1.2 Technical overview

The starting point of our proof is a simple geometric observation, versions of which have been used before for constructing inductive proofs of zero-freeness of partition functions (see, e.g., [Bar17, BDPR18]). Fix a vertex v in the graph G. Given $w \in \mathbb{C}$, and a color $k \in [q]$, let $Z_v^{(k)}(w)$ denote the *restricted partition function* in which one only includes those colorings σ in which $\sigma(v) = k$. Then, since $Z_G(w) = \sum_{k\in[q]} Z_v^{(k)}(w)$, the zero-freeness of Z_G will follow if the angles between the complex numbers $Z_v^{(k)}(w)$, viewed as vectors in $\mathbb{R}^2$, are all small, and provided that at least one of the $Z_v^{(k)}$ is non-zero. (In fact, this condition on angles can be relaxed for those $Z_v^{(k)}(w)$ that are sufficiently small in magnitude, and this flexibility is important when w is a complex number close to 0.) Therefore, one is naturally led to consider so-called *marginal ratios*:

$$R_{G,v}^{(i,j)}(w) := \frac{Z_v^{(i)}(w)}{Z_v^{(j)}(w)}.$$

(In the q-coloring problem, this ratio is 1 by symmetry. However, in our recursive approach, we have to handle the more general list-coloring problem, in which the ratio becomes non-trivial.)

We then require that for any two colors i, j for which $Z_v^{(k)}(w)$ is large enough in magnitude, the ratio $R_{G,v}^{(i,j)}(w)$ is a complex number with small argument. This is what we prove inductively in sections 4.4 and 4.5.

The broad contours of our approach as outlined so far are quite similar to some recent work [Bar17, BDPR18]. However, it is at the crucial step of how the marginal ratios are analyzed that we depart from these previous results. Instead of attacking the restricted partition functions or the marginal ratios directly for given $w \in \mathbb{C}$, as in these previous works, we crucially exploit the fact that for any $\tilde{w} \in [0, 1]$ close to the given w, these quantities have natural probabilistic interpretations, and hence can be much better understood via probabilistic and combinatorial methods. For instance, when $\tilde{w} \in [0, 1]$, the marginal ratio $R_{G,v}^{(i,j)}(w)$ is in fact a ratio of the marginal probabilities $\Pr_{G,\tilde{w}}[\sigma(v) = i]$ and $\Pr_{G,\tilde{w}}[\sigma(v) = j]$, under the natural probability distribution on colorings σ. In fact, our analysis cleanly breaks into two separate parts:

1. First, understand the behavior of true marginal probabilities of the form $\Pr_{G,\tilde{w}}[\sigma(v) = i]$ for $\tilde{w} \in [0, 1]$. This is carried out in section 4.3.

2. Second, argue that, for complex $w \approx \tilde{w}$, the ratios $R^{(i,j)}_{G,v}(w)$ remain well-behaved. This is carried out separately for the two cases when w is close to 0 (in section 4.4) and when w is bounded away from 0 but still in the vicinity of $[0, 1]$ (in section 4.5).

A key point in our technical analysis is the notion of "niceness" of vertices, which stipulates that the marginal probability $\Pr_{G,\tilde{w}}[\sigma(v) = i] \leq \frac{1}{\deg_G(v)+2}$ where $\deg_G(v)$ is the degree of v in G (see Definition 4.3.1). Note that this condition refers only to real non-negative $\tilde{w}$, and hence is amenable to analysis via standard combinatorial tools. Indeed, our proofs that the conditions on q and Δ in Theorems 4.1.1 and 4.1.2 imply this niceness condition are very similar to probabilistic arguments used by Gamarnik et al. [GKM15] to establish the property of "strong spatial mixing" (in the special case $\tilde{w} = 0$). We emphasize that this is the only place in our analysis where the lower bounds on q are used. One can therefore expect that combinatorial and probabilistic ideas used in the analysis of strong spatial mixing and the Glauber dynamics with smaller number of colors in special classes of graphs can be combined with our analysis to obtain deterministic algorithms for those settings, as we have demonstrated in the case of [GKM15].

The above ideas are sufficient to understand the real-valued case (part 1 above). For the complex case in part 2, we start from a recurrence for the marginal ratios $R^{(i,j)}_{G,v}$ that is a generalization (to the case $w \neq 0$) of a similar recurrence used by Gamarnik et al. [GKM15] (see Lemma 4.2.4). The inductive proofs in sections 4.4 and 4.5 use this recurrence to show that, if $\tilde{w} \in [0, 1]$ is close to $w \in \mathbb{C}$, then all the relevant $R^{(i,j)}_{G,v}(w)$ remain close to $R^{(i,j)}_{G,v}(\tilde{w})$ throughout. The actual induction, especially in the case when w is close to 0, requires a delicate choice of induction hypotheses (see Lemmas 4.4.2 and 4.5.3). The key technical idea is to use the "niceness" property of vertices established in part 1 to argue that the two recurrences (real and complex) remain close at every step of the induction. This in turn depends upon a careful application of the mean value theorem, *separately* to the real and imaginary parts (see Lemma 4.2.5), of a function f_κ that arises in the analysis of the recurrence (see Lemma 4.2.6).

4.1.3 Comparison with correlation-decay based algorithms

We conclude this overview with a brief discussion of how we are able to obtain a better bound on the number of colors than in correlation decay algorithms, such as [GK12, LY13] cited earlier. In these algorithms, one first uses recurrences similar to the one mentioned above to *compute* the marginal probabilities, and then appeals to self-reducibility to compute the partition function. Of course, expanding the full tree of computations generated by the recurrence will in general give an exponential time (but exact) algorithm. The core of the analysis of these algorithms is to show that even if this tree of computations is only expanded to depth about $O(\log(n/\varepsilon))$, and the recurrence at that point is initialized with *arbitrary* values, the computation still converges to an ε-approximation of the true value. However, the requirement that the analysis be able to deal with arbitrary initializations implies that one cannot directly use properties of the actual probability distribution (e.g., the "niceness" property alluded to above); indeed, this issue is also pointed out by Gamarnik et al. [GKM15]. In contrast, our analysis does not truncate the recurrence, and thus only has to handle initializations that make sense in the context of the graph being considered. Moreover, the exponential size of the recursion tree is no longer a barrier since, in contrast to correlation decay algorithms, we are using the tree only as a tool to establish zero-freeness; the algorithm itself follows from Barvinok's polynomial interpolation paradigm. Our approach suggests that this paradigm can be viewed as a method for using (complex-valued generalizations of) strong spatial mixing results to obtain deterministic algorithms.

4.2 Preliminary

4.2.1 Colorings and the Potts model

Throughout, we assume that the graphs that we consider are augmented with a list of colors for every vertex. Formally, a graph is a triple $G = (V, E, L)$, where V is the vertex set, E is the edge set, and $L : V \to 2^{\mathbb{N}}$ specifies a list of colors for every vertex. The partition function as defined in section 4.1 generalizes naturally to this setting: the sum is now over all those colorings σ which satisfy $\sigma(v) \in L(v)$.

We also allow graphs to contain *pinned* vertices: a vertex v is said to be *pinned* to a color c if only those colorings of G are allowed in which v has color c. Suppose that a vertex v of degree d_v in a graph G is pinned to a color c, and consider the graph G' obtained by replacing v with d_v copies of itself, each of which is pinned to c and connected to exactly one of the original neighbors of v in G. It is clear that $Z_{G'}(w) = Z_G(w)$ for all w. We will therefore assume that all pinned vertices in our graphs G have degree exactly one. The *size* of graph, denoted as $|G|$, is defined to be the number of unpinned vertices. It is worth noting that the above operation of duplicating pinned vertices does not change the size of the graph.

Let G be a graph and v an unpinned vertex in G. A color c in the list of v is said to be *good* for v if for every pinned neighbor u of v is pinned to a color different from c. The set of good colors for a vertex v in graph G is denoted $\Gamma_{G,v}$. We sometimes omit the graph G and write Γ_v when G is clear from the context. A color c that is not in Γ_v is called *bad* for v. Further, given a graph G with possibly pinned vertices, we say that the graph is *unconflicted* if no two neighboring vertices in G are pinned to the same color. Note that since all pinned vertices have degree exactly one, each conflicted graph is the vertex-disjoint union of an unconflicted graph and a collection of disjoint, conflicted edges.

We will assume throughout that all unconflicted graphs G we consider have at least one proper coloring: this will be guaranteed in our applications since we will always have $|L(u)| \geq \deg_G(u) + 1$ for every unpinned vertex u in G.

Definition 4.2.1. For a graph G, a vertex v and a color $i \in L(v)$, the *restricted partition function* $Z^{(i)}_{G,v}(w)$ is the partition function restricted to colorings in which the vertex v receives color i.

Definition 4.2.2. Let ω be a formal variable. For any G, a vertex v and colors $i, j \in L(v)$, we define the *marginal ratio* of color i to color j as $R^{(i,j)}_{G,v}(\omega) := \frac{Z^{(i)}_{G,v}(\omega)}{Z^{(j)}_{G,v}(\omega)}$. Similarly we also define formally the corresponding *pseudo marginal probability* as $\mathcal{P}_{G,\omega}[c(v) = i] := \frac{Z^{(i)}_{G,v}(\omega)}{Z_G(\omega)}$.

Remark 8. Note that when a numerical value $w \in \mathbb{C}$ is substituted in place of ω in the above formal definition, $R^{(i,j)}_{G,v}(w)$ is numerically well-defined as long as $Z^{(j)}_{G,v}(w) \neq 0$, and $\mathcal{P}_{G,w}[c(v) = i]$ is numerically well-defined as long as $Z_G(w) \neq 0$. In the proof of the main theorem in sections 4.4 and 4.5, we will ensure that the above definitions are numerically instantiated only in cases where the corresponding conditions for such an instantiation to be well-defined, as stated above, are satisfied. For instance, when $w \in [0, 1]$, this is the case for the first definition when either (i) $w \neq 0$; or (ii) $w = 0$, but G is unconflicted and

$j \in \Gamma_{G,v}$; while for the second definition, this is the case when either (i) $w \neq 0$; or (ii) $w = 0$, but G is unconflicted.

Remark 9. Note also that when $w \in [0, 1]$, the pseudo probabilities, if well-defined, are actual marginal probabilities. In this case, we will also write $\mathcal{P}_{G,w}[c(v) = i]$ as $\Pr_{G,w}[c(v) = i]$. For arbitrary complex w, this interpretation as probabilities is of course not valid (since $\mathcal{P}_{G,w}[c(v) = i]$ can be non-real), but provided that $Z_G(w) \neq 0$ it is still true that

$$\sum_{i \in L(v)} \mathcal{P}_{G,w}[c(v) = i] = \frac{1}{Z_G(w)} \sum_{i \in L(v)} Z_{G,v}^{(i)}(w) = \frac{Z_G(w)}{Z_G(w)} = 1. \tag{4.2}$$

We also note that if v is pinned to color k, then $\mathcal{P}_{G,w}[c(v) = i]$ is 1 when $k = i$ and 0 when $k \neq i$.

Notation For the case $w = 0$ we will sometimes shorten the notations $\mathcal{P}_{G,0}[c(v) = i]$ and $\Pr_{G,0}[c(v) = i]$ to $\mathcal{P}_G[c(v) = i]$ and $\Pr_G[c(v) = i]$ respectively.

Definition 4.2.3 (The graphs $G_k^{(i,j)}$). Given a graph G and a vertex u in G, let $v_1, \cdots, v_{\deg_G(u)}$ be the neighbors of u. We define $G_k^{(i,j)}$ (the vertex u will be understood from the context) to be the graph obtained from G as follows:

- first we replace vertex u with $u_1, \cdots, u_{\deg_G(u)}$, and connect u_1 to v_1, u_2 to v_2, and so on;
- next we pin vertices $u_1, \cdots, u_{k-1}$ to color i, and vertices $u_{k+1}, \cdots, u_{\deg_G(u)}$ to color j;
- finally we remove the vertex u_k.

Note that the graph $G_k^{(i,j)}$ has one fewer unpinned vertex than G. Moreover, $u_1, \cdots, u_{\deg_G(u)}$ are of degree 1, so this construction maintains the property that pinned vertices have degree 1.

We now derive a recurrence relation between the marginal ratios of the graph G and pseudo marginal probabilities of the graphs $G_k^{(i,j)}$. This is an extension to the Potts model of a similar recurrence relation derived by Gamarnik, Katz and Misra [GKM15] for the special case of colorings (that is, $w = 0$).

Lemma 4.2.4. *Let ω be a formal variable. For a graph G, a vertex u and colors $i, j \in L(u)$, we have*

$$R_{G,u}^{(i,j)}(\omega) = \frac{\prod_{k=1}^{\deg_G(u)} \Big(1 - \gamma \cdot \mathcal{P}_{G_k^{(i,j)},\omega}[c(v_k) = i]\Big)}{\prod_{k=1}^{\deg_G(u)} \Big(1 - \gamma \cdot \mathcal{P}_{G_k^{(i,j)},\omega}[c(v_k) = j]\Big)},$$

where we define $\gamma := 1 - \omega$. In particular, when a numerical value $w \in \mathbb{C}$ is substituted in place of ω, the above recurrence is valid as long as the quantities $Z_{G_k^{(i,j)}}(w)$ and $1 - \gamma \cdot \mathcal{P}_{G_k^{(i,j)},w}[c(v_k) = j]$ for $1 \le k \le \deg_G(u)$ are all non-zero.

Proof. For $0 \le k \le \deg_G(u)$, let H_k be the graph obtained from G as follows:

- first we replace vertex u with $u_1, \cdots, u_{\deg_G(u)}$, and connect u_1 to v_1, u_2 to v_2, and so on;
- we then pin vertices $u_1, \cdots, u_k$ to color i, and vertices $u_{k+1}, \cdots, u_{\deg_G(u)}$ to color j.

Note that H_k is the same as $G_k^{(i,j)}$, except that the last step of the construction of $G_k^{(i,j)}$ is skipped, i.e, the vertex u_k is not removed, and, further, u_k is pinned to color i. We can now write

$$R_{G,u}^{(i,j)}(\omega) = \frac{Z_{G,u}^{(i)}(\omega)}{Z_{G,u}^{(j)}(\omega)} = \frac{Z_{H_{\deg_G(u)}}(\omega)}{Z_{H_0}(\omega)} = \prod_{k=1}^{\deg_G(u)} \frac{Z_{H_k}(\omega)}{Z_{H_{k-1}}(\omega)}.$$

Next, for $1 \le k \le \deg_G(u)$, let $Y_k := Z_{G_k^{(i,j)}}(\omega)$ and $Y_k^{(i)} := Z_{G_k^{(i,j)},v_k}^{(i)}(\omega)$. We observe that

$$\mathcal{P}_{G_k^{(i,j)},\omega}[c(v_k) = i] = \frac{Y_k^{(i)}}{Y_k},$$
$$Z_{H_k}(\omega) = Y_k - (1-\omega) \cdot Y_k^{(i)},$$
$$Z_{H_{k-1}}(\omega) = Y_k - (1-\omega) \cdot Y_k^{(j)}.$$

Therefore we have

$$R_{G,u}^{(i,j)}(\omega) = \prod_{k=1}^{\deg_G(u)} \frac{Y_k - (1-\omega) \cdot Y_k^{(i)}}{Y_k - (1-\omega) \cdot Y_k^{(j)}} = \frac{\prod_{k=1}^{\deg_G(u)} \Big(1 - \gamma \cdot \mathcal{P}_{G_k^{(i,j)},\omega}[c(v_k) = i]\Big)}{\prod_{k=1}^{\deg_G(u)} \Big(1 - \gamma \cdot \mathcal{P}_{G_k^{(i,j)},\omega}[c(v_k) = j]\Big)},$$

where $\gamma = 1 - \omega$. The claim about the validity of the recurrence on numerical substitution then follows from the conditions outlined in Definition 4.2.2. $\square$

4.2.2 Complex analysis

We start with a consequence of the mean value theorem for complex functions, specifically tailored to our application. Let D be any domain in $\mathbb{C}$ with the following properties.

- For any $z \in D$, $\Re z \in D$.
- For any $z_1, z_2 \in D$, there exists a point $z_0 \in D$ such that one of the numbers $z_1 - z_0, z_2 - z_0$ has zero real part while the other has zero imaginary part.
- If $z_1, z_2 \in D$ are such that either $\Im z_1 = \Im z_2$ or $\Re z_1 = \Re z_2$, then the segment $[z_1, z_2]$ lies in D.

We remark that a rectangular region symmetric about the real axis will satisfy all the above properties.

Lemma 4.2.5 (Mean value theorem for complex functions). *Let f be a holomorphic function on D such that for $z \in D$, $\Im f(z)$ has the same sign as $\Im z$. Suppose further that there exist positive constants ρ_I and ρ_R such that*

- *for all $z \in D$, $|\Im f'(z)| \le \rho_I$;*
- *for all $z \in D$, $\Re f'(z) \in [0, \rho_R]$.*

Then for any $z_1, z_2 \in D$, there exists $C_{z_1,z_2} \in [0, \rho_R]$ such that

$$|\Re(f(z_1) - f(z_2)) - C_{z_1,z_2} \cdot \Re(z_1 - z_2)| \le \rho_I \cdot |\Im(z_1 - z_2)|, \text{ and}$$

$$|\Im(f(z_1) - f(z_2))| \le \rho_R \cdot \begin{cases} |\Im(z_1 - z_2)|, & \text{when } (\Im z_1) \cdot (\Im z_2) \le 0. \\ \max\{|\Im z_1|, |\Im z_2|\} & \text{otherwise.} \end{cases}$$

Proof. We write $f = u + \iota v$, where $u, v : D \to \mathbb{R}$ are seen as differentiable functions from $\mathbb{R}^2$ to $\mathbb{R}$ satisfying the Cauchy-Riemann equations

$$u^{(1,0)} = v^{(0,1)} \quad \text{and} \quad u^{(0,1)} = -v^{(1,0)}.$$

This implies in particular that $\Re f'(z) = u^{(1,0)}(z) = v^{(0,1)}(z)$ and $\Im f'(z) = v^{(1,0)}(z) = -u^{(0,1)}(z)$.

