



The Ising Partition Function: Zeros and Deterministic Approximation

Jingcheng Liu¹ · Alistair Sinclair¹ · Piyush Srivastava²

Received: 29 August 2018 / Accepted: 22 November 2018 / Published online: 8 December 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

We study the problem of approximating the partition function of the ferromagnetic Ising model with both pairwise as well as higher order interactions (equivalently, in graphs as well as hypergraphs). Our approach is based on the classical Lee–Yang theory of phase transitions, along with a new Lee–Yang theorem for the Ising model with higher order interactions, and on an extension of ideas developed recently by Barvinok, and Patel and Regts that can be seen as an algorithmic realization of the Lee–Yang theory. Our first result is a *deterministic* polynomial time approximation scheme (an FPTAS) for the partition function in bounded degree graphs that is valid over the entire range of parameters β (the interaction) and λ (the external field), except for the case $|\lambda| = 1$ (the “zero-field” case). A polynomial time *randomized* approximation scheme (FPRAS) for all graphs and all β, λ , based on Markov chain Monte Carlo simulation, has long been known. Unlike most other deterministic approximation algorithms for problems in statistical physics and counting, our algorithm does not rely on the “decay of correlations” property, but, as pointed out above, on Lee–Yang theory. This approach extends to the more general setting of the Ising model on hypergraphs of bounded degree and edge size, where no previous algorithms (even randomized) were known for a wide range of parameters. In order to achieve this latter extension, we establish a tight version of the Lee–Yang theorem for the Ising model on hypergraphs, improving a classical result of Suzuki and Fisher.

Keywords Lee–Yang theorems · Ising model · Partition functions · Higher-order interactions

An extended abstract of this paper appeared in the proceedings of the *58th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, 2017, pp. 986–997.

JL and AS are supported in part by US NSF Grants CCF-1420934 and CCF-1815328. PS is supported by a Ramanujan Fellowship of the Indian Department of Science and Technology. Some of this work was done at the Simons Institute for the Theory of Computing at UC Berkeley.

✉ Alistair Sinclair
sinclair@cs.berkeley.edu

Extended author information available on the last page of the article

1 Introduction

The Ising model, first studied a century ago as a model for magnetic materials by Lenz and Ising [26], has since become an important tool for the modeling of interacting systems. In the Ising model, such a system is represented as a graph $G = (V, E)$, so that the individual 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 $\sigma : V \rightarrow \{+, -\}$ of one of two possible values (often called “spins”) to each vertex. The model then induces a probability distribution (known as a *Gibbs distribution*) over these global configurations in terms of local parameters that model the interactions of the vertices.

In our setting, it will be convenient to parameterize these interactions in terms of a *vertex activity* or *fugacity* λ , that models an “external field” and determines the propensity of a vertex to be in the $+$ configuration, and an *edge activity* $\beta \geq 0$ that models the tendency of vertices to agree with their neighbors. The model assigns to each configuration σ a weight

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

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

$$Z_G^\beta(\lambda) := \sum_{\sigma: V \rightarrow \{+, -\}} w(\sigma) = \sum_{S \subseteq V} \beta^{|E(S, \bar{S})|} \lambda^{|S|}. \quad (1)$$

Notice that the partition function may be interpreted combinatorially as a cut generating polynomial in the graph G .

In this paper we focus on the original *ferromagnetic* case in which $\beta < 1$, 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 paper 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 $\beta_e < 1$ 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.

As in almost all statistical physics and graphical models, the partition function captures the computational complexity of the Ising model. Partition functions are $\#P$ -hard¹ 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 [28] 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.

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 intractability of efficient exact computation. For a more detailed account of this phenomenon in the context of partition functions, see, e.g., [47, Appendix A].

For several statistical physics models on bounded degree graphs (including the anti-ferromagnetic Ising model [31,48] and the “hard core”, or independent set model [55]), 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 complementary results prove that the partition function is NP-hard to approximate outside this regime [20,50]. 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 [28]: to the best of our knowledge, the best deterministic algorithm, due to Zhang, Liang and Bai [57], 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 [48, Appendix 2] for a more detailed discussion of this point.

The first goal of this paper is to supply such a deterministic algorithm which covers almost the entire range of parameters of the model except for the “zero-field” case $|\lambda| = 1$:

Theorem 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 (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 1.3 and 1.4 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 [31,49,55] as well as our present analytic continuation approach. In contrast, approximate counting algorithms based on Markov chain Monte Carlo (e.g., [18,27,35]) often have the desirable feature that they are in a sense “fixed parameter tractable”: even as the fixed parameters of the problems are moved close to the boundary up to which the algorithm is applicable, the degree of the polynomial bounding their 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 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

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. (3)]. The running time of the algorithm is polynomial in n and $1/\varepsilon$.

sub-graphs may grow as $n^{\Theta(\log \Delta)}$ (see Sect. 3, and in particular, the proof of Lemma 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 [21, 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 graphs. This leads to the following tantalizing possibility: extending Theorem 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 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 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 the paper.

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 [6, 7] (see also the recent monograph by Barvinok [8] 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 $\log Z$ [56].

To apply this approach in the case of the ferromagnetic Ising model, we may appeal to the famous Lee–Yang theorem of the 1950s [30], 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 $Z(\lambda) = \lambda^n Z(\frac{1}{\bar{\lambda}})$.

Remark 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 [4, 5]. 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 time³ in order to achieve a $(1 \pm 1/\text{poly}(n))$ -factor multiplicative approximation of Z (equivalently, a $1/\text{poly}(n)$ additive approximation of $\log Z$), since this requires the Taylor series for $\log Z$ to be evaluated to $k = \Omega(\log n)$ terms [7,9,10].

Recently, Patel and Regts [42] 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 [15]. 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 [42, 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 (where in fact the correlation decay property also holds), specifically the regime $|\beta - 1| \leq 0.34/\Delta$ and $\lambda = 1$.

In this paper, 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 Sect. 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 1.1. We note that if one is only interested in deriving Theorem 1.1, then this can also be done using the notion of *fragments*, developed by Patel and Regts [42] 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 [54], and also more recently in the literature on approximate counting [19,33,51]. 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 [28] 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

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

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 (1), where the cut $E(S, \bar{S})$ consists of those hyperedges with at least one vertex in each of S and \bar{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 [54] prove a weak version of the Lee–Yang theorem for hypergraphs (see Theorem 4.3 in Sect. 4), which we are able to strengthen to obtain the following optimal statement, which is of independent interest:

Theorem 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 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 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 1.2 follows along the lines of Asano’s inductive proof of the Lee–Yang theorem [3], 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 4.5.

Combining Theorem 1.2 with our earlier algorithmic approach immediately yields the following generalization of Theorem 1.1 to hypergraphs.

Theorem 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 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 [54]. A two-spin system on a hypergraph $H = (V, E)$ is specified by hyperedge activities $\varphi_e : \{+, -\}^{|e|} \rightarrow \mathbb{C}$ for $e \in E$, and a vertex activity $\psi : \{+, -\} \rightarrow \mathbb{C}$. (Note that we treat each hyperedge e as a set of vertices.) Then the partition function is defined as:

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

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

$$Z_H^\varphi(\lambda) = \sum_{\sigma: V \rightarrow \{+, -\}} \prod_{e \in E} \varphi_e(\sigma|_e) \lambda^{|v: \sigma(v)=+|}. \quad (2)$$

We call a hypergraph two-spin system *symmetric* if $\varphi_e(\sigma) = \overline{\varphi_e(-\sigma)}$. Suzuki and Fisher [54] proved a Lee–Yang theorem for symmetric hypergraph two-spin systems (which is weaker than our Theorem 1.2 above when specialized to the Ising model). Combining this with our general algorithmic approach yields our final result:

Theorem 1.4 Fix any $\Delta > 0$ and $k \geq 2$ and a family of symmetric edge activities $\varphi = \{\varphi_e\}$ satisfying $|\varphi_e(+, \dots, +)| \geq \frac{1}{4} \sum_{\sigma \in \{+, -\}^V} |\varphi_e(\sigma)|$. Then there exists an FPTAS for the partition function $Z_H^\varphi(\lambda)$ 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 $|\lambda| \neq 1$.

The remainder of the paper is organized as follows. In Sect. 2, we spell out Baryinok's approach to approximating partition functions using Taylor series. Section 3 introduces the notion of insects and shows how to use them to efficiently compute the lowest-degree coefficients of the partition function in the general context of hypergraphs; as discussed above, this machinery applied to graphs, in conjunction with the Lee–Yang theorem, implies Theorem 1.1. Finally, in Sect. 4 we prove our extension of the Lee–Yang theorem to the hypergraph Ising model (Theorem 1.2), and then use it and the Suzuki–Fisher theorem to prove our algorithmic results for hypergraphs, Theorems 1.3 and 1.4.

1.1 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., [16, 22]).

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 [29]. 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 [28] 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 [43].) This was later extended to ferromagnetic two-spin systems by Goldberg, Jerrum and Paterson [23]. Similar techniques have been applied recently to the related random-cluster model by Guo and Jerrum [24].

Much detailed work has been done on MCMC for Ising spin configurations for several important classes of graphs, including two-dimensional lattices (e.g., [34, 38, 39]), random graphs and graphs of bounded degree (e.g., [41]), the complete graph (e.g., [32]) and trees (e.g., [11, 40]); 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 [55] 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) [31, 48]. A complementary result of Sly and Sun [50] (see also [20]) 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 [33], 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 et al. [57] handle only the “high-temperature” regime of the Ising model, while the recent results for ferromagnetic two-spin systems of Guo and Lu [25] 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 [6,7,9,10]. More recently, Patel and Regts [42] 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 [56] in connection with the analysis of phase transitions. By proving the famous Lee–Yang circle theorem for the ferromagnetic Ising model [30], they were able to conclude that there can be at most one phase transition for the model. Asano [3] extended the Lee–Yang theorem to the Heisenberg model, and provided a simpler proof. Asano’s work was generalized further by Suzuki and Fisher [54], while Sinclair and Srivastava [47] 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 [44]. 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 [46] and Scott and Sokal [45] 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 [36]. Such Lee–Yang type theorems are exemplars of the more general stability theory of polynomials [12,13], a field of study that has had numerous recent applications to theoretical computer science and combinatorics (see, e.g., [1,2,14,36,37,47,53]).