Let z_0 be a point in D such that $\Re(z_2 - z_0) = 0$ and $\Im(z_1 - z_0) = 0$ (by the conditions imposed on D, such a z_0 exists, possibly after interchanging z_1 and z_2). Now we have

$$\begin{aligned}\Re(f(z_1) - f(z_2)) &= u(z_1) - u(z_0) + u(z_0) - u(z_2)\\ &= u^{(1,0)}(z') \cdot \Re(z_1 - z_0) + u(z_0) - u(z_2),\end{aligned}$$

where z' is a point lying on the segment $[z_0, z_1]$, obtained by applying the standard mean value theorem to the function u along this segment (note that the segment is parallel to the real axis). On the other hand, since the segment $[z_0, z_2]$ is parallel to the imaginary axis, we apply the standard mean value theorem to the real valued function u to get (after recalling that $\left|u^{(0,1)}(z)\right| = |\Im f'(z)| \le \rho_I$ for all $z \in D$)

$$|u(z_0) - u(z_2)| \le \rho_I \left|\Im(z_2 - z_0)\right| = \rho_I \left|\Im(z_2 - z_1)\right|.$$

This proves the first part, once we set $C_{z_1,z_2} = u^{(1,0)}(z') = \Re f'(z')$, which must lie in $[0, \rho_R]$ since $z' \in D$.

For the second part, we note that since $\Im f(z) = 0$ when $\Im z = 0$, we have for $z \in D$,

$$\begin{aligned}\Im f(z) = \Im(f(z) - f(\Re z)) &= v(z) - v(\Re z)\\ &= v^{(0,1)}(z') \cdot \Im z,\end{aligned}$$

where z' is a point lying on the segment $[z, \Re z]$, obtained by applying the standard mean value theorem to the function v along this segment (note that the segment is parallel to the imaginary axis).

Since $v^{(0,1)}(z') = u^{(1,0)}(z') \in [0, \rho_R]$ for all $z' \in D$, there exist $a, b \in [0, \rho_R]$ such that

$$|\Im(f(z_1) - f(z_2))| = |a\Im z_1 - b\Im z_2|,$$

so that we get

$$|\Im(f(z_1) - f(z_2))| = |a\Im z_1 - b\Im z_2| \le \rho_R \cdot \begin{cases} |\Im(z_1 - z_2)|, & \text{when } (\Im z_1) \cdot (\Im z_2) \le 0. \\ \max\{|\Im z_1|, |\Im z_2|\} & \text{otherwise.} \end{cases}$$

□

We will apply the above lemma to the function

$$f_\kappa(x) := -\ln(1 - \kappa e^x), \tag{4.3}$$

which, as we shall see later, will play a central role in our proofs. (We note that here, and also later in the paper, we use ln to denote the principal branch of the complex logarithm; i.e., if $z = re^{\iota\theta}$ with $r > 0$ and $\theta \in (-\pi, \pi)$, then $\ln z = \ln r + \iota\theta$.) Below we verify that such an application is valid, and record the consequences.

Lemma 4.2.6. *Consider the domain D given by*

$$D := \{z \mid \Re z \in (-\infty, -\zeta) \text{ and } |\Im z| < \tau\},$$

where $\tau < 1/2$ and ζ are positive real numbers such that $\tau^2 + e^{-\zeta} < 1$. Suppose $\kappa \in [0, 1]$ and consider the function f_κ as defined in eq. (4.3). Then,

1. *The function f_κ and the domain D satisfy the hypotheses of Lemma 4.2.5, if ρ_R and ρ_I in the statement of the theorem are taken to be $\frac{e^{-\zeta}}{1-e^{-\zeta}}$ and $\frac{\tau \cdot e^{-\zeta}}{\left(1-e^{-\zeta}\right)^2}$, respectively.*

2. *If $\varepsilon > 0$ and κ' are such that $|\kappa' - \kappa| < \varepsilon$ and $(1 + \varepsilon) < e^\zeta$, then for any $z \in D$,*

$$|f_{\kappa'}(z) - f_\kappa(z)| \le \frac{\varepsilon}{e^\zeta - 1 - \varepsilon}.$$

In particular, we note that the domain D is indeed rectangular and symmetric about the real axis.

Proof. The domain D is rectangular and symmetric about the real axis, so it clearly satisfies the conditions. We also note that since $\kappa \le 1$, $f_\kappa(z)$ is well defined when $\Re z < 0$, and maps real numbers in D to real numbers. Further, a direct calculation shows that $\Im f_\kappa(z) = -\arg(1 - \kappa e^z)$ has the same sign as $\sin(\Im z)$ when $\Re z < 0$ (since $\kappa \in [0, 1]$). Since $|\Im z| \le \tau < \pi$, we see therefore that $\Im f_\kappa(z)$ has the same sign as $\Im z$, and hence f_κ satisfies the hypothesis of Lemma 4.2.5.

Note that $f'_\kappa(z) = \frac{\kappa e^z}{1-\kappa e^z}$. A direct calculation then shows that $\Re f'_\kappa(z) = \frac{\kappa \Re e^z - \kappa^2 |e^z|^2}{|1-\kappa e^z|^2}$ and $\Im f'_\kappa(z) = \frac{\kappa \Im e^z}{|1-\kappa e^z|^2}$. Now, for $z \in D$, $|\arg e^z| \le \tau$, so that $\Re e^z \ge |e^z| \cos \arg e^z \ge |e^z| (1 - \tau^2)$. Thus, we see that $\kappa \Re e^z - \kappa^2 |e^z|^2 \ge \kappa |e^z| \left(1 - \tau^2 - \kappa |e^z|\right) \ge \kappa |e^z| \left(1 - \tau^2 - \kappa e^{-\zeta}\right)$. Since $\kappa \in [0, 1]$ and $\tau^2 + e^{-\zeta} < 1$ by assumption, we therefore have $\Re f'_\kappa(z) \ge 0$. Further $\Re f'_\kappa(z) \le |f'_\kappa(z)| = \frac{\kappa |e^z|}{|1-\kappa e^z|} \le \frac{\kappa |e^z|}{1-\kappa |e^z|} \le \frac{\kappa e^{-\zeta}}{1-e^{-\zeta}}$, since $\kappa \in [0, 1]$. Together, these show that $\Re f'_\kappa(z) \in \left[0, \frac{e^{-\zeta}}{1-e^{-\zeta}}\right]$ for $z \in D$, so that the claimed choice of the parameter ρ_R in Lemma 4.2.5 is justified.

Similarly, for the imaginary part, we have $|\Im f'_\kappa(z)| = \frac{\kappa|\Im e^z|}{|1-\kappa e^z|^2}$, which in turn is at most $\frac{\kappa\cdot\tau\cdot e^{-\zeta}}{(1-\kappa e^{-\zeta})^2}$ for $z \in D$. Since $\kappa \in [0,1]$, this justifies the choice of the parameter ρ_I.

We now turn to the second item of the observation. The derivative of $f_x(z)$ with respect to x is $\frac{e^z}{1-xe^z}$, which for x within distance ε (satisfying $(1+\varepsilon) < e^\zeta$) of κ and $z \in D$ has length at most $\frac{1}{e^\zeta-1-\varepsilon}$. Thus, the standard mean value theorem applied along the segment $[\kappa, \kappa']$ (which is of length at most ε) yields the claim. □

We will also need the following simple geometric lemma, versions of which have been used in the work of Barvinok [Bar17] and also Bencs et al. [BDPR18].

Lemma 4.2.7. *Let $z_1, z_2, \ldots, z_n$ be complex numbers such that the angle between any two non-zero z_i is at most $\alpha \in [0, \pi/2)$. Then $|\sum_{i=1}^n z_i| \geq \cos(\alpha/2)\sum_{i=1}^n |z_i|$.*

Proof. Fix a non-zero z_i, and without loss of generality let z_1 and z_2 be the non-zero elements giving the maximum and minimum values, respectively, of the quantity $\arg(z_j/z_i)$, as z_j varies over all the non-zero elements (breaking ties arbitrarily). Consider the ray z bisecting the angle between z_1 and z_2. Then, by the assumption, the angle made by z and any of the non-zero z_i is at most $\alpha/2$, so that the projection of z_i on z is of length at least $|z_i|\cos(\alpha/2)$ and is in the same direction as z. Thus, denoting by S' the projection of $S = \sum_{i=1}^n z_i$ on z, we have

$$|S| \geq |S'| \geq \sum_{i=1}^n |z_i| \cos(\alpha/2). \qquad \square$$

4.2.3 Sketch of the algorithm

In this subsection we outline how to apply Barvinok's algorithmic paradigm to translate our zero-freeness result (Theorem 4.1.4) into the FPTAS claimed in Theorem 4.1.3. Let G be a graph with n vertices and m edges and maximum degree Δ. Recall that our goal is to obtain a $1 \pm \varepsilon$ approximation of the Potts model partition function $Z_G(w)$ at any point $w \in [0,1]$. Note that Z_G is a polynomial of degree m, and that computing Z_G at $w = 1$ is trivial since $Z_G(1) = q^n$. Recall also that Theorem 4.1.4 ensures that Z_G has no zeros in the region $\mathcal{D}_\Delta$ of width τ_Δ around the real interval $[0,1]$. For technical convenience we will actually work with a slightly smaller zero-free region consisting of the rectangle

$$\mathcal{D}'_\Delta = \{w \in \mathbb{C} : -\tau'_\Delta \leq \Re w \leq 1 + \tau'_\Delta;\ |\Im w| \leq \tau'_\Delta\},$$

where $\tau'_\Delta = \tau_\Delta/\sqrt{2}$. Note that $\mathcal{D}'_\Delta \subset \mathcal{D}_\Delta$ so $\mathcal{D}'_\Delta$ is also zero-free. In the rest of this section, we drop the subscript Δ from these quantities.

Now let $f(z)$ be a complex polynomial of degree d for which $f(0)$ is easy to evaluate, and suppose we wish to approximate $f(1)$. Barvinok's basic paradigm [Bar17, Section 2.2] achieves this under the assumption that f has no zeros in the open disk $\mathcal{B}(0, 1+\delta)$ of radius $1+\delta$ centered at 0: the approximation simply consists of the first $k = O(\frac{1}{\delta}\log(\frac{d}{\varepsilon\delta}))$ terms of the Taylor expansion of $\log f$ around 0. (Note that this expansion is absolutely convergent within $\mathcal{B}(0, 1+\delta)$ by the zero-freeness of f.) These terms can in turn be expressed as linear combinations of the first k coefficients of f itself. We now sketch how to reduce our computation of $Z_G(w)$ to this situation.

First, for any fixed $w \in [0,1]$, define the polynomial $g(z) := Z_G(z(w-1)+1)$. Note that $g(0) = Z_G(1)$ is trivial, while $g(1) = Z_G(w)$ is the value we are trying to compute. Moreover, plainly $g(z) \neq 0$ for all $z \in \mathcal{D}'$. Next, define a polynomial $\phi : \mathbb{C} \to \mathbb{C}$ that maps the disk $\mathcal{B}(0, 1+\delta)$ into the rectangle $\mathcal{D}'$, so that $\phi(0) = 0$ and $\phi(1) = 1$; Barvinok [Bar17, Lemma 2.2.3] gives an explicit construction of such a polynomial, with degree $N = \exp(\Theta(\tau^{-1}))$ and with $\delta = \exp(-\Theta(\tau^{-1}))$. Now we have reduced the computation of $Z_G(w)$ to that of $f(1)$, where $f(z) := g(\phi(z))$ is a polynomial of degree $\deg(g)\cdot\deg(\phi) = mN$ that is non-zero on the disk $\mathcal{B}(0, 1+\delta)$, so the framework of the previous paragraph applies. Note that the number of terms required in the Taylor expansion of $\log f$ is $k = O(\frac{1}{\delta}\log(\frac{mN}{\varepsilon\delta})) = \exp(O(\tau^{-1}))\log(\frac{n\Delta}{\varepsilon})$.

Naive computation of these k terms requires time $n^{\Theta(k)}$, which yields only a quasi-polynomial algorithm since k contains a factor of $\log n$. This complexity comes from the need to enumerate all colorings of subgraphs induced by up to k edges. However, a technique of Patel and Regts [PR17a], based on Newton's identities and an observation of Csikvari and Frenkel [CF16], can be used to reduce this computation to an enumeration over subgraphs induced by *connected* sets of edges (see [PR17a, Section 6] for details). Since G has bounded degree, this reduces the complexity to $\Delta^{O(k)} = (\frac{n\Delta}{\varepsilon})^{\log(\Delta)\exp(O(\tau^{-1}))}$. For any fixed Δ this is polynomial in (n/ε), thus satisfying the requirement of a FPTAS.

Note that the degree of the polynomial is exponential in τ^{-1}; since τ^{-1} in turn is exponential in Δ (see the discussion following the proof of Theorem 4.1.4), the degree of the polynomial is doubly exponential in Δ. The same discussion explains how this can be improved to singly exponential for the case of uniformly large list sizes.

4.3 Properties of the real-valued recurrence

In this section we prove some basic properties of the real-valued recurrence established in Lemma 4.2.4, that is, in the case where $w \in [0,1]$ is real (and hence, $\gamma = 1 - w \in [0,1]$).

We remark that in all graphs G appearing in our analysis, we will be able to assume that for any unpinned vertex u in G, $|L(u)| \geq \deg_G(u) + 1$. Thus, $Z_G(w) \neq 0$ whenever either (i) $w \in (0,1]$; or (ii) $w = 0$, but G is unconflicted. As discussed in the previous section, this implies that the marginal ratios and the pseudo marginal probabilities are well-defined, and, further, the latter are actual probabilities. Moreover, if G is not connected, and G' is the connected component containing u, then we have $R^{(i,j)}_{G,u}(w) = R^{(i,j)}_{G',u}(w)$ and $\mathcal{P}_{G,w}[c(u) = i] = \mathcal{P}_{G',w}[c(u) = i]$. Thus without loss of generality, we will only consider connected graphs in this section.

We now formally state the conditions on the list sizes under which our main theorem holds.

Condition 1 (**Large lists**). *The graph G satisfies at least one of the following two conditions.*

1. $|L(v)| \geq \max\{2, 2 \cdot \deg_G(v)\}$ *for each unpinned vertex v in G.*

2. *The graph G is triangle-free and further, for each vertex v of G,*

$$|L(v)| \geq \alpha \cdot \deg_G(v) + \beta,$$

where α is any fixed constant larger than the unique positive solution $\alpha^\star$ of the equation $xe^{-\frac{1}{x}} = 1$ and $\beta = \beta(\alpha) \geq 2\alpha$ is a constant chosen so that $\alpha \cdot e^{-\frac{1}{\alpha}(1+\frac{1}{\beta})} \geq 1$. We note that $\alpha^\star$ lies in the interval $[1.763, 1.764]$, and β as chosen above is at least $7/2$.

Remark 10. Note that the condition $|L(v)| \geq 2$ imposed in case 1 above is without loss of generality, since any vertex with $|L(v)| = 1$ can be removed from G after removing the unique color in its list from the lists of its neighbors, without changing the number of colorings of G.

As stated in section 4.1, an important element of our analysis is going to be the fact that under Condition 1, one can show that certain vertices are “nice” in the sense of the following definition. We emphasize that Condition 1 is ancillary to our main technical development: any condition under which the probability bounds imposed in the following definition can be proved (as is done in Lemma 4.3.2 below) will be sufficient for the analysis.

Definition 4.3.1. Given a graph G and an unpinned vertex u in G, let d be the number of unpinned neighbors of u. We say the vertex u is *nice* in G if for any $w \in [0,1]$ and any color $i \in L(u)$, $\Pr_{G,w}[c(u) = i] \leq \frac{1}{d+2}$.

Remark 11. We adopt the convention that if G is a conflicted graph (so that it has no proper colorings) and $w = 0$, then $\Pr_{G,w}[c(u) = i] = 0$ for every color i and every unpinned vertex u in G. This is just to simplify the presentation in this section by avoiding the need to explicitly exclude this case from the lemmas below. In the proof of our main result in sections 4.4 and 4.5, we will never consider conflicted graphs in a situation where w could be 0, so that this convention will then be rendered moot.

Lemma 4.3.2. *If G satisfies Condition 1 then for any vertex u in G, and any unpinned neighbor v_k of u, we have that v_k is nice in $G_k^{(i,j)}$.*

We prove this lemma separately for each of the two cases in Condition 1.

4.3.1 Analysis for case 1 of Condition 1

Lemma 4.3.3. *Let G be a graph that satisfies case 1 of Condition 1. Then for any unpinned vertex u in G, and any unpinned neighbor v_k of u, we have that v_k is nice in $G_k^{(i,j)}$.*

Proof. For ease of notation, we denote $G_k^{(i,j)}$ by H and v_k by v. Since G satisfies case 1 of Condition 1, and $\deg_H(v) = \deg_G(v_k) - 1$ (since the neighbor u of v_k in G is dropped in the construction of $H = G_k^{(i,j)}$), we have $|L_H(v)| = |L_G(v_k)| \geq 2\deg_G(v_k) \geq 2 \cdot \deg_H(v) + 2$.

Consider any valid coloring[3] σ' of the neighbors of v in H. For $k \in L_H(v)$, let n_k denote the number of neighbors of v that are colored k in σ'. Then for any $w \in [0,1]$ and $i \in L_H(v)$,

$$\Pr_{H,w}\left[c(v) = i|\sigma'\right] = \frac{w^{n_i}}{\sum_{j \in L_H(v)} w^{n_j}} \leq \frac{1}{|L_H(v)| - \deg_H(v)},$$

since at most $\deg_H(v)$ of the n_j can be positive. Note in particular that if i is not a good color for v in H, then the probability is 0. Since this holds for any coloring σ', we have

[3]Here, we say that a coloring σ is *valid* if the color σ assigns to any vertex v is from $L(v)$, and further, in case $w = 0$, no two neighbors are assigned the same color by σ.

$\Pr_{H,w}[c(v) = i] \leq \frac{1}{|L_H(v)| - \deg_H(v)}$. Now, let d be the number of unpinned neighbors of v in H. Noting that $\deg_H(v) \geq d$, and recalling the observation above that $|L_H(v)| \geq 2\deg_H(v) + 2$, we thus have

$$\Pr_{G_k^{(i,j)},w}[c(v_k) = i] = \Pr_{H,w}[c(v) = i] \leq \frac{1}{|L_H(v)| - \deg_H(v)} \leq \frac{1}{d+2}.$$

Thus v_k is nice in $G_k^{(i,j)}$. □

4.3.2 Analysis for case 2 of Condition 1

Notice that if G satisfies case 2 of Condition 1, then so does $G_k^{(i,j)}$. Thus in order to show that v_k is nice in $G_k^{(i,j)}$, it suffices to show the following more general fact.

Lemma 4.3.4. *Let G be any graph that satisfies case 2 of Condition 1, and let u be any unpinned vertex in G, then u is nice in G.*

The proof of this lemma is almost identical to arguments that appear in the work of Gamarnik, Katz and Misra [GKM15] on strong spatial mixing; we include a proof here for completeness.

Proof. We show first that $\Pr_{G,w}[c(u) = i] \leq \frac{1}{\beta}$ whenever $L_G(u) \geq \deg_G(u) + \beta$; this will be required later in the proof. To do so, we repeat the arguments in the proof of Lemma 4.3.3 to see that $\Pr_{G,w}[c(u) = i] \leq \frac{1}{|L(u)| - \deg_G(u)}$. The claimed bound then follows since $|L(u)| - \deg_G(u) \geq \beta$.

Next we show that the upper bound of $\frac{1}{d+2}$, where d is the number of unpinned neighbors of u in G, holds conditioned on every coloring of the neighbors of the (unpinned) neighbors of u, by following a similar path as in [GKM15]. Consider any valid coloring[4] σ' of the vertices at distance *two* from u. Since G is triangle free, we claim that conditional on σ' there is a tree T of depth 2 rooted at u, with all the leaves pinned according to σ', such that

$$\Pr_{G,w}\left[c(u) = i | \sigma'\right] = \Pr_{T,w}[c(u) = i]. \tag{4.4}$$

To see this, notice that once we condition on the coloring of the vertices at distance 2 from u, the distribution of the color at u becomes independent of the distribution of colors of vertices at distance 3 or more. Further, because of triangle freeness, no two neighbors of

[4]Here, we say that a coloring σ is *valid* if the color σ assign to any vertex v is from $L(v)$, and further, in case $w = 0$, no two neighbors are assigned the same color by σ.

u have an edge between them, and hence any cycle in the distance-2 neighborhood, if one exists, must go through at least one pinned vertex. We then observe that such a cycle can be broken by replacing any pinned vertex v' in it with $\deg(v')$ copies, one for each of its neighbor: as discussed earlier, this operation cannot change the partition function or probabilities. This operation therefore ensures that every pinned vertex in the resulting graph is now a leaf of a tree T of depth 2 rooted at u. Further, in T, the root u has d unpinned children, and all vertices at depth 2 are pinned according to σ'.

Let $v_1, \cdots, v_d$ be the d unpinned neighbors of u in T, and let $T_1, \cdots, T_d$ be the subtrees rooted at $v_1, \cdots, v_d$ respectively. For each $k \in L_G(u)$, let n_k be the number of neighbors of u that are pinned to color k. Then by Lemma 4.2.4,

$$R^{(j,i)}_{T,u}(w) = \frac{w^{n_j} \cdot \prod_{k=1}^{d}(1 - \gamma \cdot \mathcal{P}_{T_k,w}[c(v_k) = j]\,)}{w^{n_i} \cdot \prod_{k=1}^{d}(1 - \gamma \cdot \mathcal{P}_{T_k,w}[c(v_k) = i]\,)}.$$

Define $t_{kj} := \gamma \cdot \Pr_{T_k,w}[c(v_k) = j]$, and note that from the calculation at the beginning of the proof, we have $0 \le t_{kj} \le \frac{\gamma}{\beta} \le \frac{1}{\beta} \le 1/2$. Note also that $t_{kj} = 0$ if $j \notin L(v_k)$. Thus, we have

$$\sum_{j\in\Gamma_u} t_{kj} = \gamma \sum_{j\in\Gamma_u\cap L(v_k)} \Pr_{T_k,w}[c(v_k) = j] \le \gamma \le 1. \tag{4.5}$$

Therefore,

$$\Pr_{T,w}[c(u) = i] = \frac{1}{\sum_{j\in L(v)} R^{(j,i)}_{T,v}(w)} = \frac{w^{n_i} \cdot \prod_{k=1}^{d}(1 - t_{ki})}{\sum_{j\in L(u)} w^{n_j} \cdot \prod_{k=1}^{d}(1 - t_{kj})} \le \frac{1}{\sum_{j\in\Gamma_u} \prod_{k=1}^{d}(1 - t_{kj})}, \tag{4.6}$$

where, in the last inequality we use that $n_j = 0$ when j is good for u in G, and also that $w \in [0, 1]$.

Since $\Pr_{G,w}[c(u) = i|\sigma'] = \Pr_{T,w}[c(u) = i]$, it remains to lower bound the denominator $\sum_{j\in\Gamma_u} \prod_{k=1}^{d}(1 - t_{kj})$. We begin by recalling the following standard consequence of the Taylor expansion of $\ln(1 - x)$ around 0: when $0 \le x \le \frac{1}{\beta} < 1$, and β is such that $(1 - 1/\beta)^2 \ge 1/2$,

$$\ln(1 - x) \ge -x - \frac{x^2}{2(1 - 1/\beta)^2} \ge -x - x^2 \ge -\left(1 + \frac{1}{\beta}\right) x. \tag{4.7}$$

Note that the condition required of β is satisfied since $\beta \ge 2\alpha \ge 7/2$, as stipulated in case 2 of Condition 1. Since $0 \le t_{kj} \le 1/\beta$, we therefore obtain, for every $j \in \Gamma_u$,

$$\prod_{k=1}^{d}(1 - t_{kj}) \ge \prod_{k=1}^{d} \exp\left(-\left(1 + \frac{1}{\beta}\right) t_{kj}\right) = \exp\left(-\left(1 + \frac{1}{\beta}\right) \sum_{k=1}^{d} t_{kj}\right). \tag{4.8}$$

For convenience of notation, we denote $|\Gamma_u|$ by q_u. Note that since $|L(u)| \geq \alpha \deg(u) + \beta$, and u has $\deg(u) - d$ pinned neighbors, we have

$$q_u \geq |L(u)| - (\deg(u) - d) \geq |L(u)| - \alpha(\deg(u) - d) \geq \alpha d + \beta,$$

where in the second inequality we use $\alpha \geq 1$. Now, by the AM-GM inequality, we get

$$\begin{aligned}
\sum_{j \in \Gamma_u} \prod_{k=1}^{d} (1 - t_{kj}) &\geq q_u \left(\prod_{j \in \Gamma_u} \prod_{k=1}^{d} (1 - t_{kj}) \right)^{\frac{1}{q_u}} \\
&\geq q_u \exp\left(-\frac{1}{q_u}(1 + 1/\beta) \cdot \sum_{k=1}^{d} \sum_{j \in \Gamma_u} t_{kj} \right), && \text{using eq. (4.8)} \\
&\geq (\alpha d + \beta) \exp\left(-\frac{d(1 + 1/\beta)}{\alpha d + \beta} \right), && \text{using eq. (4.5) and } q_u \geq \alpha d + \beta \\
&\geq (d + 2)\alpha \cdot \exp\left(-\frac{(1 + 1/\beta)}{\alpha} \right), && \text{using } \beta \geq 2\alpha \\
&\geq (d + 2),
\end{aligned}$$

where the last line uses the stipulation in case 2 of Condition 1 that α and β satisfy $\alpha \cdot \exp\left(-\frac{(1+1/\beta)}{\alpha}\right) \geq 1$. From eqs. (4.4) and (4.6) we therefore get

$$\Pr_{G,w}\left[c(u) = i | \sigma'\right] \leq \frac{1}{d + 2}.$$

Since this holds for any conditioning σ' of the colors of the neighbors of the neighbors of u in G, we then have

$$\Pr_{G,w}[c(u) = i] \leq \frac{1}{d + 2},$$

which concludes the proof.

□

The proof of Lemma 4.3.2 is immediate from Lemmas 4.3.3 and 4.3.4.

Proof of Lemma 4.3.2. If G satisfies case 1 of Condition 1 then we apply Lemma 4.3.3. If G satisfies case 2 of Condition 1 then we apply Lemma 4.3.4 after noting that if G satisfies case 2 of Condition 1, then so does $G_k^{(i,j)}$, and further that, as assumed in the hypothesis of Lemma 4.3.2, v_k is unpinned in $G_k^{(i,j)}$. □

We conclude this section by noting that, the niceness condition can be strengthened in the case when all the list sizes are uniformly large (e.g., as in the case of q-colorings).

Remark 12. In Condition 1, if we replace the degree of a vertex by the maximum degree Δ (e.g., in case 1 of the condition, if we assume $|L(v)| \geq 2\Delta$, instead of $2\deg_G(v)$, for each v), then for every vertex v in the graph G, it holds that $\Pr_{G,w}[c(v) = i] < \min\left\{\frac{4}{3\Delta}, 1\right\}$.

To see this, notice that the same calculation as in the proof of Lemma 4.3.3 above gives $\Pr_{G,w}[c(v) = i] \leq \frac{1}{|L(v)|-\Delta} \leq \frac{1}{(\alpha-1)\Delta+\beta} \leq \frac{1}{(\alpha-1)\Delta} < \frac{4}{3\Delta}$. We will refer to this stronger condition on list sizes (which holds, in particular, when one is considering the case of q-colorings), as the *uniformly large list size* condition.

4.4 Zero-free region for small $|w|$

As explained in section 4.2.3, all our algorithmic results follow from Theorem 4.1.4, which establishes a zero-free region for the partition function $Z_G(w)$ around the interval $[0, 1]$ in the complex plane. We split the proof of Theorem 4.1.4 into two parts: in this section, we establish the existence of a zero-free disk around the endpoint $w = 0$ (see Theorem 4.4.1): this is the most delicate case because $w = 0$ corresponds to proper colorings. Then in section 4.5 (see Theorem 4.5.1) we derive a zero-free region around the remainder of the interval, using a similar but less delicate approach. Taken together, Theorems 4.4.1 and 4.5.1 immediately imply Theorem 4.1.4, so this will conclude our analysis.

Theorem 4.4.1. *Fix a positive integer Δ. There exists a $\nu_w = \nu_w(\Delta)$ such that the following is true. Let G be a graph of maximum degree Δ satisfying Condition 1, and having no pinned vertices. Then, $Z_G(w) \neq 0$ for any w satisfying $|w| \leq \nu_w$.*

In the proof, we will encounter several constants which we now fix. Given the degree bound $\Delta \geq 1$, we define

$$\varepsilon_R := \frac{0.01}{\Delta^2},\ \varepsilon_I := \varepsilon_R \cdot \frac{0.01}{\Delta^2},\ \text{and } \varepsilon_w := \varepsilon_I \cdot \frac{0.01}{\Delta^3}. \tag{4.9}$$

We will then see that the quantity ν_w in the statement of the theorem can be chosen to be $0.2\varepsilon_w/2^{\Delta}$. (In fact, we will show that if one has the slightly stronger assumption of uniformly large list sizes considered in Remark 12, then ν_w can be chosen to be $\varepsilon_w/(300\Delta)$).

Throughout the rest of this section, we fix Δ to be the maximum degree of the graphs, and let $\varepsilon_w, \varepsilon_I, \varepsilon_R$ be as above.

We now briefly outline our strategy for the proof. Recall that, for a vertex u and colors i, j, the marginal ratio is given by $R_{G,u}^{(i,j)}(w) = \frac{Z_{G,u}^{(i)}(w)}{Z_{G,u}^{(j)}(w)}$. When G is an unconflicted graph, $R^{(i,j)}_{G,u}(0)$ is always a well-defined non-negative real number. Intuitively, we would like to show that $R^{(i,j)}_{G,u}(w) \approx R^{(i,j)}_{G,u}(0)$, independent of the size of G, when $w \in \mathbb{C}$ is close to 0. Given such an approximation one can use a simple geometric argument (see Consequence 4.4.3) to conclude that the partition function does not vanish for such w. In order to prove the above approximate equality inductively for a given graph G, we take an approach that exploits the properties of the "real" case (i.e., of $R^{(i,j)}_{G,u}(0)$) and then uses the notion of "niceness" of certain vertices described earlier to control the accumulation of errors. To this end, we will prove the following lemma via induction on the number of unpinned vertices in G. Theorem 4.4.1 will follow almost immediately from the lemma; see the end of this section for the details.

Lemma 4.4.2. *Let G be an unconflicted graph of maximum degree Δ satisfying Condition 1, and u be any unpinned vertex in G. Then, the following are true (with $\varepsilon_w, \varepsilon_I$, and ε_I as defined in eq. (4.9)):*

1. *For $i \in \Gamma_u$, $\left|Z^{(i)}_{G,u}(w)\right| > 0$.*
2. *For $i, j \in \Gamma_u$, if u has all neighbors pinned, then $R^{(i,j)}_{G,u}(w) = R^{(i,j)}_{G,u}(0) = 1$.*
3. *For $i, j \in \Gamma_u$, if u has $d \geq 1$ unpinned neighbors, then*
$$\frac{1}{d}\left|\Re \ln R^{(i,j)}_{G,u}(w) - \Re \ln R^{(i,j)}_{G,u}(0)\right| < \varepsilon_R.$$
4. *For any $i, j \in \Gamma_u$, if u has $d \geq 1$ unpinned neighbors, we have $\frac{1}{d}\left|\Im \ln R^{(i,j)}_{G,u}(w)\right| < \varepsilon_I$.*
5. *For any $i \notin \Gamma_u, j \in \Gamma_u$, then $\left|R^{(i,j)}_{G,u}(w)\right| \leq \varepsilon_w$.*

We will refer to items 1 to 5 as "items of the induction hypothesis". The rest of this section is devoted to the proof of this lemma via induction on the number of unpinned vertices in G.

We begin by verifying that the induction hypothesis holds in the base case when u is the only unpinned vertex in an unconflicted graph G. In this case, items 3 and 4 are vacuously true since u has no unpinned neighbors. Since all neighbors of u in G are pinned, the fact that all pinned vertices have degree at most one implies that G can be decomposed into two disjoint components G_1 and G_2, where G_1 consists of u and its pinned neighbors, while G_2 consists of a disjoint union of unconflicted edges (since G is unconflicted). Now, since G_1 and G_2 are disjoint components, we have $Z^{(i)}_{G,u}(w) = Z_{G_2}(w) = 1$ for all $i \in \Gamma_{G,u}$ and all $w \in \mathbb{C}$. This proves items 1 and 2. Similarly, when $i \notin \Gamma_{G,u}$, we have $Z^{(i)}_{G,u}(w) = w^{n_i}$, where $n_i \geq 1$ is the number of neighbors of u pinned to color i. This gives

$$\left|R^{(i,j)}_{G,u}(w)\right| \leq |w|^{n_i} \leq \varepsilon_w,$$

since $|w| \leq \varepsilon_w \leq 1$, and proves item 5.

We now derive some consequences of the above induction hypothesis that will be helpful in carrying out the induction. Throughout, we assume that G is an unconflicted graph satisfying Condition 1.

Consequence 4.4.3. *If $|L(u)| \geq \deg_G(u) + 1$ then*

$$|Z_G(w)| \geq 0.9 \min_{i \in \Gamma_u} \left|Z^{(i)}_{G,v}(w)\right| > 0.$$

Proof. Note that $Z_G(w) = \sum_{i \in L(u)} Z^{(i)}_{G,u}(w)$. From item 4, we see that the angle between the complex numbers $Z^{(i)}_{G,u}(w)$ and $Z^{(j)}_{G,u}(w)$, when $i, j \in \Gamma_u$, is at most $d\varepsilon_I$. Applying Lemma 4.2.7 to the terms corresponding to the good colors and item 5 to the terms corresponding to the bad colors, we then have

$$\begin{aligned}
\left|\sum_{i \in L(u)} Z^{(i)}_{G,u}(w)\right| &\geq \left(|\Gamma_u| \cos\frac{d\varepsilon_I}{2} - |L(u) \setminus \Gamma_u|\, \varepsilon_w\right) \min_{i \in \Gamma_u} \left|Z^{(i)}_{G,u}(w)\right| \\
&\geq \left((|L(u)| - \deg_G(u)) \cos\frac{d\varepsilon_I}{2} - |\deg_G(u)|\, \varepsilon_w\right) \min_{i \in \Gamma_u} \left|Z^{(i)}_{G,u}(w)\right| \\
&\geq \left(\cos\frac{d\varepsilon_I}{2} - |\deg_G(u)|\, \varepsilon_w\right) \min_{i \in \Gamma_u} \left|Z^{(i)}_{G,u}(w)\right|,
\end{aligned}$$

where we use the fact that $|L(u) \setminus \Gamma_u| \leq \deg_G(u)$ in the second inequality, and $|L(u)| \geq \deg_G(u) + 1$ in the last inequality. Since $d\varepsilon_I \leq 0.01$ and $\varepsilon_w \leq 0.01/\Delta$, we then have $\left|\sum_{i \in L(u)} Z^{(i)}_{G,u}(w)\right| \geq 0.9 \min_{i \in \Gamma_u} \left|Z^{(i)}_{G,v}(w)\right|$, which in turn is positive from item 1. ☐

Consequence 4.4.4. *The pseudo-probabilities approximate the real probabilities in the following sense:*

1. *for any $i \notin \Gamma_u$, $|\mathcal{P}_{G,w}[c(u) = i]| \leq 1.2\varepsilon_w$.*

2. *for any $j \in \Gamma_u$,*

$$\left|\Im \ln \frac{\mathcal{P}_{G,w}[c(u) = j]}{\mathcal{P}_G[c(u) = j]}\right| = |\Im \ln \mathcal{P}_{G,w}[c(u) = j]| \leq d\varepsilon_I + 2\Delta\varepsilon_w, \text{ and}$$

$$\left|\Re \ln \frac{\mathcal{P}_{G,w}[c(u) = j]}{\mathcal{P}_G[c(u) = j]}\right| \leq d\varepsilon_R + d\varepsilon_I + 2\Delta\varepsilon_w,$$

where d is the number of unpinned neighbors of u in G.

Proof. For part (1), by Consequence 4.4.3 we have

$$|\mathcal{P}_{G,w}[c(u)=i]| = \frac{\left|Z^{(i)}_{G,u}(w)\right|}{|Z_G(w)|} \le \frac{\left|Z^{(i)}_{G,u}(w)\right|}{0.9\min_{j\in\Gamma_u}\left|Z^{(j)}_{G,u}(w)\right|} \le 1.2\varepsilon_w,$$

where the last inequality follows from induction hypothesis item 5.