2 Approximation of the Log-Partition Function by Taylor Series

In this section we present an approach due to Barvinok [7] for approximating the partition function of a physical system by truncating the Taylor series of its logarithm, as discussed in the introduction. 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), such a system has partition function

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

Our goal is an FPTAS for $Z_H^\varphi(\lambda)$, 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 $Z_H^\varphi(\lambda)$, i.e., a number \hat{Z} satisfying

$$|\hat{Z} - Z_H^\varphi(\lambda)| \leq \varepsilon |Z_H^\varphi(\lambda)|. \quad (3)$$

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

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

$$f(\lambda) = \sum_{j=0}^{\infty} f^{(j)}(0) \cdot \frac{\lambda^j}{j!}, \quad (4)$$

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 $f, \tilde{f} \in \mathbb{C}$ such that $|f - \tilde{f}| \leq \varepsilon$, we have

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

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 $f'(\lambda) = \frac{1}{Z(\lambda)} \frac{dZ(\lambda)}{d\lambda}$, or $\frac{dZ(\lambda)}{d\lambda} = f'(\lambda)Z(\lambda)$. Thus for any $m \geq 1$,

$$\frac{d^m}{d\lambda^m} Z(\lambda) = \sum_{j=0}^{m-1} \binom{m-1}{j} \frac{d^j}{d\lambda^j} Z(\lambda) \cdot \frac{d^{m-j}}{d\lambda^{m-j}} f(\lambda). \quad (5)$$

Given $\frac{d^j}{d\lambda^j} Z(\lambda) \big|_{\lambda=0}$ for $j = 0, \dots, m$, Eq. (5) is a triangular system of linear equations in $\{f^{(j)}(0)\}_{j=1}^m$ of representation length $\text{poly}(m)$, and is non-degenerate since $Z(0) = 1$; hence it can be solved in $\text{poly}(m)$ time.

We can now specify the algorithm: first compute $\left\{ \frac{d^j}{d\lambda^j} Z(\lambda) \big|_{\lambda=0} \right\}_{j=0}^m$; next, use the system in Eq. (5) to solve for $\{f^{(j)}(0)\}_{j=1}^m$; and finally, compute and output 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, \dots, r_n of Z . Throughout this paper, we will be using (some variant of) the Lee–Yang theorem to argue that, for the range of interactions φ 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 $\varphi_e(-, \dots, -) = 1$, the constant term $\prod_{i=1}^n (-r_i)$ of $Z(\lambda)$ is 1, and hence we have $Z(\lambda) = \prod_{i=1}^n (1 - \frac{\lambda}{r_i})$. The log partition function can then be written as

$$f(\lambda) = \log Z(\lambda) = \sum_{i=1}^n \log \left(1 - \frac{\lambda}{r_i} \right) = - \sum_{i=1}^n \sum_{j=1}^{\infty} \frac{1}{j} \left(\frac{\lambda}{r_i} \right)^j. \quad (6)$$

Note that due to the uniqueness of the Taylor expansion of meromorphic functions, the two power series expansions of $f(\lambda)$ in Eqs. (4) and (6) are identical in the domain of their convergence. Denoting the first m terms of the above expansion by $f_m(\lambda) = - \sum_{i=1}^n \sum_{j=1}^m \frac{1}{j} \left(\frac{\lambda}{r_i} \right)^j$, 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 2.1 *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 $\left\{ \frac{d^j}{d\lambda^j} Z(\lambda) \Big|_{\lambda=0} \right\}_{j=0}^m$, $f_m(\lambda)$ can be computed in time $\text{poly}(n/\varepsilon)$. 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 2.1 into an FPTAS is the computation of the derivatives $\frac{d^j}{d\lambda^j} Z(\lambda) \Big|_{\lambda=0}$ 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 2.1 gives an FPTAS.

3 Computing Coefficients via insects

As discussed in the introduction, Patel and Regts [42] 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 [42] as much as possible. From now on, unless otherwise stated, we will use G to denote a hypergraph. Recall from the introduction the partition function of a two-spin system on a hypergraph $G = (V, E)$:

$$Z_G^\varphi(\lambda) = \sum_{\sigma: V \rightarrow \{+, -\}} \prod_{e \in E} \varphi_e(\sigma|_e) \lambda^{\{v: \sigma(v) = +\}}. \quad (7)$$

Due to the normalization $\varphi_e(-, \dots, -) = 1$, each term in the summation depends only on the set $S = \{v : \sigma(v) = +\}$ and the labelled induced sub-hypergraph $(S \cup \partial S, E[S] \cup E(S, \bar{S}))$, 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 σ^S be the configuration where the set of vertices assigned $+$ -spins is S , that is, $\sigma^S(v) = +$ for $v \in S$ and $\sigma^S(v) = -$ otherwise. We will also write $\varphi_e(S) := \varphi_e(\sigma^S|_e)$ 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, \dots, r_n , where $n = |V|$. As explained in the paragraph preceding Eq. (6), the assumption $\varphi_e(-, \dots, -) = 1$ 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 $(\frac{1}{r_1}, \dots, \frac{1}{r_n})$.