For part (2), by items 2 to 4 of the induction hypothesis, there exist complex numbers ξ_i (for all $i \in \Gamma_u$) satisfying $|\Re\xi_i| \le d\varepsilon_R$ and $|\Im\xi_i| \le d\varepsilon_I$ such that

$$\begin{aligned}\frac{1}{\mathcal{P}_{G,w}[c(u)=j]} &= \sum_{i\in L(u)} \frac{Z^{(i)}_{G,u}(w)}{Z^{(j)}_{G,u}(w)} \\ &= \underbrace{\sum_{i\in\Gamma_u} \frac{Z^{(i)}_{G,u}(0)}{Z^{(j)}_{G,u}(0)} e^{\xi_i}}_{:=A} + \underbrace{\sum_{i\in L(u)\setminus\Gamma_u} \frac{Z^{(i)}_{G,u}(w)}{Z^{(j)}_{G,u}(w)}}_{:=B}.\end{aligned}$$

Next we show that $A \approx \frac{1}{\mathcal{P}_G[c(u)=j]}$ and B is negligible. From item 5 of the induction hypothesis we have

$$\mathcal{P}_G[c(u)=j]\cdot|B| \le \Delta\varepsilon_w. \tag{4.10}$$

Now, note that $\sum_{i\in\Gamma_u}\frac{Z^{(i)}_{G,u}(0)}{Z^{(j)}_{G,u}(0)} = \frac{1}{\mathcal{P}_G[c(u)=j]}$. Further, when $\varepsilon_I \le 0.1/\Delta$, we also have[5]

$$\Re e^{\xi_i} \in (\exp(-d\varepsilon_R) - d^2\varepsilon_I^2, \exp(d\varepsilon_R)), \text{ and } |\arg e^{\xi_i}| \le d\varepsilon_I. \tag{4.11}$$

The above will therefore be true also for any convex combination of the e^{ξ_i}. Noting that $\mathcal{P}_G[c(u)=j]\cdot A$ is just such a convex combination (as the coefficients of the e^{ξ_i} are non-negative reals summing to 1), we have

$$\mathcal{P}_G[c(u)=j]\cdot\Re A \in (\exp(-d\varepsilon_R) - d^2\varepsilon_I^2, \exp(d\varepsilon_R)), \text{ and} \tag{4.12}$$

$$|\arg(\mathcal{P}_G[c(u)=j]\cdot A)| \le d\varepsilon_I. \tag{4.13}$$

Together, eqs. (4.10), (4.12) and (4.13) imply that if $C := \frac{\mathcal{P}_G[c(u)=j]}{\mathcal{P}_{G,w}[c(u)=j]}$ then (using the values

[5]Here, we also use the elementary facts that if z is a complex number satisfying $\Re z = r$ and $|\Im z| = \theta \le 0.1$ then $|\arg e^z| = |\Im z| = \theta$, and $e^r \ge \Re e^z = e^r\cos\theta = \exp(r + \ln\cos\theta) \ge \exp(r-\theta^2) \ge e^r - e^r\theta^2$. Hence if $r < 0$, we have $\Re e^z \ge e^r - \theta^2$.

of $\varepsilon_R, \varepsilon_I$, and ε_w)[6]

$$\begin{aligned}\Re C &\in \left(\exp(-d\varepsilon_R) - d^2\varepsilon_I^2 - \Delta\varepsilon_w, \exp(d\varepsilon_R) + \Delta\varepsilon_w\right), \text{ and}\\ \arg C &\in \left(-d\varepsilon_I - 2\Delta\varepsilon_w, d\varepsilon_I + 2\Delta\varepsilon_w\right).\end{aligned}$$

Thus, since $\varepsilon_I, \varepsilon_R$ are small enough and $\varepsilon_w \leq 0.01 \min\{\varepsilon_I, \varepsilon_R\}$, we have

$$\begin{aligned}|\Re \ln C| &\leq d\varepsilon_R + d\varepsilon_I + 2\Delta\varepsilon_w, \text{ and}\\ |\Im \ln C| &\leq d\varepsilon_I + 2\Delta\varepsilon_w.\end{aligned}$$

Here we use the elementary fact that for $z \in \mathbb{C}$, $\Re \ln z = \ln |z|$ and $\Im \ln z = \arg z$. Further, for z satisfying $\Re z = r \in [0.9, 1.1]$ and $|\arg z| = \theta \leq 0.1$, we also have $\ln r \leq \Re \ln z \leq \ln r + \ln \sec \theta \leq \ln r + \theta^2$. □

In the next consequence, we show that the error contracts during the induction. We first set up some notation. For a graph G, a vertex u, and a color $i \in \Gamma_u$, we let $a^{(i)}_{G,u}(w) = \ln \mathcal{P}_{G,w}[c(u) = i]$. We also recall that $\gamma := 1 - w$, and the definition of the function $f_\gamma(x) := -\ln(1 - \gamma e^x)$ from eq. (4.3).

Consequence 4.4.5. *There exists a positive constant $\eta \in [0.9, 1)$ so that the following is true. Let d be the number of unpinned neighbors of u. Assume further that u is nice in G. Then, for any colors $i, j \in \Gamma_u$, there exists a real constant $c = c_{G,u,i} \in [0, \frac{1}{d+\eta}]$ such that*

$$\left|\Re f_\gamma(a^{(i)}_{G,u}(w)) - f_1(a^{(i)}_{G,u}(0)) - c \cdot \Re\left(a^{(i)}_{G,u}(w) - a^{(i)}_{G,u}(0)\right)\right| \leq \varepsilon_I + \varepsilon_w. \tag{4.14}$$

$$\left|\Im f_\gamma(a^{(i)}_{G,u}(w)) - f_\gamma(a^{(j)}_{G,u}(w))\right| \leq \frac{1}{d+\eta} \cdot (d\varepsilon_I + 4\Delta\varepsilon_w) + 2\varepsilon_w. \tag{4.15}$$

$$\left|\Im f_\gamma(a^{(i)}_{G,u}(w))\right| \leq \frac{1}{d+\eta} \cdot (d\varepsilon_I + 4\Delta\varepsilon_w) + \varepsilon_w. \tag{4.16}$$

Proof. Since u is nice in G, the bound $\mathcal{P}_{G,0}[c(u) = k] \leq \frac{1}{d+2}$ (for any $k \in \Gamma_{G,u}$) applies. Combining them with Consequence 4.4.4 we see that $a^{(i)}_{G,u}(w), a^{(i)}_{G,u}(0), a^{(j)}_{G,u}(w), a^{(j)}_{G,u}(0)$ lie in a domain D as described in Lemma 4.2.6 (with the parameter κ therein set to 1), with

[6] Here, for the second inclusion, we use the following elementary computation. Let z, s be complex numbers such that $\Re z = r \in [0.9, 1.1]$, $|\arg z| = \theta \leq 0.1$ and $|s| \leq 0.1$. Then, we have $\Re(z + s) \geq r - |s|$ and $|\Im(z + s)| \leq r\theta + |s|$. Thus, $|\arg(z + s)| \leq \frac{|\Im(z+s)|}{|\Re(z+s)|} \leq \frac{r\theta+|s|}{r-|s|} = \theta + |s| \cdot \frac{1+\theta}{r-|s|} \leq \theta + 2\,|s|$.

the parameters ζ and τ in that observation chosen as

$$\begin{aligned} \zeta &= \ln(d+2) - d\varepsilon_R - d\varepsilon_I - 2\Delta\varepsilon_w \text{ and} \\ \tau &= d\varepsilon_I + 2\Delta\varepsilon_w. \end{aligned} \tag{4.17}$$

Here, for the bound on ζ, we use the fact that for $j \in \Gamma_{G,u}$, $\mathcal{P}_G[c(u) = j] \leq \frac{1}{d+2}$, which is due to u being nice in G.

The bounds on $\varepsilon_w, \varepsilon_I$ and ε_R now imply $e^\zeta \geq (d+2)\left(1 - \frac{0.02}{\Delta}\right) \geq d + 1.94$, and also that $\tau \leq 0.02/\Delta$. Thus, the conditions required on ζ and τ in Lemma 4.2.6 (i.e. that $\tau < 1/2$ and $\tau^2 + e^{-\zeta} < 1$) are satisfied. Further, ρ_R and ρ_I as set in the observation satisfy $\rho_R \leq \frac{1}{d+\eta}$, where η can be taken to be 0.94, and $\rho_I < 3\varepsilon_I$.

Using Lemma 4.2.5 followed by the value of ε_w, and noting that $a^{(i)}_{G,u}(0)$ is a real number, we then have

$$\begin{aligned} \left|\Re f_1(a^{(i)}_{G,u}(w)) - f_1(a^{(i)}_{G,u}(0)) - c \cdot \Re\left(a^{(i)}_{G,u}(w) - a^{(i)}_{G,u}(0)\right)\right| &\leq \rho_I \cdot \left|\Im\left(a^{(i)}_{G,u}(w) - a^{(i)}_{G,u}(0)\right)\right| \\ &\leq 3\varepsilon_I(d\varepsilon_I + 2\Delta\varepsilon_w) \leq 4d\varepsilon_I^2 \leq \varepsilon_I, \end{aligned} \tag{4.18}$$

for an appropriate positive $c \leq 1/(d+\eta)$. This is almost eq. (4.14), whose difference will be handled later.

Similarly, applying Lemma 4.2.5 to the imaginary part we have

$$\begin{aligned} &\left|\Im f_1(a^{(i)}_{G,u}(w)) - f_1(a^{(j)}_{G,u}(w))\right| \\ &\qquad \leq \rho_R \cdot \max\left\{\left|\Im\left(a^{(i)}_{G,u}(w) - a^{(j)}_{G,u}(w)\right)\right|, \left|\Im a^{(i)}_{G,u}(w)\right|, \left|\Im a^{(j)}_{G,u}(w)\right|\right\}, \end{aligned} \tag{4.19}$$

where, as noted above, $\rho_R \leq \frac{1}{d+\eta}$. Now, note that the first term in the above maximum is less than $d\varepsilon_I$ by item 4 of the induction hypothesis, while the other two terms are at most $d\varepsilon_I + 2\Delta\varepsilon_w$ from item 2 of Consequence 4.4.4. This is almost the bound in eq. (4.15), whose difference will be handled later.

To prove the bound in eq. (4.16), we first apply the imaginary part of Lemma 4.2.5 along with the fact that $\Im a^{(i)}_{G,u}(0) = 0$ to get

$$\left|\Im f_1(a^{(i)}_{G,u}(w))\right| = \left|\Im f_1(a^{(i)}_{G,u}(w)) - f_1(a^{(i)}_{G,u}(0))\right| \leq \rho_R \cdot \left|\Im\left(a^{(i)}_{G,u}(w)\right)\right| \leq \frac{1}{d+\eta}(d\varepsilon_I + \Delta\varepsilon_w). \tag{4.20}$$

Finally, we use item 2 of Lemma 4.2.6 (with the parameter κ' therein set to γ) to conclude the proofs of eqs. (4.14) to (4.16). To this end, we note that γ satisfies $|\gamma - 1| \leq \varepsilon_w$, so that the condition $(1 + \varepsilon_w) < e^{\zeta}$ required for item 2 to apply is satisfied. Thus we see that for any $z \in D$,

$$|f_\gamma(z) - f_1(z)| \le \varepsilon_w,$$

so that the quantities $|\Re f_\gamma(a^{(i)}_{G,u}(w)) - \Re f_1(a^{(i)}_{G,u}(w))|$, $|\Im f_\gamma(a^{(i)}_{G,u}(w)) - \Im f_1(a^{(i)}_{G,u}(w))|$, $|\Im f_\gamma(a^{(j)}_{G,u}(w)) - \Im f_1(a^{(j)}_{G,u}(w))|$, and $|\Im f_\gamma(a^{(j)}_{G,u}(w)) - \Im f_1(a^{(j)}_{G,u}(w))|$ are all at most ε_w. The desired bounds of eqs. (4.14) to (4.16) now follow from the triangle inequality and the bounds in eqs. (4.18) to (4.20). $\square$

We set up some further notation for the next consequence. For a color $i \in L(u) \setminus \Gamma_u$ we let $b^{(i)}_{G,u}(w) = \mathcal{P}_{G,w}[c(u) = i]$. We then consider the function $g_\gamma(x) := -\ln(1 - \gamma x)$.

Consequence 4.4.6. *For every color $i \notin \Gamma_u$, $\left|g_\gamma(b^{(i)}_{G,u}(w))\right| \le 2\varepsilon_w$.*

Proof. Item 1 of Consequence 4.4.4 implies that $\left|b^{(i)}_{G,u}(w)\right| \le 1.2\varepsilon_w$. Thus, recalling that $|\gamma - 1| \le \varepsilon_w$, we get that for all $\varepsilon_w < 0.01$, $\left|g_\gamma(b^{(i)}_{G,u}(w))\right| = \left|\ln(1 - \gamma b^{(i)}_{G,u}(w))\right| \le 2\varepsilon_w$. $\square$

Inductive proof of Lemma 4.4.2

We are now ready to see the induction step in the proof of Lemma 4.4.2; recall that the base case was already established following the statement of the lemma. Let G be any unconflicted graph which satisfies Condition 1 and had at least two unpinned vertices (the base case when $|G| = 1$ was already handled above). We first prove induction item 1 for any vertex $u \in G$. Consider the graph G' obtained from G by pinning vertex u to color i. Note that by the definition of the pinning operation, $Z^{(i)}_{G,u}(w) = Z_{G'}(w)$, and when $i \in \Gamma_{G,u}$, the graph G' is also unconflicted and satisfies Condition 1, and has one fewer unpinned vertex than G. Thus, from Consequence 4.4.3 of the induction hypothesis applied to G', we have that $\left|Z^{(i)}_{G,u}(w)\right| = |Z_{G'}(w)| > 0$.

We now consider item 2. When all neighbors of u in G are pinned, the fact that all pinned vertices have degree at most one implies that G can be decomposed into two disjoint components G_1 and G_2, where G_1 consists of u and its pinned neighbors, while G_2 is also unconflicted (when G is unconflicted) and has one fewer unpinned vertex than G. Now, since G_1 and G_2 are disjoint components, we have $Z^{(k)}_{G,u}(x) = Z_{G_2}(x)$ for all $k \in \Gamma_{G,u}$ and all $x \in \mathbb{C}$. Further, from Consequence 4.4.3 of the induction hypothesis applied to G_2, we also have that $Z_{G_2}(w)$ and $Z_{G_2}(0)$ are both non-zero. It therefore follows that when $i, j \in \Gamma_{G,u}$, $R^{(i,j)}_{G,u}(w) = R^{(i,j)}_{G,u}(0) = 1$.

We now consider items 3 and 4. Recall that by Lemma 4.2.4, we have

$$R_{G,u}^{(i,j)}(w) = \prod_{k=1}^{\deg_G(u)} \frac{\left(1-\gamma\mathcal{P}_{G_k^{(i,j)},w}[c(v_k)=i]\right)}{\left(1-\gamma\mathcal{P}_{G_k^{(i,j)},w}[c(v_k)=j]\right)}. \tag{4.21}$$

For simplicity we write $G_k := G_k^{(i,j)}$. Note that when $i, j \in \Gamma_{G,u}$, and G is unconflicted, so are the G_k. Further, each G_k has exactly one fewer unpinned vertex than G, so that the induction hypothesis applies to each G_k. Note also that when $i, j \in \Gamma_{G,u}$, we can restrict the product above to the d unpinned neighbors of u, since for such i, j, the contribution of the factor corresponding to a pinned neighbor is 1, irrespective of the value of w. Without loss of generality, we relabel these unpinned neighbors as $v_1, v_2, \ldots, v_d$.

Now, as before, for $s \in \Gamma_{G_k,v_k}$ we define $a_{G_k,v_k}^{(s)}(w) := \ln \mathcal{P}_{G_k,w}[c(v_k)=s]$; while for $t \in L(v_k) \setminus \Gamma_{G_k,v_k}$ we let $b_{G_k,v_k}^{(t)}(w) := \mathcal{P}_{G_k,w}[c(v_k)=t]$. For a graph G, a vertex u and a color s, we let $B_{G,u}(s)$ be the set of those neighbors of u for which s is a bad color in $G \setminus \{u\}$. For simplicity we will also write $B(s) := B_{G,u}(s)$ when it is clear from the context. As before, we have $\gamma = 1-w$, $f_\gamma(x) = -\ln(1-\gamma e^x)$, $g_\gamma(x) = -\ln(1-\gamma x)$. From the above recurrence, we then have,

$$\begin{aligned}
-\ln R_{G,u}^{(i,j)}(w) = &\sum_{v_k \in \overline{B(i)} \cap \overline{B(j)}} f_\gamma\left(a_{G_k,v_k}^{(i)}(w)\right) - f_\gamma\left(a_{G_k,v_k}^{(j)}(w)\right) \\
&+ \left(\sum_{v_k \in \overline{B(i)} \cap B(j)} f_\gamma\left(a_{G_k,v_k}^{(i)}(w)\right)\right) - \left(\sum_{v_k \in B(i) \cap \overline{B(j)}} f_\gamma\left(a_{G_k,v_k}^{(j)}(w)\right)\right) \\
&- \left(\sum_{v_k \in \overline{B(i)} \cap B(j)} g_\gamma\left(b_{G_k,v_k}^{(j)}(w)\right)\right) + \left(\sum_{v_k \in B(i) \cap \overline{B(j)}} g_\gamma\left(b_{G_k,v_k}^{(i)}(w)\right)\right) \\
&+ \left(\sum_{v_k \in B(i) \cap B(j)} g_\gamma\left(b_{G_k,v_k}^{(i)}(w)\right) - g_\gamma\left(b_{G_k,v_k}^{(j)}(w)\right)\right).
\end{aligned} \tag{4.22}$$

Note that the same recurrence also applies when w is replaced by 0 (and hence γ by 1), except in that case the last three sums are 0 (as, when i is bad for v_k in G_k, we have $b_{G_k,v_k}^{(i)}(0) := \Pr_{G_k}[c(v_k) = i] = 0$):

$$\begin{aligned}
-\ln R_{G,u}^{(i,j)}(0) = &\sum_{v_k \in \overline{B(i)} \cap \overline{B(j)}} f_1\left(a_{G_k,v_k}^{(i)}(0)\right) - f_1\left(a_{G_k,v_k}^{(j)}(0)\right) \\
&+ \left(\sum_{v_k \in \overline{B(i)} \cap B(j)} f_1\left(a_{G_k,v_k}^{(i)}(0)\right)\right) - \left(\sum_{v_k \in B(i) \cap \overline{B(j)}} f_1\left(a_{G_k,v_k}^{(j)}(0)\right)\right).
\end{aligned} \tag{4.23}$$

Further, by Consequence 4.4.6 of the induction hypothesis applied to the graph G_k at a vertex $v_k \in B(i)$ (respectively, $v_k \in B(j)$) we see that $\left|g_\gamma\left(b^{(i)}_{G_k,v_k}(w)\right)\right| \leq 2\varepsilon_w$ (respectively, $g_\gamma\left(b^{(j)}_{G_k,v_k}(w)\right) \leq 2\varepsilon_w$). Thus, applying the triangle inequality to the real part of the difference of the two recurrences, we get

$$\begin{aligned}
\frac{1}{d}\left|\Re \ln R^{(i,j)}_{G,u}(0) - \ln R^{(i,j)}_{G,u}(w)\right| \leq 2\Delta\varepsilon_w & \\
+ \max\Bigg\{ \max_{v_k \in \overline{B(i)} \cap \overline{B(j)}} & \Bigg\{\Bigg|\left(\Re f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - f_1\left(a^{(i)}_{G_k,v_k}(0)\right)\right) \\
& - \left(\Re f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right) - f_1\left(a^{(j)}_{G_k,v_k}(0)\right)\right)\Bigg|\Bigg\}, \\
\max_{v_k \in \overline{B(i)} \cap B(j)} & \left\{\left|\Re f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - f_1\left(a^{(i)}_{G_k,v_k}(0)\right)\right|\right\}, \\
\max_{v_k \in \overline{B(j)} \cap B(i)} & \left\{\left|\Re f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right) - f_1\left(a^{(j)}_{G_k,v_k}(0)\right)\right|\right\}\Bigg\}.
\end{aligned} \tag{4.24}$$

In what follows, we let v_k be the vertex that maximizes the above expression, and d_k be the number of unpinned neighbors of v_k in G_k. Before proceeding with the analysis, we note that the graphs G_k are unconflicted and satisfy Condition 1, and further that v_k is nice in G_k (this last fact follows from Lemma 4.3.2 and the fact that G satisfies Condition 1). Thus, the preconditions of Consequence 4.4.5 apply to the vertex v_k in graph G_k. We now proceed with the analysis.

We first consider $v_k \in \overline{B(i)} \cap B(j)$. Note that this implies that $i \in \Gamma_{G_k,v_k}$. Thus, the conditions of Consequence 4.4.5 of the induction hypothesis instantiated on G_k apply to v_k with color i, and we thus have from eq. (4.14) that

$$\left|\Re f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - f_1\left(a^{(i)}_{G_k,v_k}(0)\right)\right| \leq \frac{1}{d_k+\eta}\left|\Re a^{(i)}_{G_k,v_k}(w) - a^{(i)}_{G_k,v_k}(0)\right| + \varepsilon_I + \varepsilon_w,$$

where d_k is the number of unpinned neighbors of v_k and $\eta \in [0.9, 1)$ is as in the statement of Consequence 4.4.5. Applying item 2 of Consequence 4.4.4 (which, again, is applicable because $i \in \Gamma_{G_k,v_k}$), we then have $\left|\Re a^{(i)}_{G_k,v_k}(w) - a^{(i)}_{G_k,v_k}(0)\right| \leq d_k(\varepsilon_R + \varepsilon_I) + 2\Delta\varepsilon_w$, so that

$$\left|\Re f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - f_1\left(a^{(i)}_{G_k,v_k}(0)\right)\right| \leq \frac{d_k}{d_k+\eta}\varepsilon_R + 2\varepsilon_I + 3\Delta\varepsilon_w. \tag{4.25}$$

By interchanging the roles of i and j in the above argument, we see that, for $v_k \in \overline{B(j)} \cap B(i)$

$$\left|\Re f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right) - f_1\left(a^{(j)}_{G_k,v_k}(0)\right)\right| \leq \frac{d_k}{d_k+\eta}\varepsilon_R + 2\varepsilon_I + 3\Delta\varepsilon_w. \tag{4.26}$$

We now consider $v_k \in \overline{B(i)} \cap \overline{B(j)}$. Note that both i and j are good for v_k in G_k, so that

$$\begin{aligned}
&\left|\left(\Re f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - f_1\left(a^{(i)}_{G_k,v_k}(0)\right)\right) - \left(\Re f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right) - f_1\left(a^{(j)}_{G_k,v_k}(0)\right)\right)\right| \\
&\le \max_{i',j' \in \Gamma_{G_k,v_k}} \left|\left(\Re f_\gamma\left(a^{(i')}_{G_k,v_k}(w)\right) - f_1\left(a^{(i')}_{G_k,v_k}(0)\right)\right) - \left(\Re f_\gamma\left(a^{(j')}_{G_k,v_k}(w)\right) - f_1\left(a^{(j')}_{G_k,v_k}(0)\right)\right)\right|,
\end{aligned}$$

Now, for any color $s \in \Gamma_{G_k,v_k}$, Consequence 4.4.5 of the induction hypothesis instantiated on G_k and applied to v_k and s shows that there exists a $C_s = C_{s,v_k,G_k} \in [0, 1/(d_k+\eta)]$ such that

$$\left|\Re f_\gamma\left(a^{(s)}_{G_k,v_k}(w)\right) - f_1\left(a^{(s)}_{G_k,v_k}(0)\right) - C_s(\Re a^{(s)}_{G_k,v_k}(w) - a^{(s)}_{G_k,v_k}(0))\right| \le \varepsilon_I + \varepsilon_w.$$

Substituting this in the previous display shows that

$$\begin{aligned}
&\left|\left(\Re f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - f_1\left(a^{(i)}_{G_k,v_k}(0)\right)\right) - \left(\Re f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right) - f_1\left(a^{(j)}_{G_k,v_k}(0)\right)\right)\right| \\
&\quad \le \max_{i',j' \in \Gamma_{G_k,v_k}} \left|C_{i'}(\Re a^{(i')}_{G_k,v_k}(w) - a^{(i')}_{G_k,v_k}(0)) - C_{j'}(\Re a^{(j')}_{G_k,v_k}(w) - a^{(j')}_{G_k,v_k}(0))\right| + 2\varepsilon_I + 2\varepsilon_w \\
&\quad = 2\varepsilon_I + 2\varepsilon_w + \max_{i',j' \in \Gamma_{G_k,v_k}} \left|C_{i'}\Re\xi_{i'} - C_{j'}\Re\xi_{j'}\right|, \\
&\quad = 2\varepsilon_I + 2\varepsilon_w + C_s\Re\xi_s - C_t\Re\xi_t, \qquad (4.27)
\end{aligned}$$

where $\xi_l := a^{(l)}_{G_k,v_k}(w) - a^{(l)}_{G_k,v_k}(0)$ for $l \in \Gamma_{G_k,v_k}$, and s and t are given by

$$s := \operatorname*{arg\,max}_{i' \in \Gamma_{G_k,v_k}} C_{i'}\Re\xi_{i'} \quad \text{and} \quad t := \operatorname*{arg\,min}_{i' \in \Gamma_{G_k,v_k}} C_{i'}\Re\xi_{i'}.$$

We now have the following two cases:

Case 1: $(\Re\xi_s)\cdot(\Re\xi_t) \le 0$. Recall that C_s, C_t are non-negative and lie in $[0, 1/(d_k+\eta)]$. Thus, in this case, we must have $\Re\xi_s \ge 0$ and $\Re\xi_t \le 0$, so that

$$C_s\Re\xi_s - C_t\Re\xi_t = C_s\Re\xi_s + C_t\left|\Re\xi_t\right| \le \frac{1}{d_k+\eta}(\Re\xi_s + \left|\Re\xi_t\right|) = \frac{1}{d_k+\eta}\left|\Re\xi_s - \Re\xi_t\right|. \qquad (4.28)$$

Now, note that

$$\begin{aligned}
\Re\xi_s - \Re\xi_t &= \Re\ln\frac{\mathcal{P}_{G_k,w}[c(v_k)=s]}{\mathcal{P}_{G_k}[c(v_k)=s]} - \Re\ln\frac{\mathcal{P}_{G_k,w}[c(v_k)=t]}{\mathcal{P}_{G_k}[c(v_k)=t]} \\
&= \Re\ln\frac{\mathcal{P}_{G_k,w}[c(v_k)=s]}{\mathcal{P}_{G_k,w}[c(v_k)=t]} - \Re\ln\frac{\mathcal{P}_{G_k}[c(v_k)=s]}{\mathcal{P}_{G_k}[c(v_k)=t]} \\
&= \Re\ln R^{(s,t)}_{G_k,v_k}(w) - \ln R^{(s,t)}_{G_k,v_k}(0).
\end{aligned}$$

Note that all the logarithms in the above are well defined from Consequence 4.4.4 of the induction hypothesis applied to G_k and v_k (as $s, t \in \Gamma_{G_k,v_k}$). Further from items 2 and 3 of the induction hypothesis, the last term is at most $d_k \varepsilon_R$ in absolute value. Substituting this in eq. (4.28), we get

$$C_s \Re \xi_s - C_t \Re \xi_t \leq \frac{d_k}{d_k + \eta} \varepsilon_R. \tag{4.29}$$

This concludes the analysis of Case 1.

Case 2: $\Re \xi_{i'}$ for $i' \in \Gamma_{G_k,v_k}$ all have the same sign. Suppose first that $\Re \xi_{i'} \geq 0$ for all $i' \in \Gamma_{G_k,v_k}$. Then, we have

$$0 \leq C_s \Re \xi_s - C_t \Re \xi_t \leq \frac{1}{d_k + \eta} \Re \xi_s \leq \frac{d_k}{d_k + \eta} \varepsilon_R + \varepsilon_I + 4\Delta \varepsilon_w, \tag{4.30}$$

where the last inequality follows from item 2 of Consequence 4.4.5 of the induction hypothesis applied to G_k at vertex v_k with color s, which states that $|\Re \xi_s| \leq d_k(\varepsilon_R + \varepsilon_I) + 4\Delta \varepsilon_w$. Similarly, when $\Re \xi_{i'} \leq 0$ for all $i' \in \Gamma_{G_k,v_k}$, we have

$$\begin{aligned} 0 \leq C_s \Re \xi_s - C_t \Re \xi_t &= C_t |\Re \xi_t| - C_s |\Re \xi_s| \\ &\leq \frac{1}{d_k + \eta} |\Re \xi_t| \leq \frac{d_k}{d_k + \eta} \varepsilon_R + \varepsilon_I + 4\Delta \varepsilon_w, \end{aligned} \tag{4.31}$$

where the last inequality follows from item 2 of Consequence 4.4.5 of the induction hypothesis applied to G_k at vertex v_k with color t, which states that $|\Re \xi_t| \leq d_k(\varepsilon_R + \varepsilon_I) + 4\Delta \varepsilon_w$. This concludes the analysis of Case 2.

Now, substituting eqs. (4.29) to (4.31) into eq. (4.27), we get

$$\begin{aligned} &\left| \left(\Re f_\gamma \left(a^{(i)}_{G_k,v_k}(w) \right) - f_1 \left(a^{(i)}_{G_k,v_k}(0) \right) \right) - \left(\Re f_\gamma \left(a^{(j)}_{G_k,v_k}(w) \right) - f_1 \left(a^{(j)}_{G_k,v_k}(0) \right) \right) \right| \\ &\qquad \leq \frac{d_k}{d_k + \eta} \varepsilon_R + 3\varepsilon_I + 5\Delta \varepsilon_w. \end{aligned} \tag{4.32}$$

Substituting eqs. (4.25), (4.26) and (4.32) into eq. (4.24), we get

$$\frac{1}{d} \left| \Re \ln R^{(i,j)}_{G,u}(w) - \ln R^{(i,j)}_{G,u}(0) \right| \leq \frac{d_k}{d_k + \eta} \varepsilon_R + 3\varepsilon_I + 7\Delta \varepsilon_w < \varepsilon_R,$$

where the last inequality follows since $\eta \varepsilon_R > (\Delta + 1)(3\varepsilon_I + 7\Delta \varepsilon_w)$ (recalling that $0 \leq d_k \leq \Delta$ and $\eta \in [0.9, 1)$). This verifies item 3 of the induction hypothesis.

For item 4, we consider the imaginary part of eq. (4.22). As in the derivation of eq. (4.24), we use the fact that the induction hypothesis applied to the graph G_k at the vertex $v_k \in B(i)$ (respectively, $v_k \in B(j)$) implies that $\left|g_\gamma\left(b^{(i)}_{G_k,v_k}(w)\right)\right| \le 2\varepsilon_w$ (respectively, $g_\gamma\left(b^{(j)}_{G_k,v_k}(w)\right) \le 2\varepsilon_w$). This yields

$$\begin{aligned}\frac{1}{d}\left|\Im \ln R^{(i,j)}_{G,u}(w)\right| &\le 2\Delta\varepsilon_w \\ &+ \max\left\{\max_{v_k \in \overline{B(i)\cap B(j)}}\left|\Im f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - \Im f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right)\right|\right. \\ &\left.\max_{v_k \in \overline{B(i)}\cap B(j)}\left|\Im f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right)\right|, \max_{v_k \in \overline{B(j)}\cap B(i)}\left|\Im f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right)\right|\right\}.\end{aligned} \tag{4.33}$$

Again, let v_k be the vertex that maximizes the above expression, and d_k be the number of unpinned neighbors of v_k in G_k. We first consider $v_k \in \overline{B(i)} \cap \overline{B(j)}$. Applying eq. (4.15) of Consequence 4.4.5 of the induction hypothesis to the graph G_k at vertex v_k with colors $i, j \in \Gamma_{G_k,v_k}$ gives

$$\left|\Im f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - \Im f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right)\right| \le \frac{d_k}{d_k+\eta}\varepsilon_I + 6\Delta\varepsilon_w. \tag{4.34}$$

Now consider $v_k \in \overline{B(i)} \cap B(j)$. For this case, eq. (4.16) of Consequence 4.4.5 of the induction hypothesis applied to G_k at vertex v_k with color $i \in \Gamma_{G_k,v_k}$ gives

$$\left|\Im f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right)\right| \le \frac{d_k}{d_k+\eta}\varepsilon_I + 5\Delta\varepsilon_w. \tag{4.35}$$

Similarly, for $v_k \in \overline{B(j)} \cap B(i)$. For this case, eq. (4.16) of Consequence 4.4.5 of the induction hypothesis applied to G_k at vertex v_k with color $j \in \Gamma_{G_k,v_k}$ gives

$$\left|\Im f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right)\right| \le \frac{d_k}{d_k+\eta}\varepsilon_I + 5\Delta\varepsilon_w. \tag{4.36}$$

Substituting eqs. (4.34) to (4.36) into eq. (4.33) we then have

$$\frac{1}{d}\left|\Im \ln R^{(i,j)}_{G,u}(w)\right| \le \frac{d_k}{d_k+\eta}\varepsilon_I + 8\Delta\varepsilon_w < \varepsilon_I,$$

where the last inequality holds since $\eta\varepsilon_I > 8(\Delta+1)\Delta\varepsilon_w$ (recalling that $0 \le d_k \le \Delta$ and $\eta \in [0.9, 1)$). This completes the proof of item item 4 of the induction hypothesis.

Finally, we prove item 5. Since $i \notin \Gamma_u$, there exist $n_i > 0$ neighbors of u that are pinned to color i. Let H be the graph obtained by removing these neighbors of u from G. Then, H is an unconflicted graph with the *same* number of unpinned vertices as G which also satisfies $i, j \in \Gamma_{H,u}$; we can therefore apply the already proved items 1 to 3 to H to conclude that

$$\left|R_H^{(i,j)}(w)\right| \le \left|R_H^{(i,j)}(0)\right| \exp(d\varepsilon_R). \tag{4.37}$$

Now, since $i, j \in \Gamma_{H,u}$, we can apply the recurrence of Lemma 4.2.4 in the same way as in the derivation of eq. (4.21) above to get

$$R_{H,u}^{(i,j)}(w) = \prod_{k=1}^{\deg_H(u)} \frac{\left(1 - \mathcal{P}_{H_k^{(i,j)},w}[c(v_k) = i]\right)}{\left(1 - \mathcal{P}_{H_k^{(i,j)},w}[c(v_k) = j]\right)}, \tag{4.38}$$

where, for the reasons described in the discussion following eq. (4.21), the product can be restricted to unpinned neighbors of u in H. Renaming these unpinned neighbors as $v_1, v_2, \ldots, v_d$, we then have

$$0 \le R_H^{(i,j)}(0) = \prod_{k=1}^{d} \frac{(1 - \mathcal{P}_{H_k}[c(v_k) = i]\,)}{(1 - \mathcal{P}_{H_k}[c(v_k) = j]\,)}, \tag{4.39}$$

where as before, $H_k := H_k^{(i,j)}$. Now, since G satisfies Condition 1, so does H. Thus, for $1 \le k \le d$, v_k is nice in H_k (Lemma 4.3.2), and hence, $\mathcal{P}_{H_k}[c(v_k) = j] \le \frac{1}{d_k+2}$ for $1 \le k \le d$, where $d_k \ge 0$ is the number of unpinned neighbors of v_k in H_k. We then have

$$0 \le R_H^{(i,j)}(0) = \prod_{k=1}^{d} \frac{(1 - \mathcal{P}_{H_k}[c(v_k) = i]\,)}{(1 - \mathcal{P}_{H_k}[c(v_k) = j]\,)} \le \prod_{k=1}^{d} \frac{1}{1 - \frac{1}{d_k+2}} = \prod_{k=1}^{d} \frac{d_k + 2}{d_k + 1} \le 2^{\Delta}.$$

(As an aside, we note that one could get a better bound under the slightly stronger assumption of uniformly large list sizes considered in Remark 12. Under the conditions of that remark, we have $\mathcal{P}_{H_k}[c(v_k) = j] < \min\left\{\frac{4}{3\Delta}, 1\right\}$, so that the above upper bound can be improved to $R_H^{(i,j)}(0) \le e^4$ for $\Delta > 1$.)

Combining the estimate with eq. (4.37), we get $\left|R_H^{(i,j)}(w)\right| \le 5 \cdot 2^{\Delta}$ since $d\varepsilon_R \le 1/2$. Now note that since $j \in \Gamma_{G,u}$,

$$Z_{G,u}^{(i)}(w) = w^{n_i} Z_{H,u}^{(i)}(w), \quad \text{and} \quad Z_{G,u}^{(j)}(w) = Z_{H,u}^{(j)}(w),$$

so that $\left|R_{G,u}^{(i,j)}(w)\right| = |w|^{n_i} \left|R_{H,u}^{(i,j)}(w)\right| \le 5 \cdot 2^{\Delta} \cdot |w|^{n_i}$. The latter is at most ε_w whenever $|w| \le 0.2\varepsilon_w/2^{\Delta}$. This proves item 5, and also completes the inductive proof of Lemma 4.4.2. (Note also that using the stronger upper bound above under the condition of uniformly large list sizes, we can in fact relax the requirement further to $|w| \le \varepsilon_w/(300\Delta)$.) □

We conclude this section by using Lemma 4.4.2 to prove Theorem 4.4.1.