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

$$e_i(G) = (-1)^i \sum_{\substack{S \subseteq V \\ |S|=i}} \prod_{e: e \cap S \neq \emptyset} \varphi_e(S). \quad (8)$$

Once we have computed the first m coefficients of Z (i.e., the values $e_i(G)$ for $i = 1, \dots, m$), where $m = \Omega\left(\frac{\log(n/\varepsilon) - \log(1-|\lambda|)}{\log(1-|\lambda|)}\right)$, we can use Lemma 2.1 to obtain an FPTAS as claimed in Theorems 1.1, 1.3 and 1.4. The main result in this section will be an algorithm for computing these coefficients $e_i(G)$:

Theorem 3.1 *Fix $k, \Delta \in \mathbb{N}$ and $C > 0$. There exists a deterministic $\text{poly}(n/\varepsilon)$ -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 $i = 1, \dots, m$, where $m = \lceil C \log(n/\varepsilon) \rceil$.*

3.1 Insects in a Hypergraph

To take advantage of the fact that each term in Eq. (7) only depends on the set S and the induced sub-hypergraph $(S \cup \partial S, E[S] \cup E(S, \bar{S}))$, we define the following structure.

Definition 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 $B(H) := (\bigcup_{e \in E} e) \setminus S$ is called the *boundary set*.

Given an insect H , we use the notation $V(H)$ for its label set. The *size* $|H|$ of the insect H is defined to be $|V(H)|$. 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 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. (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 $S \subseteq V(G)$ such that $G^+[S] = H$.

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 $f(G) = \sum_{S \subseteq V(G)} a_{G^+[S]}$ 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 [17], also employed in the case of graph invariants by Patel and Regts [42], can then be adapted to give the following:

Lemma 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 . \square

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

- 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. \square

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 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,

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

where in the second line we have used Lemma 3.5, and where

$$c_H := \sum_{\substack{H_1, H_2 \text{ compatible} \\ H = H_1 \cup H_2}} a_{H_1} b_{H_2}. \quad (9)$$

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. \square

3.3 Enumerating Connected Sub-insects

We observe next that $e_i(G)$, as defined in Eq. (8), can be written as a weighted sub-insect count. Accordingly, we generalize Eq. (8) to arbitrary insects G of maximum degree Δ and hyperedge size k as follows:

$$e_i(G) = (-1)^i \sum_{\substack{S \subseteq V(G) \\ |S|=i}} \prod_{\substack{e: e \cap S \neq \emptyset}} \varphi_e(S) = \sum_{H \in \mathcal{G}_i^{\Delta, k}} \mu_H \cdot \mathbf{1}[H \hookrightarrow G], \quad (10)$$

where $\mu_H := (-1)^i \prod_{e: e \cap V(H) \neq \emptyset} \varphi_e(\underline{V}(H))$. Note that this definition coincides with Eq. (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 $\text{poly}(|H|)$; however, as discussed in the introduction, 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. (10) would be inefficient. To prove Theorem 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 [42], we use the following calculation of Borgs et al. [15, Lemma 2.1 (c)].

Lemma 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 [52, 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 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 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 [42, Lemma 3.4]. However, for the sake of completeness we also include a direct proof.

Lemma 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 $\tilde{O}(nt^3(e\Delta k)^{t+1})$. Here \tilde{O} hides factors of the form $\text{polylog}(n)$, $\text{polylog}(\Delta k)$ and $\text{polylog}(t)$.*

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 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 $\tilde{O}(t)$), in time $\tilde{O}(nt^3(e\Delta k)^t)$.

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

3.4 Proof of Theorem 3.1

The results in the previous subsection allow us to efficiently enumerate connected sub-insects. To prove Theorem 3.1, it remains to reduce the sum over all (possibly disconnected) H in Eq. (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 [42] 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 , $1 \leq i \leq n$, of $Z_G(\lambda)$ [extended to insects G as in the paragraph following Eq. (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

$$p_t = \sum_{i=1}^{t-1} (-1)^{i-1} p_{t-i} e_i + (-1)^{t-1} t e_t. \quad (11)$$

Recall from Eq. (10) that e_i is a weighted sub-insect count supported on insects of size $\leq i$, and also from Corollary 3.6 that the product of two weighted sub-insect counts supported on insects of size $\leq t_1, t_2$ respectively is a weighted sub-insect count supported on insects of size $\leq t_1 + t_2$. Therefore, by Eq. (11) and induction, each p_t is also a weighted sub-insect count supported on insects of size $\leq t$. Thus, for any insect G , we may write

$$p_t(G) = \sum_{H \in \mathcal{G}_t^{\Delta, k}} a_H^{(t)} \cdot \mathbf{1}[H \hookrightarrow G] \quad (12)$$

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 3.3, the coefficients of p_t are supported on *connected* insects, and we may write Eq. (12) as

$$p_t(G) = \sum_{H \in \mathcal{C}_t^{\Delta,k}} a_H^{(t)} \cdot \mathbf{1}[H \hookrightarrow G]. \quad (13)$$

Notice that by Corollary 3.8, there are at most $n(e\Delta k)^t$ non-zero terms in this sum.

Lemma 3.10 *There is a $\text{poly}(n/\varepsilon)$ -time algorithm to compute all the coefficients $a_H^{(t)}$ in Eq. (13), for $t \leq O(\log(n/\varepsilon))$.*

Proof By Lemma 3.9, we compute \mathcal{T}_t , consisting of all $S \subseteq V(G)$ such that $|S| \leq t$ 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. (11), for $t = 1$ we have $p_1 = e_1$, so by Eq. (10) we can read off the coefficients $a_H^{(1)}$ from $e_1(G)$. Next suppose we have computed $a_{H'}^{(t')}$ for $|H'| \leq t' < t$, 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. (11), it suffices to compute the coefficient corresponding to H in $p_{t-i}e_i$ for each $1 \leq i \leq k-1$ (since the contribution of the last term in Eq. (11) is simply $(-1)^{t-1}t\mu_H$ if $|H| = t$ and 0 otherwise). By Eqs. (9) and (13), this coefficient is given by

$$\sum_{\substack{H_1 \in \mathcal{G}_i^{\Delta,k}, H_2 \in \mathcal{C}_{(t-i)}^{\Delta,k} \\ H_1 \text{ compatible with } H_2 \\ H_1 \cup H_2 = H}} a_{H_2}^{(t-i)} \mu_{H_1}^{(i)} = \sum_{\substack{(S_1, S_2) \\ S_1 \cup S_2 = \underline{V}(H) \\ H^+[S_1] \in \mathcal{G}_i^{\Delta,k}, H^+[S_2] \in \mathcal{C}_{(t-i)}^{\Delta,k}}} a_{H^+[S_2]}^{(t-i)} \mu_{H^+[S_1]}. \quad (14)$$

Since $t \leq O(\log(n/\varepsilon))$, the second sum involves at most $4^t = \text{poly}(n/\varepsilon)$ terms. Moreover, due to Corollary 3.8, there are at most $nt(e\Delta k)^t = \text{poly}(n/\varepsilon)$ previously computed $a_{H'}^{(t')}$, where H' is a connected sub-insect of G and $|H'| \leq t' < t$. In order to look up $a_{H^+[S]}^{(t-i)}$, one can do a linear scan, which also takes time $\text{poly}(n/\varepsilon)$ for $t \leq O(\log(n/\varepsilon))$. The coefficients $\mu_{H^+[S]}$ can simply be read off from their definition in Eq. (10).

To conclude, because $t \leq O(\log(n/\varepsilon))$, Eq. (13) only contains $\text{poly}(n/\varepsilon)$ terms. And for each term, $a_H^{(t)}$ can be computed using the above dynamic programming scheme in $\text{poly}(n/\varepsilon)$ time. \square

Finally, now that we can compute $a_{H,t}$ efficiently, by Eq. (13) we can compute p_k using the sub-insect enumerator in Lemma 3.9, and we can then compute e_k using Newton's identities as in Eq. (11), which completes the proof of Theorem 3.1.

3.5 Proofs of Main Theorems

Our first main result in the introduction, the FPTAS for the Ising model on graphs throughout the ferromagnetic regime with non-zero field stated in Theorem 1.1, now follows by combining Theorem 3.1 with Lemma 2.1 and the Lee–Yang theorem [30] (also stated as Theorem 4.2 in the next section). Recall from the introduction that the Lee–Yang theorem ensures that the partition function has no zeros inside the unit disk.

Similarly, Theorem 1.4, the FPTAS for two-spin systems on hypergraphs, follows by combining Theorem 3.1 with Lemma 2.1 and the Suzuki–Fisher theorem [54] (also stated as Theorem 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 1.4.

To establish our final main algorithmic result, Theorem 1.3, we first need to prove a new Lee–Yang theorem for the hypergraph Ising model as stated in Theorem 1.2 in the introduction. This will be the content of the next and final section of our paper. Once we have that, Theorem 1.3 follows immediately by the same route as above.

4 A Lee–Yang Theorem for Hypergraphs

In this section we prove a tight Lee–Yang theorem for the hypergraph Ising model (Theorem 1.2 in the introduction). 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 $\lambda_1, \lambda_2, \dots, \lambda_n$ and hyperedge activities $\beta = (\beta_e)$, we define

$$Z_G^\beta(\lambda_1, \dots, \lambda_n) = \sum_{S \subseteq V} \prod_{e \in E(S, \bar{S})} \beta_e \prod_{i \in S} \lambda_i,$$

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

$$Z_G^\beta(\lambda_1, \dots, \lambda_n) = \prod_{i=1}^n \lambda_i \cdot Z_G^\beta\left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\right). \quad (15)$$

We use the following definition of the Lee–Yang property. This definition is based on the results of Asano [3] and Suzuki and Fisher [54], and somewhat stricter than the definition used by Ruelle [44].

Definition 4.1 (*Lee–Yang property*). Let $P(z_1, z_2, \dots, z_n)$ be a multilinear polynomial. P is said to have the *Lee–Yang property* (sometimes written as “ P is LY”) if for any complex numbers $\lambda_1, \dots, \lambda_n$ such that $|\lambda_1| \geq 1, \dots, |\lambda_n| \geq 1$, and $|\lambda_i| > 1$ for some i , it holds that $P(\lambda_1, \dots, \lambda_n) \neq 0$.

Then the seminal Lee–Yang theorem [30] can be stated as follows:

Theorem 4.2 *Let G be a connected undirected graph, and suppose $0 < \beta < 1$. Then the Ising partition function $Z_G^\beta(\lambda_1, \dots, \lambda_n)$ 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 [54]. Again the theorem is stated in the multivariate setting, where in the two-spin partition function in Eq. (7) each vertex i has a distinct activity λ_i .