Proof of Theorem 4.4.1. Let G be a graph satisfying Condition 1. Since G has no pinned vertices, G is unconflicted. Let u be an unpinned vertex in G. By Consequence 4.4.3 of the induction hypothesis (which we proved in Lemma 4.4.2), we then have $Z_w(G) \neq 0$ provided $\nu_w \leq 0.2\varepsilon_w/2^\Delta$.

Furthermore, as discussed above, under a slightly stronger assumption of uniformly large list sizes considered in Remark 12, ν_w can be chosen to be $\varepsilon_w/(300\Delta)$. □

4.5 Zero-free region around the interval $(0, 1]$

In this section, we consider the case of w close to $[0, 1]$ but bounded away from 0. In particular, we prove the following theorem, which complements Theorem 4.4.1.

Theorem 4.5.1. *Fix a positive integer Δ and let $\nu_w = \nu_w(\Delta)$ be as in Theorem 4.4.1. Then, for any w satisfying*

$$\Re w \in [\nu_w/2, 1 + \nu_w^2/8] \qquad \text{and} \qquad |\Im w| \leq \nu_w^2/8, \tag{4.40}$$

and any graph G satisfying Condition 1, we have $Z_G(w) \neq 0$.

(Here, we recall that as described in the discussion following Theorem 4.4.1, ν_w can be chosen to be $\varepsilon_w/(300\Delta)$ when the uniformly large list size condition of Remark 12 is satisfied. However, as in that theorem, in the case of general list coloring, one chooses $\nu_w = 0.2\varepsilon_w/2^\Delta$.)

For w as in eq. (4.40), we define $\tilde{w}$ to be the point on the interval $[0, 1]$ which is closest to w. Thus

$$\tilde{w} := \begin{cases} \Re w & \text{when } \Re w \in [\nu_w/2, 1]; \\ 1 & \text{when } \Re w \in (1, 1 + \nu_w^2/8]. \end{cases}$$

We also define, in analogy with the last section, $\gamma := 1 - w$ and $\tilde{\gamma} := 1 - \tilde{w}$. We record a few properties of these quantities in the following observation.

Observation 4.5.2. *With $w, \gamma, \tilde{w}$ and $\tilde{\gamma}$ as above, we have*

1. $0 \leq \tilde{\gamma}, |\gamma| < 1$.
2. $|\ln w - \ln \tilde{w}| \leq \nu_w$.

Proof. We have $\tilde{\gamma} \in [0, 1-\nu_w/2]$, $\Re\gamma \in [-\nu_w^2/8, 1-\nu_w/2]$ and $|\Im\gamma| \leq \nu_w^2/8$. Since $\nu_w \leq 0.01$, these bounds taken together imply item 1. We also have $0 \leq \tilde{w} \leq |w| \leq \tilde{w} + \nu_w^2/4$ and $\tilde{w} \geq \nu_w/2$. Thus

$$0 \leq \Re(\ln w - \ln \tilde{w}) = \ln \frac{|w|}{\tilde{w}} \leq \ln\left(1 + \frac{\nu_w^2}{4\tilde{w}}\right) \leq \frac{\nu_w}{2}.$$

Similarly, $\Im(\ln w - \ln \tilde{w}) = \Im \ln w = \arg w$, so that

$$|\Im(\ln w - \ln \tilde{w})| \leq |\arg w| \leq \frac{|\Im w|}{\Re w} \leq \frac{\nu_w}{4}.$$

Together, the above two bounds imply item 2. □

In analogous fashion to the proof of Theorem 4.4.1, we would like to show that $R_{G,u}^{(i,j)}(w) \approx R_{G,u}^{(i,j)}(\tilde{w})$ independent of the size of G. (Note that for positive $\tilde{w}$, $R_{G,u}^{(i,j)}(\tilde{w})$ is a well defined positive real number for any graph.) To this end, we will prove the following analog of Lemma 4.4.2 for any graph G satisfying Condition 1 and any vertex u in G, via an induction on the number of unpinned vertices in G. The induction is very similar in structure to that used in the proof of Lemma 4.4.2, except that the fact that w has strictly positive real part allows us to simplify several aspects of the proof. In particular, we do not need to consider good and bad colors separately, and do not require the underlying graphs to be unconflicted.

As in the previous section, we assume that all graphs in this section have maximum degree at most $\Delta \geq 1$, and define the quantities $\varepsilon_w, \varepsilon_R, \varepsilon_I$ in terms of Δ using eq. (4.9).

Lemma 4.5.3. *Let G be a graph of maximum degree Δ satisfying Condition 1 and let u be any unpinned vertex in G. Then, the following are true (here, $\varepsilon_w, \varepsilon_I, \varepsilon_R$ are as defined in eq. (4.9)):*

1. *For $i \in L(u)$, $\left|Z_{G,u}^{(i)}(w)\right| > 0$.*
2. *For $i, j \in L(u)$, if u has all neighbors pinned, then $|\ln R_{G,u}^{(i,j)}(w) - \ln R_{G,u}^{(i,j)}(\tilde{w})| < \varepsilon_w$.*
3. *For $i, j \in L(u)$, if u has $d \geq 1$ unpinned neighbors, then*

$$\frac{1}{d}\left|\Re \ln R_{G,u}^{(i,j)}(w) - \Re \ln R_{G,u}^{(i,j)}(\tilde{w})\right| < \varepsilon_R.$$

4. *For any $i, j \in L(u)$, if u has $d \geq 1$ unpinned neighbors, then $\frac{1}{d}\left|\Im \ln R_{G,u}^{(i,j)}(w)\right| < \varepsilon_I$.*

We will refer to items 1 to 4 as "items of the induction hypothesis". The rest of this section is devoted to the proof of this lemma via an induction on the number of unpinned vertices in G.

We begin by verifying that the induction hypothesis holds in the base case when u is the only unpinned vertex in a graph G. In this case, items 3 and 4 are vacuously true since u has no unpinned neighbors. Since all neighbors of u in G are pinned, the fact that all pinned vertices have degree at most one implies that G can be decomposed into two disjoint components G_1 and G_2, where G_1 consists of u and its pinned neighbors, while G_2 consists of a disjoint union of edges with pinned end-points. Let m be the number of conflicted edges on G_2, and let n_k denote the number of neighbors of u pinned to color k. We then have $Z^{(k)}_{G,u}(x) = x^{n_k} Z_{G_2}(x) = x^{n_k+m}$ for all $x \in \mathbb{C}$. This already proves item 1 since $w, \tilde{w} \neq 0$. Item 2 follows via the following computation (which uses item 2 of Observation 4.5.2):

$$|\ln R^{(i,j)}_{G,u}(w) - \ln R^{(i,j)}_{G,u}(\tilde{w})| = |n_i - n_j| \cdot |\ln w - \ln \tilde{w}| \leq \Delta\nu_w < \varepsilon_w.$$

We now derive some consequences of the above induction hypothesis that will be helpful in carrying out the induction.

Consequence 4.5.4. *If $|L(u)| \geq 1$, then $|Z_G(w)| > 0$.*

Proof. Note that $Z_G(w) = \sum_{i \in L(u)} Z^{(i)}_{G,u}(w)$. From item 4, we see that the angle between the complex numbers $Z^{(i)}_{G,u}(w)$ and $Z^{(j)}_{G,u}(w)$, for all $i, j \in L(u)$, is at most $d\varepsilon_I$. Applying Lemma 4.2.7 we then have

$$\left| \sum_{i \in L(u)} Z^{(i)}_{G,u}(w) \right| \geq |L(u)| \cos\frac{d\varepsilon_I}{2} \cdot \min_{i \in \Gamma_u} \left| Z^{(i)}_{G,u}(w) \right| \geq 0.9 \min_{i \in \Gamma_u} \left| Z^{(i)}_{G,u}(w) \right|,$$

when $|L(u)| \geq 1$ and $d\varepsilon_I \leq 0.01$. This last quantity is positive from item 1. ☐

Consequence 4.5.5. *For all $\varepsilon_R, \varepsilon_I, \varepsilon_w$ small enough such that $\varepsilon_I \leq \varepsilon_R$ and $\varepsilon_w \leq 0.01\varepsilon_I$, the pseudo-probabilities approximate the real probabilities in the following sense: for any $j \in L(u)$,*

$$\left| \Im \ln \frac{\mathcal{P}_{G,w}[c(u) = j]}{\mathcal{P}_{G,\tilde{w}}[c(u) = j]} \right| = |\Im \ln \mathcal{P}_{G,w}[c(u) = j]| \leq d\varepsilon_I + 2\Delta\varepsilon_w, \text{ and}$$
$$\left| \Re \ln \frac{\mathcal{P}_{G,w}[c(u) = j]}{\mathcal{P}_{G,\tilde{w}}[c(u) = j]} \right| \leq d\varepsilon_R + d\varepsilon_I + 2\Delta\varepsilon_w,$$

where d is the number of unpinned neighbors of u in G.

Proof. Using items 2 to 4 of the induction hypothesis, there exist complex numbers ξ_i (for all $i \in \Gamma_u$) satisfying $|\Re\xi_i| \le d\varepsilon_R + \varepsilon_w$ and $|\Im\xi_i| \le d\varepsilon_I + \varepsilon_w$ such that

$$\frac{\mathcal{P}_{G,\tilde{w}}[c(u)=j]}{\mathcal{P}_{G,w}[c(u)=j]} = \mathcal{P}_{G,\tilde{w}}[c(u)=j] \sum_{i \in L(u)} \frac{Z^{(i)}_{G,u}(w)}{Z^{(j)}_{G,u}(w)} = \mathcal{P}_{G,\tilde{w}}[c(u)=j] \sum_{i \in L(u)} \frac{Z^{(i)}_{G,u}(\tilde{w})}{Z^{(j)}_{G,u}(\tilde{w})} e^{\xi_i} \quad (4.41)$$

Now, note that $\sum_{i \in L(u)} \frac{Z^{(i)}_{G,u}(\tilde{w})}{Z^{(j)}_{G,u}(\tilde{w})} = \frac{1}{\mathcal{P}_{G,\tilde{w}}[c(u)=j]}$, so that the sum above is a convex combination of the $\exp(\xi_i)$. From the bounds on the real and imaginary parts of the ξ_i quoted above, by a calculation similar to that in eq. (4.11), we also have (when $\varepsilon_I, \varepsilon_w \le 0.01/\Delta$)

$$\Re e^{\xi_i} \in (\exp(-d\varepsilon_R - \varepsilon_w) - (d\varepsilon_I + \varepsilon_w)^2, \exp(d\varepsilon_R + \varepsilon_w)), \text{ and } |\arg e^{\xi_i}| \le d\varepsilon_I + \varepsilon_w. \quad (4.42)$$

The above will therefore be true also for any convex combination of the e^{ξ_i}, in particular the one in eq. (4.41). We therefore have for $C := \frac{\mathcal{P}_{G,\tilde{w}}[c(u)=j]}{\mathcal{P}_{G,w}[c(u)=j]}$

$$\Re C \in (\exp(-d\varepsilon_R - \varepsilon_w) - (d\varepsilon_I + \varepsilon_w)^2, \exp(d\varepsilon_R + \varepsilon_w)), \text{ and} \quad (4.43)$$

$$|\arg C| \le d\varepsilon_I + \varepsilon_w. \quad (4.44)$$

Now recall that for $|\theta| \le \pi/4$, we have $-\theta^2 \le \ln\cos\theta \le -\theta^2/2$. Thus, using the values of $\varepsilon_w, \varepsilon_I$ and ε_R, we have

$$|\Re \ln C| \le d\varepsilon_R + d\varepsilon_I + 2\Delta\varepsilon_w, \text{ and}$$
$$|\Im \ln C| \le d\varepsilon_I + \varepsilon_w. \qquad \square$$

As before we define $a^{(i)}_{G,u}(w) = \ln \mathcal{P}_{G,w}[c(u)=i]$, and recall the definition of the function $f_\gamma(x) := -\ln(1-\gamma e^x)$.

Consequence 4.5.6. *There exists a positive constant $\eta \in [0.9, 1)$ so that the following is true. Let d be the number of unpinned neighbors of u. Assume further that the vertex u is nice in G. Then, for any colors $i, j \in L(u)$, there exist a real constant $c = c_{G,u,i} \in [0, \frac{1}{d+\eta}]$ such that*

$$\left|\Re f_\gamma(a^{(i)}_{G,u}(w)) - f_{\tilde{\gamma}}(a^{(i)}_{G,u}(\tilde{w})) - c \cdot \Re\Big(a^{(i)}_{G,u}(w) - a^{(i)}_{G,u}(\tilde{w})\Big)\right| \le \varepsilon_I + \varepsilon_w. \quad (4.45)$$

$$\left|\Im f_\gamma(a^{(i)}_{G,u}(w)) - f_\gamma(a^{(j)}_{G,u}(w))\right| \le \frac{1}{d+\eta} \cdot (d\varepsilon_I + 4\Delta\varepsilon_w) + 2\varepsilon_w. \quad (4.46)$$

Proof. Since u is nice in G, the bound $\mathcal{P}_{G,\tilde{w}}[c(u) = k] \leq \frac{1}{d+2}$ (for any $k \in L(u)$) applies. Combining them with Consequence 4.5.5 we see that $a^{(i)}_{G,u}(w), a^{(i)}_{G,u}(\tilde{w}), a^{(j)}_{G,u}(w), a^{(j)}_{G,u}(\tilde{w})$ lie in a domain D as described in Lemma 4.2.6, with the parameters ζ and τ in that lemma chosen as

$$\begin{aligned} \zeta &= \ln(d+2) - d\varepsilon_R - d\varepsilon_I - 2\Delta\varepsilon_w \text{ and} \\ \tau &= d\varepsilon_I + 2\Delta\varepsilon_w. \end{aligned} \tag{4.47}$$

Here, for the bound on ζ, we use the fact that for $k \in L(u)$, $\mathcal{P}_{G,\tilde{w}}[c(u) = k] \leq \frac{1}{d+2}$, since u is nice in G. As in the proof of Consequence 4.4.5, we use the values of $\varepsilon_w, \varepsilon_I, \varepsilon_R$ to verify that the condition $\tau < 1/2$ and $\tau^2 + e^{-\zeta} < 1$ are satisfied, so that item 1 of Lemma 4.2.6 applies (with the parameter κ therein set to $\tilde{\gamma}$) and further that ρ_R and ρ_I as set there satisfy $\rho_R \leq \frac{1}{d+\eta}$ and $\rho_I < 3\varepsilon_I$, with $\eta = 0.94$. Using Lemma 4.2.5 followed by the bound on ε_w, we then have

$$\left|\Re f_{\tilde{\gamma}}(a^{(i)}_{G,u}(w)) - f_{\tilde{\gamma}}(a^{(i)}_{G,u}(\tilde{w})) - c \cdot \Re\Big(a^{(i)}_{G,u}(w) - a^{(i)}_{G,u}(\tilde{w})\Big)\right| \leq 3\varepsilon_I(d\varepsilon_I + 2\Delta\varepsilon_w) \leq 4d\varepsilon_I^2 \leq \varepsilon_I, \tag{4.48}$$

for an appropriate positive $c \leq 1/(d+\eta)$. This is almost eq. (4.45), whose difference will be handled later.

Similarly, applying Lemma 4.2.5 to the imaginary part we have

$$\begin{aligned} &\left|\Im f_{\tilde{\gamma}}(a^{(i)}_{G,u}(w)) - f_{\tilde{\gamma}}(a^{(j)}_{G,u}(w))\right| \\ &\qquad \leq \rho_R \cdot \max\left\{\left|\Im\Big(a^{(i)}_{G,u}(w) - a^{(j)}_{G,u}(w)\Big)\right|, \left|\Im a^{(i)}_{G,u}(w)\right|, \left|\Im a^{(j)}_{G,u}(w)\right|\right\}, \end{aligned} \tag{4.49}$$

where, as noted above, $\rho_R \leq \frac{1}{d+\eta}$. Now, note that the first term in the above maximum is less than $d\varepsilon_I + \varepsilon_w$ by items 2 and 4 of the induction hypothesis, while the other two are at most $d\varepsilon_I + 2\Delta\varepsilon_w$ from item 2 of Consequence 4.5.5.

Finally, we use item 2 of Lemma 4.2.6 with the parameter κ' therein set to γ. To this end, we note that $|\gamma - \tilde{\gamma}| \leq \varepsilon_w$, and that with the fixed values of $\varepsilon_w, \varepsilon_R$, and ε_I, the condition $(1 + \varepsilon_w) < e^{\zeta}$ is satisfied, so that the item applies. Using the item, we then see that for any $z \in D$,

$$|f_\gamma(z) - f_{\tilde{\gamma}}(z)| \leq \varepsilon_w.$$

Thus, the following quantities: $|\Re f_\gamma(a^{(i)}_{G,u}(w)) - \Re f_{\tilde{\gamma}}(a^{(i)}_{G,u}(w))|$, $|\Im f_\gamma(a^{(i)}_{G,u}(w)) - \Im f_{\tilde{\gamma}}(a^{(i)}_{G,u}(w))|$, $|\Im f_\gamma(a^{(j)}_{G,u}(w)) - \Im f_{\tilde{\gamma}}(a^{(j)}_{G,u}(w))|$, and $|\Im f_\gamma(a^{(j)}_{G,u}(w)) - \Im f_{\tilde{\gamma}}(a^{(j)}_{G,u}(w))|$ are all at most ε_w. The desired bounds now follow from the triangle inequality and the bounds in eqs. (4.48) and (4.49). □

Inductive proof of Lemma 4.5.3

We are now ready to see the inductive proof of Lemma 4.5.3; recall that the base case was already established following the statement of the lemma. Let G be any graph which satisfies Condition 1 and had at least two unpinned vertices (the base case when $|G| = 1$ was already handled above). We first prove induction item 1 for any vertex u in G. Consider the graph G' obtained from G by pinning vertex u to color i. Note that by the definition of the pinning operation, $Z^{i}_{G,u}(w) = Z_{G'}(w)$. Further, the graph G' also satisfies Condition 1, and has one fewer unpinned vertex than G. Thus, from Consequence 4.5.4 of the induction hypothesis applied to G', we have that $\left|Z^{(i)}_{G,u}(w)\right| = |Z_{G'}(w)| > 0$.