Theorem 4.3 *Consider any symmetric hypergraph two-spin system, with a connected hypergraph G and edge activities $\{\varphi_e\}$. Then the partition function $Z_G^\varphi(\lambda_1, \dots, \lambda_n)$ has the Lee–Yang property if $|\varphi_e(+, \dots, +)| \geq \frac{1}{4} \sum_{\sigma \in \{+, -\}^V} |\varphi_e(\sigma)|$ for every hyperedge e .*

Theorem 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 4.4 Let $G = (V, E)$ be a connected hypergraph, and $\beta = (\beta_e)_{e \in E}$ be a vector of real valued hyperedge activities so that the activity of edge $e \in E$ is β_e . Then Z_G^β 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 4.2).

The following corollary for the univariate polynomial $Z_G^\beta(\lambda)$ follows immediately via Eq. (15) and the fact that, by Hurwitz’s theorem, the zeros of $Z_G^\beta(\lambda)$ are continuous functions of β and thus remain on the unit circle after taking the limit in the range of each β_e .

Corollary 4.5 Let $G = (V, E)$ be a connected hypergraph, and $\beta = (\beta_e)_{e \in E}$ 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 $Z_G^\beta(\lambda)$ 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 1.2 in the introduction, and hence also Theorem 1.3 as explained at the end of the previous section. The second part of Theorem 1.2, which asserts that the range of edge activities under which the theorem holds is optimal, is proven in Sect. 4.1. (Note that the optimality for the univariate case claimed in Theorem 1.2 does not directly follow from the optimality for the multivariate case guaranteed by Theorem 4.4 above.)

Remark As a comparison, the result of Suzuki and Fisher, which we restated in Theorem 4.3, implies that a sufficient condition for the Lee–Yang property of $Z_G^\beta(\lambda)$ 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 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 4.4 in Sect. 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 [44], 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 [44].

Lemma 4.6 *Given a multilinear polynomial $P(z_1, z_2, \dots, z_n)$ with real coefficients define, for each $1 \leq j \leq n$, multilinear polynomials A_j and B_j in the variables $z_1, \dots, z_{j-1}, z_{j+1}, \dots, z_n$ 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 $A_j(z_1, \dots, z_{j-1}, z_{j+1}, \dots, z_n) \neq 0$ 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 $\lambda_2, \dots, \lambda_n$ satisfying $|\lambda_i| \geq 1$ such that $A_1(\lambda_2, \dots, \lambda_n) = 0$. Since P is LY, it follows that $B_1(\lambda_2, \dots, \lambda_n) \neq 0$ (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 $(\lambda_2, \dots, \lambda_n)$ in \mathbb{C}^{n-1} . In particular, let S_ε be the open set

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

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 $(\lambda_2, \dots, \lambda_n)$ it follows from continuity that for $\varepsilon < \varepsilon_0$ small enough, $|A_1| \leq \delta_0/2$ in S_ε . Since A_1 does not vanish identically on S_ε , there must exist a point (y_2, \dots, y_n) in S_ε such that $0 < |A_1(y_2, \dots, y_n)| < \delta_0/2$. Since $|B_1(y_2, \dots, y_n)| \geq \delta_0$ by the choice of ε_0 , it follows that if we define $y_1 = -B_1(y_2, \dots, y_n)/A_1(y_2, \dots, y_n)$ then $2 < |y_1| < \infty$. However, we then have $P(y_1, y_2, \dots, y_n) = 0$ even though $|y_1| > 1$ and $|y_i| \geq 1$ for all i . This contradicts the Lee–Yang property of P . \square

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

Corollary 4.7 *Let $P(z_1, z_2, \dots, z_n)$ be a multilinear polynomial with non-zero real coefficients, i.e.,*

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

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

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

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

We next show that Lemma 4.6 has a partial converse for symmetric multilinear functions.

Lemma 4.8 *Let $P(z_1, z_2, \dots, z_n)$ be a symmetric multilinear polynomial with non-zero real coefficients, i.e.,*

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

where $p_S \neq 0$ for all $S \subseteq [n]$ and $p_S = p_{\bar{S}}$.

Assume further that the polynomials A_j as defined in Lemma 4.6 all have the property that they are non-zero when all their arguments z_i satisfy $|z_i| \geq 1$. 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 $i \geq 2$, then any root $z_1 = \xi_1$ of the equation $A_1 z_1 + B_1 = 0$ satisfies $|\xi_1| = 1$. (Here A_1 and B_1 are in the notation of Lemma 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

$$\begin{aligned} |\xi_1| = \left| \frac{B_1}{A_1} \right| &= \left| \frac{\sum_{S \subseteq [2, n]} p_S \prod_{i \in S} z_i}{\sum_{S \subseteq [2, n]} p_{S \cup \{1\}} \prod_{i \in S} z_i} \right| = \left| \left(\prod_{i \in [2, n]} z_i \right) \frac{\sum_{S \subseteq [2, n]} p_S \prod_{i \notin S} (1/z_i)}{\sum_{S \subseteq [2, n]} p_{S \cup \{1\}} \prod_{i \in S} z_i} \right| \\ &\stackrel{(\star)}{=} \left| \frac{\sum_{S \subseteq [2, n]} p_S \prod_{i \notin S} \bar{z}_i}{\sum_{S \subseteq [2, n]} p_{S \cup \{1\}} \prod_{i \in S} z_i} \right| \stackrel{(\dagger)}{=} \left| \frac{\sum_{S \subseteq [2, n]} p_{S \cup \{1\}} \prod_{i \in S} \bar{z}_i}{\sum_{S \subseteq [2, n]} p_{S \cup \{1\}} \prod_{i \in S} z_i} \right| = 1. \end{aligned} \quad (16)$$

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, \dots, 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 $(\xi_1, \xi_2, \xi_3, \dots, \xi_n)$ of P such that $|\xi_i| \geq 1$ for all i , and exactly $k - 1$ of the ξ_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

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

where $\alpha_{12}, \alpha_1, \alpha_2$ and α_{\emptyset} are non-zero polynomials in z_3, \dots, 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 4.6, $\alpha_{12}(z_3, \dots, z_n) \neq 0$ when $|z_i| \geq 1$ for $i \geq 3$. Now, again by hypothesis, $A_2 \neq 0$ when $|z_1|$ and $|z_3|, \dots, |z_n|$ are at least 1, while $z_1 = -\frac{\alpha_2(z_3, \dots, z_n)}{\alpha_{12}(z_3, \dots, z_n)}$ gives $A_2 = 0$. Thus, we must have that

$$\frac{|\alpha_2(z_3, \dots, z_n)|}{|\alpha_{12}(z_3, \dots, z_n)|} < 1 \text{ when } |z_i| \geq 1 \text{ for } i \geq 3. \quad (17)$$

We now set $\xi_i = z_i$ for $i \geq 3$, and consider z_1 as a function of z_2 . The equality $P(z_1, z_2, \xi_3, \dots, \xi_n) = 0$ is then equivalent to

$$z_1 = -\frac{\alpha_2 z_2 + \alpha_{\emptyset}}{\alpha_{12} z_2 + \alpha_1}, \quad (18)$$

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

$$\lim_{z_2 \rightarrow \infty} |z_1| = \frac{|\alpha_2|}{|\alpha_{12}|} < 1. \quad (19)$$

Initially, both z_1 and z_2 lie outside the closed unit disk. Thus, by Eq. (19) and continuity, we can take z_2 large enough in absolute value such that z_1 as defined in Eq. (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 $P(\zeta_1, \dots, \zeta_n) = 0$ 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. \square

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

Lemma 4.9 *Let m be any integer, and k a positive integer such that $2|m| \leq k$. Consider the maximization problem*

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

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,

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

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 $2|m| \leq k$. \square

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

Lemma 4.10 *Fix an integer $k \geq 2$ and a hyperedge activity $\beta \in \mathbb{R}$. Let $G = (V = \{v_1, v_2, \dots, v_k\}, E = \{\{v_1, v_2, \dots, v_k\}\})$ 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 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 [30] (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 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

$$\frac{1}{\beta} = 1 - \prod_{i=2}^k \left(1 + \frac{1}{z_i}\right). \quad (20)$$

To establish the lemma, we therefore only need to show that for the claimed values of β , Eq. (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. (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 $y = r \exp(i\theta)$ for $\theta \in [-\pi/2, \pi/2]$ and $0 \leq r \leq 2 \cos \theta$. Consider now the set $\mathbb{R} \cap \{\prod_{i=2}^k y_i \mid y_i \in D \text{ for } 2 \leq i \leq k\}$. 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. (20) cannot hold.

Recalling that each $y \in D$ can be written in the form $r \exp(i\theta)$ 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 $k \geq 3$):

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

Using Lemma 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. \square

We now proceed to an inductive proof of Theorem 4.4, using Lemma 4.10 as the base case.

Proof of Theorem 4.4 The case $k = 2$ is a special case of the Lee–Yang theorem [30] (though, as with the proof of Lemma 4.10, the argument below can again be specialized to directly establish this). We assume therefore that $k \geq 3$.

The proof uses the inductive method of Asano [3]. When the hypergraph consists of a single hyperedge of size $k' \leq k$, it follows from Lemma 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 $k' \leq k$ 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 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 4.7, it can be proved, by adapting an argument first developed by Asano [3], that these two operations preserve the Lee–Yang property. Asano’s method has by now become standard (see, e.g., [54, 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 $P(z_1, z_2, \dots, z_n) = A(z_2, \dots, z_n)z_1 + B(z_2, \dots, z_n)$ and $Q(y_1, y_2, \dots, y_{k'}) = C(y_2, \dots, y_{k'})y_1 + D(y_2, \dots, y_{k'})$ 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

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

where z is a new variable corresponding to the new vertex created by the merger of u_1 and v_1 . Let $\lambda_2, \dots, \lambda_n$, and $\mu_2, \dots, \mu_{k'}$ be complex numbers lying outside the open unit disk. In order to prove that R is LY, we need to show that (i) $R(z, \lambda_2, \dots, \lambda_n, \mu_2, \dots, \mu_{k'}) = 0$ implies that $|z| \leq 1$; and (ii) when at least one of these complex numbers lies strictly outside the *closed* unit disk then $R(z, \lambda_2, \dots, \lambda_n, \mu_2, \dots, \mu_{k'}) = 0$ implies that $|z| < 1$. Now, since P and Q are assumed to be LY, Lemma 4.6 implies that $A = A(\lambda_2, \dots, \lambda_n)$ and $C = C(\mu_2, \dots, \mu_{k'})$ are both non-zero. Thus, $R = 0$ implies that

$$|z| = |B/A| \cdot |D/C|, \quad (21)$$

where $B = B(\lambda_2, \dots, \lambda_n)$ and $D = D(\mu_2, \dots, \mu_{k'})$. 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. (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. (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