We now consider item 2. When all neighbors of u in G are pinned, the fact that all pinned vertices have degree at most one implies that G can be decomposed into two disjoint components G_1 and G_2, where G_1 consists of u and its pinned neighbors, while G_2 has one fewer unpinned vertex than G. Let n_k be the number of neighbors of u pinned to color k. Now, since G_1 and G_2 are disjoint components, we have $Z^{(k)}_{G,u}(x) = x^{n_k} Z_{G_2}(x)$ for all $k \in L(u)$ and all $x \in \mathbb{C}$. Further, from Consequence 4.5.4 of the induction hypothesis applied to G_2, we also have that $Z_{G_2}(w)$ and $Z_{G_2}(\tilde{w})$ are both non-zero. It therefore follows that

$$|\ln R^{(i,j)}_{G,u}(w) - \ln R^{(i,j)}_{G,u}(\tilde{w})| = |n_i - n_j| \cdot |\ln w - \ln \tilde{w}| \le \Delta \nu_w < \varepsilon_w.$$

We now consider items 3 and 4. Recall that by Lemma 4.2.4, we have

$$R^{(i,j)}_{G,u}(w) = \prod_{k=1}^{\deg_G(u)} \frac{\left(1 - \gamma \mathcal{P}_{G^{(i,j)}_k, w}[c(v_k) = i]\right)}{\left(1 - \gamma \mathcal{P}_{G^{(i,j)}_k, w}[c(v_k) = j]\right)}.$$

As before, for simplicity we write $G_k := G^{(i,j)}_k$. Note that each G_k has exactly one fewer unpinned vertex than G, so that the induction hypothesis applies to each G_k. Without loss of generality, we relabel the unpinned neighbors of u as $v_1, v_2, \ldots, v_d$. Let n_k be the number of neighbors of u pinned to color k. Recalling that $1 - \gamma = w$, we can then simplify the above recurrence to

$$R^{(i,j)}_{G,u}(w) = w^{n_i - n_j} \prod_{k=1}^{d} \frac{\left(1 - \gamma \mathcal{P}_{G^{(i,j)}_k, w}[c(v_k) = i]\right)}{\left(1 - \gamma \mathcal{P}_{G^{(i,j)}_k, w}[c(v_k) = j]\right)}.$$

Now, as before, for $s \in L(v_k)$ we define $a^{(s)}_{G_k, v_k}(w) := \ln \mathcal{P}_{G_k, w}[c(v_k) = s]$. From the above recurrence, we then have,

$$-\ln R^{(i,j)}_{G,u}(w) = (n_i - n_j)\ln w + \sum_{k=1}^{d} f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right). \tag{4.50}$$

Note that the same recurrence also applies when w is replaced by $\tilde{w}$ (and hence γ by $\tilde{\gamma}$):

$$-\ln R^{(i,j)}_{G,u}(\tilde{w}) = (n_i - n_j)\ln \tilde{w} + \sum_{k=1}^{d} f_{\tilde{\gamma}}\left(a^{(i)}_{G_k,v_k}(\tilde{w})\right) - f_{\tilde{\gamma}}\left(a^{(j)}_{G_k,v_k}(\tilde{w})\right). \tag{4.51}$$

(Recall that since $\Re w, \tilde{w} > 0$, $\ln w$ and $\ln \tilde{w}$ are well defined).

Using item 2 of Observation 4.5.2, $|n_i - n_j| \leq \Delta$, and the fact that $\Delta \nu_w \leq \varepsilon_w$, we have

$$|n_i - n_j|\,|\ln w - \ln \tilde{w}| \leq \varepsilon_w.$$

Applying the triangle inequality to the real part of the difference of the two recurrences, we therefore get

$$\begin{aligned}\frac{1}{d}\left|\Re \ln R^{(i,j)}_{G,u}(w) - \ln R^{(i,j)}_{G,u}(\tilde{w})\right| &\leq \varepsilon_w + \\ \max_{1\leq k\leq d}\Bigg\{\Bigg|\left(\Re f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - f_{\tilde{\gamma}}\left(a^{(i)}_{G_k,v_k}(\tilde{w})\right)\right) &- \left(\Re f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right) - f_{\tilde{\gamma}}\left(a^{(j)}_{G_k,v_k}(\tilde{w})\right)\right)\Bigg|\Bigg\}.\end{aligned} \tag{4.52}$$

In what follows, we let v_k be the vertex that maximizes the above expression, and d_k be the number of unpinned neighbors of v_k in G_k. Before proceeding with the analysis, we note that the graphs G_k satisfy Condition 1, and further that v_k is nice in G_k (the latter fact follows from Lemma 4.3.2 and the fact that G has Condition 1). Thus, the preconditions of Consequence 4.5.6 applies to the vertex v_k in graph G_k. We now proceed with the analysis.

We begin by noting that

$$\begin{aligned}&\left|\left(\Re f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - f_{\tilde{\gamma}}\left(a^{(i)}_{G_k,v_k}(\tilde{w})\right)\right) - \left(\Re f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right) - f_{\tilde{\gamma}}\left(a^{(j)}_{G_k,v_k}(\tilde{w})\right)\right)\right| \\ &\leq \max_{i',j'\in L(v_k)}\left|\left(\Re f_\gamma\left(a^{(i')}_{G_k,v_k}(w)\right) - f_{\tilde{\gamma}}\left(a^{(i')}_{G_k,v_k}(\tilde{w})\right)\right) - \left(\Re f_\gamma\left(a^{(j')}_{G_k,v_k}(w)\right) - f_{\tilde{\gamma}}\left(a^{(j')}_{G_k,v_k}(\tilde{w})\right)\right)\right|.\end{aligned}$$

On the other hand, for any color $s \in L(v_k)$, Consequence 4.5.6 of the induction hypothesis instantiated on G_k and applied to v_k and s shows that there exists a $C_s = C_{s,v_k,G_k} \in [0, 1/(d_k + \eta)]$ such that

$$\left|\Re f_\gamma\left(a^{(s)}_{G_k,v_k}(w)\right) - f_{\tilde{\gamma}}\left(a^{(s)}_{G_k,v_k}(\tilde{w})\right) - C_s(\Re a^{(s)}_{G_k,v_k}(w) - a^{(s)}_{G_k,v_k}(\tilde{w}))\right| \leq \varepsilon_I + \varepsilon_w.$$

Substituting this in the previous display shows that

$$\begin{aligned}
&\left|\left(\Re f_{\gamma}\left(a^{(i)}_{G_k,v_k}(w)\right) - f_{\tilde{\gamma}}\left(a^{(i)}_{G_k,v_k}(\tilde{w})\right)\right) - \left(\Re f_{\gamma}\left(a^{(j)}_{G_k,v_k}(w)\right) - f_{\tilde{\gamma}}\left(a^{(j)}_{G_k,v_k}(\tilde{w})\right)\right)\right| \\
&\quad\le \max_{i',j'\in L(v_k)} \left|C_{i'}(\Re a^{(i')}_{G_k,v_k}(w) - a^{(i')}_{G_k,v_k}(\tilde{w})) - C_{j'}(\Re a^{(j')}_{G_k,v_k}(w) - a^{(j')}_{G_k,v_k}(\tilde{w}))\right| + 2\varepsilon_I + 2\varepsilon_w \\
&\quad= 2\varepsilon_I + 2\varepsilon_w + \max_{i',j'\in L(v_k)} \left|C_{i'}\Re\xi_{i'} - C_{j'}\Re\xi_{j'}\right|, \\
&\quad= 2\varepsilon_I + 2\varepsilon_w + C_s\Re\xi_s - C_t\Re\xi_t, \qquad (4.53)
\end{aligned}$$

where $\xi_l := a^{(l)}_{G_k,v_k}(w) - a^{(l)}_{G_k,v_k}(\tilde{w})$ for $l \in \Gamma_{G_k,v_k}$, and s and t are given by

$$s := \arg\max_{i'\in L(v_k)} C_{i'}\Re\xi_{i'} \quad \text{and} \quad t := \arg\min_{i'\in L(v_k)} C_{i'}\Re\xi_{i'}.$$

We now have the following two cases:

Case 1: $(\Re\xi_s)\cdot(\Re\xi_t) \le 0$. Recall that C_s, C_t are non-negative and lie in $[0, 1/(d_k+\eta)]$. Thus, in this case, we must have $\Re\xi_s \ge 0$ and $\Re\xi_t \le 0$, so that

$$C_s\Re\xi_s - C_t\Re\xi_t = C_s\Re\xi_s + C_t\left|\Re\xi_t\right| \le \frac{1}{d_k+\eta}(\Re\xi_s + \left|\Re\xi_t\right|) = \frac{1}{d_k+\eta}\left|\Re\xi_s - \Re\xi_t\right|. \qquad (4.54)$$

Now, note that

$$\begin{aligned}
\Re\xi_s - \Re\xi_t &= \Re\ln\frac{\mathcal{P}_{G_k,w}[c(v_k)=s]}{\mathcal{P}_{G_k,\tilde{w}}[c(v_k)=s]} - \Re\ln\frac{\mathcal{P}_{G_k,w}[c(v_k)=t]}{\mathcal{P}_{G_k,\tilde{w}}[c(v_k)=t]} \\
&= \Re\ln\frac{\mathcal{P}_{G_k,w}[c(v_k)=s]}{\mathcal{P}_{G_k,w}[c(v_k)=t]} - \Re\ln\frac{\mathcal{P}_{G_k,\tilde{w}}[c(v_k)=s]}{\mathcal{P}_{G_k,\tilde{w}}[c(v_k)=t]} \\
&= \Re\ln R^{(s,t)}_{G_k,v_k}(w) - \ln R^{(s,t)}_{G_k,v_k}(\tilde{w}).
\end{aligned}$$

Note that all the logarithms in the above are well defined from Consequence 4.5.5 of the induction hypothesis applied to G_k and v_k. Further, from items 2 and 3 of the induction hypothesis, the last term is at most $d_k\varepsilon_R + \varepsilon_w$ in absolute value. Substituting this in eq. (4.54), we get

$$C_s\Re\xi_s - C_t\Re\xi_t \le \frac{d_k}{d_k+\eta}\varepsilon_R + \varepsilon_w. \qquad (4.55)$$

This conclude the analysis of Case 1.

Case 2: $\Re\xi_{i'}$ for $i' \in L(v_k)$ all have the same sign. Suppose first that $\Re\xi_{i'} \geq 0$ for all $i' \in L(v_k)$. Then, we have

$$0 \leq C_s\Re\xi_s - C_t\Re\xi_t \leq \frac{1}{d_k+\eta}\Re\xi_s \leq \frac{d_k}{d_k+\eta}\varepsilon_R + \varepsilon_I + 4\Delta\varepsilon_w, \tag{4.56}$$

where the last inequality follows from item 2 of Consequence 4.5.6 of the induction hypothesis applied to G_k at vertex v_k with color s, which states that $|\Re\xi_s| \leq d_k(\varepsilon_R + \varepsilon_I) + 4\Delta\varepsilon_w$. Similarly, when $\Re\xi_{i'} \leq 0$ for all $i' \in \Gamma_{G_k,v_k}$, we have

$$\begin{aligned}0 \leq C_s\Re\xi_s - C_t\Re\xi_t &= C_t|\Re\xi_t| - C_s|\Re\xi_s| \\ &\leq \frac{1}{d_k+\eta}|\Re\xi_t| \leq \frac{d_k}{d_k+\eta}\varepsilon_R + \varepsilon_I + 4\Delta\varepsilon_w,\end{aligned} \tag{4.57}$$

where the last inequality follows from item 2 of Consequence 4.5.6 of the induction hypothesis applied to G_k at vertex v_k with color t, which states that $|\Re\xi_t| \leq d_k(\varepsilon_R + \varepsilon_I) + 4\Delta\varepsilon_w$. This concludes the analysis of Case 2.

Now, substituting eqs. (4.55) to (4.57) into eq. (4.53), we get

$$\begin{aligned}&\left|\left(\Re f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - f_{\tilde\gamma}\left(a^{(i)}_{G_k,v_k}(\tilde w)\right)\right) - \left(\Re f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right) - f_{\tilde\gamma}\left(a^{(j)}_{G_k,v_k}(\tilde w)\right)\right)\right| \\ &\qquad\leq \frac{d_k}{d_k+\eta}\varepsilon_R + 3\varepsilon_I + 5\Delta\varepsilon_w.\end{aligned} \tag{4.58}$$

Substituting eq. (4.58) into eq. (4.52), we get

$$\frac{1}{d}\left|\Re\ln R^{(i,j)}_{G,u}(w) - \ln R^{(i,j)}_{G,u}(\tilde w)\right| \leq \frac{d_k}{d_k+\eta}\varepsilon_R + 3\varepsilon_I + 7\Delta\varepsilon_w < \varepsilon_R,$$

where the last inequality holds since $\eta\varepsilon_R > (\Delta+1)(3\varepsilon_I + 7\Delta\varepsilon_w)$ (recalling that $0 \leq d_k \leq \Delta$ and $\eta =\in [0.9, 1)$). This verifies item 3 of the induction hypothesis.

Finally, for proving item 4, we consider the imaginary part of eq. (4.50). We first note that

$$|n_i - n_j|\,|\Im\ln w| \leq \Delta\,|\ln w - \ln\tilde w| \leq \Delta\nu_w \leq \varepsilon_w.$$

We then have

$$\frac{1}{d}\left|\Im\ln R^{(i,j)}_{G,u}(w)\right| \leq \varepsilon_w + \max_{1\leq k\leq d}\left|\Im f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - \Im f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right)\right|. \tag{4.59}$$

Again, let v_k be the vertex that maximizes the above expression, and d_k be the number of unpinned neighbors of v_k in G_k. Applying eq. (4.46) of Consequence 4.5.6 of the induction hypothesis to the graph G_k at vertex v_k with colors $i, j \in L(v_k)$ gives

$$\left|\Im f_\gamma\left(a^{(i)}_{G_k,v_k}(w)\right) - \Im f_\gamma\left(a^{(j)}_{G_k,v_k}(w)\right)\right| \leq \frac{d_k}{d_k+\eta}\varepsilon_I + 6\Delta\varepsilon_w. \tag{4.60}$$

Substituting eq. (4.60) into eq. (4.59) we then have

$$\frac{1}{d}\left|\Im \ln R_{G,u}^{(i,j)}(w)\right| \le \frac{d_k}{d_k+\eta}\varepsilon_I + 8\Delta\varepsilon_w < \varepsilon_I,$$

where the last inequality holds since $\eta\varepsilon_I > 8(\Delta+1)\Delta\varepsilon_w$ (recalling that $0 \le d_k \le \Delta$ and $\eta \in [0.9, 1)$). This proves item 4, and also completes the inductive proof of Lemma 4.5.3. □

We now use Lemma 4.5.3 to prove Theorem 4.5.1.

Proof of Theorem 4.5.1. Let G be any graph of maximum degree Δ satisfying Condition 1. If G has no unpinned vertices, then $Z_G(w) = 1$ and there is nothing to prove. Otherwise, let u be an unpinned vertex in G. By Consequence 4.5.4 of the induction hypothesis (which we proved in Lemma 4.5.3), we then have $Z_w(G) \neq 0$ for w as in the statement of the theorem. □

The proof of Theorem 4.1.4 is now immediate.

Proof of Theorem 4.1.4. Let the quantity $\nu_w = \nu_w(\Delta)$ be as in the statements of Theorems 4.4.1 and 4.5.1. Fix the maximum degree Δ, and suppose that w satisfies

$$-\nu_w^2/8 \le \Re w \le 1 + \nu_w^2/8 \text{ and } |\Im w| \le \nu_w^2/8. \tag{4.61}$$

Let G be a graph of maximum degree Δ satisfying Condition 1. When w satisfying eq. (4.61) is such that $\Re w \le \nu_w/2$, we have $|w| \le \nu_w$, so that $Z_G(w) \neq 0$ by Theorem 4.4.1, while when such a w satisfies $\Re w \ge \nu_w/2$, we have $Z_G(w) \neq 0$ from Theorem 4.5.1. It therefore follows that $Z_G(w) \neq 0$ for all w satisfying eq. (4.61), and thus the quantity τ_Δ in the statement of Theorem 4.1.4 can be taken to be $\nu_w^2/8$. □

We conclude with a brief discussion of the dependence of τ_Δ on Δ. We saw above that τ_Δ can be taken to be $\nu_w(\Delta)^2/8$, so it is sufficient to consider the dependence of $\nu_w = \nu_w(\Delta)$ on Δ. Let $c = 10^{-6}$. As stated in the discussion following eq. (4.9), ν_w can be chosen to be $0.2c/(2^\Delta\Delta^7)$ for the case of general list colorings, or $c/(300\Delta^8)$ with the assumption of uniformly large list sizes (which, we recall from Remark 12, is satisfied in the case of uniform q-colorings). We have not tried to optimize these bounds, and it is conceivable that a more careful accounting of constants in our proofs can improve the value of the constant c by a few orders of magnitude.

Bibliography

[AG15] Nima Anari and Shayan Oveis Gharan. The Kadison-Singer problem for strongly Rayleigh measures and applications to Asymmetric TSP. In *Proc. 56th IEEE FOCS*, 2015.

[AG17] Nima Anari and Shayan Oveis Gharan. A generalization of permanent inequalities and applications in counting and optimization. In *Proc. 49th ACM STOC*, pages 384–396, 2017. arXiv:1702.02937.

[Asa70] Taro Asano. Lee-Yang theorem and the Griffiths inequality for the anisotropic Heisenberg ferromagnet. *Phys. Rev. Let.*, 24(25):1409–1411, 1970.

[Bar15a] Alexander Barvinok. Computing the partition function for cliques in a graph. *Theory Comput.*, 11(13):339–355, 2015.

[Bar15b] Alexander Barvinok. Computing the permanent of (some) complex matrices. *Found. Comput. Math.*, 16(2):329–342, 2015.

[Bar17] A. Barvinok. *Combinatorics and Complexity of Partition Functions*. Algorithms and Combinatorics. Springer, 2017.

[BB09a] Julius Borcea and Petter Brändén. The Lee-Yang and Pólya-Schur programs. I. Linear operators preserving stability. *Invent. Math.*, 177(3):541–569, 2009.

[BB09b] Julius Borcea and Petter Brändén. The Lee-Yang and Pólya-Schur programs. II. Theory of stable polynomials and applications. *Comm. Pure Appl. Math.*, 62(12):1595–1631, 2009.

[BBL09] Julius Borcea, Petter Brändén, and Thomas Liggett. Negative dependence and the geometry of polynomials. *J. AMS*, 22(2):521–567, 2009.

[BC18] Ferenc Bencs and Péter Csikvári. Note on the zero-free region of the hard-core model, July 2018. arXiv:1807.08963.

[BCKL13] Christian Borgs, Jennifer Chayes, Jeff Kahn, and László Lovász. Left and right convergence of graphs with bounded degree. *Random Struct. Algorithms*, 42(1):1–28, 2013.

[BD97] Russ Bubley and Martin Dyer. Path coupling: A technique for proving rapid mixing in Markov chains. In *Proceedings 38th Annual Symposium on Foundations of Computer Science*, pages 223–231. IEEE, 1997.

[BDPR18] Ferenc Bencs, Ewan Davies, Viresh Patel, and Guus Regts. On zero-free regions for the anti-ferromagnetic Potts model on bounded-degree graphs. *arXiv preprint arXiv:1812.07532*, 2018.

[BG01] João C. A. Barata and Pedro S. Goldbaum. On the distribution and gap structure of Lee–Yang zeros for the Ising model: Periodic and aperiodic couplings. *J. Stat. Phys.*, 103(5-6):857–891, 2001.

[BG08] Antar Bandyopadhyay and David Gamarnik. Counting without sampling: Asymptotics of the log-partition function for certain statistical physics models. *Random Structures & Algorithms*, 33(4):452–479, December 2008.

[BGK+07] Mohsen Bayati, David Gamarnik, Dimitriy Katz, Chandra Nair, and Prasad Tetali. Simple deterministic approximation algorithms for counting matchings. In *Proceedings of the thirty-ninth annual ACM symposium on Theory of computing*, pages 122–127. ACM, 2007.

[BKMP05] Noam Berger, Claire Kenyon, Elchanan Mossel, and Yuval Peres. Glauber dynamics on trees and hyperbolic graphs. *Probab. Theory Relat. Fields*, 131(3):311–340, 2005.

[BM97] J. C. A. Barata and D. H. U. Marchetti. Griffiths' singularities in diluted Ising models on the Cayley tree. *J. Stat. Phys.*, 88(1-2):231–268, 1997.

[BR17] Alexander Barvinok and Guus Regts. Weighted counting of integer points in a subspace. *arXiv preprint arXiv:1706.05423*, 2017.

[Bro41] R. L. Brooks. On colouring the nodes of a network. *Mathematical Proceedings of the Cambridge Philosophical Society*, 37(2):194–197, 1941.

[BS16a] Alexander Barvinok and Pablo Soberón. Computing the partition function for graph homomorphisms. *Combinatorica*, 2016. [Online].

[BS16b] Alexander Barvinok and Pablo Soberón. Computing the partition function for graph homomorphisms with multiplicities. *J. Combin. Theory, Ser. A*, 137:1–26, 2016.

[CCL10] Jin-Yi Cai, Xi Chen, and Pinyan Lu. Graph homomorphisms with complex values: A dichotomy theorem. In *Proc. ICALP*, volume 6198 of *Lecture Notes in Computer Science*, pages 275–286. Springer, 2010.

[CDM+19] Sitan Chen, Michelle Delcourt, Ankur Moitra, Guillem Perarnau, and Luke Postle. Improved bounds for randomly sampling colorings via linear programming. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2216–2234. SIAM, 2019.

[CF16] Péter Csikvári and Péter E Frenkel. Benjamini–Schramm continuity of root moments of graph polynomials. *Eur. J. Combin.*, 52:302–320, 2016.

[DF03] Martin Dyer and Alan Frieze. Randomly coloring graphs with lower bounds on girth and maximum degree. *Random Structures & Algorithms*, 23(2):167–179, 2003.

[DFHV13] Martin Dyer, Alan Frieze, Thomas P Hayes, and Eric Vigoda. Randomly coloring constant degree graphs. *Random Structures & Algorithms*, 43(2):181–200, 2013.

[EHŠ+16] Charilaos Efthymiou, Thomas P. Hayes, Daniel Štefankovic, Eric Vigoda, and Yitong Yin. Convergence of MCMC and loopy BP in the tree uniqueness region for the hard-core model. In *Proc. 57th IEEE FOCS*, pages 704–713, 2016.

[EM18] Lior Eldar and Saeed Mehraban. Approximating the permanent of a random matrix with vanishing mean. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 23–34. IEEE, 2018.

[Fis65] M. E. Fisher. The nature of critical points. In W. E. Brittin, editor, *Lecture notes in Theoretical Physics*, volume 7c, pages 1–159. University of Colorado Press, 1965.

[FP08] Roberto Fernández and Aldo Procacci. Regions without complex zeros for chromatic polynomials on graphs with bounded degree. *Combinatorics, Probability and Computing*, 17(2):225–238, 2008.

[FV06] Alan Frieze and Juan Vera. On randomly colouring locally sparse graphs. *Discrete Mathematics & Theoretical Computer Science*, 8(1), 2006.

[FV08] Alan Frieze and Eric Vigoda. A survey on the use of Markov chains to randomly sample colourings. *Combinatorics, Complexity and Chance: A Tribute to Dominic Welsh*, pages 53–71, 2008.

[Geo11] Georgii Hans-Otto. *Gibbs Measures and Phase Transitions.* De Gruyter, Berlin, Boston, 2011.

[GG16] Andreas Galanis and Leslie Ann Goldberg. The complexity of approximately counting in 2-spin systems on k-uniform bounded-degree hypergraphs. *Inf. Comput.*, 251:36–66, 2016.

[GGJT10] Leslie Ann Goldberg, Martin Grohe, Mark Jerrum, and Marc Thurley. A complexity dichotomy for partition functions with mixed signs. *SIAM J. Comput.*, 39(7):3336–3402, 2010.

[GGŠ+14] Andreas Galanis, Qi Ge, Daniel Štefankovič, Eric Vigoda, and Linji Yang. Improved inapproximability results for counting independent sets in the hard-core model. *Random Struct. Algorithms*, 45(1):78–110, 2014.

[GJ08] Leslie Ann Goldberg and Mark Jerrum. Inapproximability of the Tutte polynomial. *Inform. Comput.*, 206(7):908 – 929, 2008.

[GJ17] Heng Guo and Mark Jerrum. Random cluster dynamics for the Ising model is rapidly mixing. In *Proc. 28th ACM-SIAM SODA*, pages 1818–1827, 2017.

[GJP03] Leslie Ann Goldberg, Mark Jerrum, and Mike Paterson. The computational complexity of two-state spin systems. *Random Struct. Algorithms*, 23(2):133–154, 2003.

[GK12] David Gamarnik and Dmitriy Katz. Correlation decay and deterministic fptas for counting colorings of a graph. *Journal of Discrete Algorithms*, 12:29 – 47, 2012. Selected papers from the 4th Algorithms and Complexity in Durham Workshop ACiD 2010.

[GKM15] David Gamarnik, Dmitriy Katz, and Sidhant Misra. Strong spatial mixing of list coloring of graphs. *Random Struct. Algorithms*, 46(4):599–613, July 2015.

[GL16] Heng Guo and Pinyan Lu. Uniqueness, spatial mixing, and approximation for ferromagnetic 2-spin systems. In *Proc. APPROX/RANDOM*, 2016.

[GLLZ19] Heng Guo, Chao Liao, Pinyan Lu, and Chihao Zhang. Zeros of Holant problems: locations and algorithms. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2262–2278. Society for Industrial and Applied Mathematics, 2019.

[GMP04] Leslie Ann Goldberg, Russell Martin, and Mike Paterson. Strong spatial mixing for lattice graphs with fewer colours. In *45th Annual IEEE Symposium on Foundations of Computer Science*, pages 562–571. IEEE, 2004.

[GŠ11] Qi Ge and Daniel Štefankovič. Strong spatial mixing of q-colorings on Bethe lattices. *arXiv:1102.2886*, February 2011.

[GŠV15] Andreas Galanis, Daniel Štefankovič, and Eric Vigoda. Inapproximability for antiferromagnetic spin systems in the tree nonuniqueness region. *J. ACM*, 62(6):50:1–50:60, December 2015.

[Hay03] Thomas P Hayes. Randomly coloring graphs of girth at least five. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 269–278. ACM, 2003.

[HPR19] Tyler Helmuth, Will Perkins, and Guus Regts. Algorithmic Pirogov-Sinai theory. In *Proceedings of the 51st annual ACM symposium on Theory of Computing (STOC), to appear*, 2019. arXiv preprint arXiv:1806.11548.

[HV03] Thomas P Hayes and Eric Vigoda. A non-Markovian coupling for randomly sampling colorings. In *44th Annual IEEE Symposium on Foundations of Computer Science, 2003. Proceedings.*, pages 618–627. IEEE, 2003.

[HV05] Tom Hayes and Eric Vigoda. Coupling with the stationary distribution and improved sampling for colorings and independent sets. In *Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 971–979. Society for Industrial and Applied Mathematics, 2005.

[Isi25] E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Z. Phys.*, 31:253–258, 1925.

[Jer95] Mark Jerrum. A very simple algorithm for estimating the number of k-colorings of a low-degree graph. *Random Structures & Algorithms*, 7(2):157–165, 1995.

[JKP19] Matthew Jenssen, Peter Keevash, and Will Perkins. Algorithms for #BIS-hard problems on expander graphs. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 2235–2247. Society for Industrial and Applied Mathematics, 2019.

[JPS13] Bill Jackson, Aldo Procacci, and Alan D. Sokal. Complex zero-free regions at large $|q|$ for multivariate Tutte polynomials (alias Potts-model partition functions) with general complex edge weights. *Journal of Combinatorial Theory, Series B*, 103(1):21 – 45, 2013.

[JS89] M. Jerrum and A. Sinclair. Approximating the permanent. *SIAM J. Comput.*, 18(6):1149–1178, 1989.

[JS93] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.*, 22(5):1087–1116, 1993.

[JVV86] Mark Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theor. Comput. Sci.*, 43:169–188, 1986.

[KHK08] Seung-Yeon Kim, Chi-Ok Hwang, and Jin Min Kim. Partition function zeros of the antiferromagnetic Ising model on triangular lattice in the complex temperature plane for nonzero magnetic field. *Nucl. Phys. B*, 805(3):441–450, December 2008.

[LLY12] Liang Li, Pinyan Lu, and Yitong Yin. Approximate counting via correlation decay in spin systems. In *Proc. 23rd ACM-SIAM Symp. Discret. Algorithms*, pages 922–940. SIAM, 2012.

[LLY13] Liang Li, Pinyan Lu, and Yitong Yin. Correlation decay up to uniqueness in spin systems. In *Proc. 24th ACM-SIAM SODA*, pages 67–84, 2013.

[LNNP14] Yun Long, Asaf Nachmias, Weiyang Ning, and Yuval Peres. A power law of order 1/4 for critical mean-field Swendsen-Wang dynamics. *Memo. AMS*, 232(1092), 2014.

[LS12] Eyal Lubetzky and Allan Sly. Critical Ising on the square lattice mixes in polynomial time. *Comm. Math. Phys.*, 313(3):815–836, 2012.

[LSS19a] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. A deterministic algorithm for counting colorings with 2Δ colors. *arXiv preprint arXiv:1906.01228*, 2019. FOCS 2019, to appear.

[LSS19b] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. Fisher zeros and correlation decay in the ising model. In *10th Innovations in Theoretical Computer Science Conference, ITCS 2019*, pages 55:1–55:8, 2019. Full version available at arXiv:1807.06577.

[LSS19c] Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. The Ising partition function: Zeros and deterministic approximation. *Journal of Statistical Physics*, 174(2):287–315, 2019. Extended abstract appeared in IEEE FOCS 2017.

[LV97] Michael Luby and Eric Vigoda. Approximately counting up to four. In *Proc. 29th ACM STOC*, pages 682–687, 1997.

[LW01] Wentao T. Lu and F. Y. Wu. Density of the Fisher Zeroes for the Ising Model. *J. Stat. Phys.*, 102(3-4):953–970, February 2001.

[LY52] T. D. Lee and C. N. Yang. Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model. *Phys. Rev.*, 87(3):410–419, 1952.

[LY13] Pinyan Lu and Yitong Yin. Improved FPTAS for multi-spin systems. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 639–654. Springer, 2013.

[Lyo89] Russell Lyons. The Ising model and percolation on trees and tree-like graphs. *Commun. Math. Phys.*, 125(2):337–353, 1989.

[LYZ16] Pinyan Lu, Kuan Yang, and Chihao Zhang. FPTAS for hardcore and Ising models on hypergraphs. In *Proc. 33rd STACS*, pages 51:1–51:14, 2016.

[LYZZ17] Pinyan Lu, Kuan Yang, Chihao Zhang, and Minshen Zhu. An FPTAS for counting proper four-colorings on cubic graphs. In *Proc. 28th ACM-SIAM SODA*, pages 1798–1817. SIAM, 2017.

[MB18] Ryan L. Mann and Michael J. Bremner. Approximation algorithms for complex-valued Ising models on bounded degree graphs, June 2018. arXiv:1806.11282.

[MO94a] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case. *Comm. Math. Phys.*, 161(3):447–486, 1994.

[MO94b] F. Martinelli and E. Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region: II. The general case. *Comm. Math. Phys.*, 161:487–514, 1994.

[Mol04] Michael Molloy. The Glauber dynamics on colorings of a graph with high girth and maximum degree. *SIAM Journal on Computing*, 33(3):721–737, 2004.

[MS13] Elchanan Mossel and Allan Sly. Exact thresholds for Ising-Gibbs samplers on general graphs. *Ann. Probab.*, 41(1):294–328, 2013.

[MSS15a] Adam Marcus, Daniel Spielman, and Nikhil Srivastava. Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Ann. Math.*, pages 307–325, 2015.

[MSS15b] Adam W Marcus, Daniel A Spielman, and Nikhil Srivastava. Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem. *Ann. Math.*, 182:327–350, 2015.

[MSW04] Fabio Martinelli, Alistair Sinclair, and Dror Weitz. Glauber dynamics on trees: Boundary conditions and mixing time. *Comm. Math. Phys.*, 250(2):301–334, 2004.

[PR17a] Viresh Patel and Guus Regts. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM J. Comput.*, 46(6):1893–1919, 2017.

[PR17b] Han Peters and Guus Regts. On a conjecture of Sokal concerning roots of the independence polynomial, January 2017. arXiv:1701.08049.

[PR18] Han Peters and Guus Regts. Location of zeros for the partition function of the Ising model on bounded degree graphs, October 2018. arXiv:1810.01699.

[Rue10] David Ruelle. Characterization of Lee-Yang polynomials. *Ann. Math.*, 171(1):589–603, 2010.

[RW99] Dana Randall and David Wilson. Sampling spin configurations of an Ising system. In *Proc. 10th ACM-SIAM SODA*, pages 959–960, 1999.

[SF71] Masuo Suzuki and Michael E. Fisher. Zeros of the partition function for the Heisenberg, Ferroelectric, and general Ising models. *J. Math. Phys.*, 12(2):235–246, 1971.

[SF99] R.P. Stanley and S. Fomin. *Enumerative Combinatorics*. Cambridge University Press, 1999.

[She85] James B. Shearer. On a problem of Spencer. *Combinatorica*, 5(3), 1985.

[Sim93] Barry Simon. *The Statistical Mechanics of Lattice Gases*, volume 1. Princeton University Press, 1993.

[Sly10] Allan Sly. Computational transition at the uniqueness threshold. In *Proc. 51st IEEE FOCS*, pages 287–296. IEEE, 2010.

[Sok01] Alan D Sokal. Bounds on the complex zeros of (di)chromatic polynomials and potts-model partition functions. *Combinatorics, Probability and Computing*, 10(1):41–77, 2001.

[Sok05] Alan D Sokal. The multivariate Tutte polynomial (alias Potts model). *Surveys in combinatorics 2005*, 327:173, 2005.

[SS97] Jesús Salas and Alan D Sokal. Absence of phase transition for antiferromagnetic Potts models via the Dobrushin uniqueness theorem. *Journal of Statistical Physics*, 86(3-4):551–579, 1997.

[SS04] Alex Scott and Alan Sokal. The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma. *J. Stat. Phys.*, 118(5-6):1151–1261, 2004.

[SS14a] Alistair Sinclair and Piyush Srivastava. Lee-Yang theorems and the complexity of computing averages. *Comm. Math. Phys.*, 329(3):827–858, 2014.

[SS14b] Allan Sly and Nike Sun. Counting in two-spin models on d-regular graphs. *Ann. Probab.*, 42(6):2383–2416, 2014.

[SSŠY16] Alistair Sinclair, Piyush Srivastava, Daniel Štefankovič, and Yitong Yin. Spatial mixing and the connective constant: Optimal bounds. *Probab. Theory Relat. Fields*, pages 1–45, 2016.

[SST14] Alistair Sinclair, Piyush Srivastava, and Marc Thurley. Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. *J. Stat. Phys.*, 155(4):666–686, 2014.

[SV17] Damian Straszak and Nisheeth K. Vishnoi. Real stable polynomials and matroids: Optimization and counting. In *Proc. 49th ACM STOC*, pages 370–383, 2017. arXiv:1611.04548.

[SYZ16] Renjie Song, Yitong Yin, and Jinman Zhao. Counting hypergraph matchings up to uniqueness threshold. In *Proc. APPROX/RANDOM*, pages 46:1–46:29, 2016.

[Vig00] Eric Vigoda. Improved bounds for sampling colorings. *Journal of Mathematical Physics*, 41(3):1555–1569, 2000.

[Wei05] Dror Weitz. Combinatorial criteria for uniqueness of Gibbs measures. *Random Structures & Algorithms*, 27(4):445–475, 2005.

[Wei06] Dror Weitz. Counting independent sets up to the tree threshold. In *Proc. 38th ACM STOC*, pages 140–149, 2006.

[YL52] C. N. Yang and T. D. Lee. Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Phys. Rev.*, 87(3):404–409, 1952.

[ZLB11] Jinshan Zhang, Heng Liang, and Fengshan Bai. Approximating partition functions of the two-state spin system. *Inf. Process. Lett.*, 111(14):702–710, 2011.