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

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

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

where z is a new variable corresponding to the vertex created by the merger of v_1 and v_2 . Now, let $\lambda_3, \dots, \lambda_n$ be complex numbers lying outside the open unit disk. Corollary 4.7 implies that $A = A(\lambda_3, \dots, \lambda_n) \neq 0$. Thus, $R(z, \lambda_3, \dots, \lambda_n) = 0$ implies that

$$|z| = |D(\lambda_3, \dots, \lambda_n)/A(\lambda_3, \dots, \lambda_n)| = |D/A|. \quad (22)$$

Now, since P is LY, both zeros of the quadratic equation $P(x, x, \lambda_3, \dots, \lambda_n) = 0$ 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. Equation (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 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

$$P(z_1, z_2, \dots, z_k) := \beta \prod_{i=1}^k (1 + z_i) + (1 - \beta) \left(1 + \prod_{i=1}^k z_i \right). \quad (23)$$

Our strategy is to show that when

$$\beta \notin \left(-\frac{1}{2^{k-1} - 1}, \frac{1}{2^{k-1} \cos^{k-1}(\frac{\pi}{k-1}) + 1} \right), \quad (24)$$

the polynomial $A_1(z_2, z_3, \dots, z_k)$, which is the coefficient of z_1 in P as defined in Lemma 4.6, vanishes at a point with $|z_i| \geq 1$ for $i \geq 2$. It then follows from Lemma 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 4.10. Let D be the closed disk of radius 1 centered at 1, as defined in the proof of that lemma. Equation (20), taken together with the discussion following it, implies that

finding a zero of $A_1(z_2, \dots, z_k)$ 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 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 $y_i = r e^{i\theta}$, $2 \leq i \leq k$, and

$$1 - \frac{1}{\beta} = (r e^{i\theta})^{k-1}. \quad (25)$$

Let $\gamma := 1 - \frac{1}{\beta}$. We now partition the condition on β in (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 (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 $\theta = \frac{\pi}{k-1}$, $r = |\gamma|^{\frac{1}{k-1}} \leq 2 \cos \theta$ gives a solution in this case. \square

4.1 Optimality of the Univariate Hypergraph Lee–Yang Theorem

We now prove the second part of the univariate hypergraph Lee–Yang theorem, Theorem 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

$$\beta \notin \left[-\frac{1}{2^{k-1}-1}, \frac{1}{2^{k-1} \cos^{k-1}(\frac{\pi}{k-1}) + 1} \right]. \quad (26)$$

We consider three exhaustive cases under which (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: $\frac{1}{2^{k-1} \cos^{k-1}(\frac{\pi}{k-1}) + 1} < \beta < 1$. 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

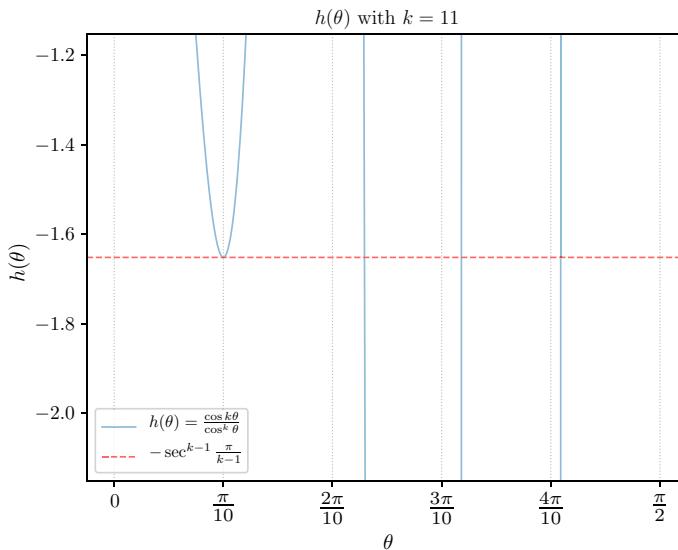


Fig. 1 The function $h(\theta)$

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 $z = e^{2i\theta}$, $\theta \in (-\pi/2, \pi/2)$, we have

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

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 $z = e^{2i\theta}$ 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 $\tau < -\sec^{k-1}(\pi / (k-1))$.

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 $0 \leq i < \lfloor k/2 \rfloor$ 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 $0 \leq i \leq \lfloor k/2 \rfloor - 1$. We note the following facts (see Fig. 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 $i < \lfloor k/2 \rfloor$, $h(\rho_i) = (-1)^i \sec^{k-1}(\frac{i\pi}{k-1})$, so that $h(\rho_i)$ is strictly positive when i is even and strictly negative when i is odd. Further, $h(\rho_1) = -\sec^{k-1}(\frac{\pi}{k-1}) > \tau$.

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} .
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.

Acknowledgements We thank Alexander Barvinok, Guus Regts and anonymous reviewers for helpful comments.

References

1. Anari, N., Gharan, S.O.: The Kadison–Singer problem for strongly Rayleigh measures and applications to Asymmetric TSP. In: Proceedings of the 56th IEEE Symposium on Foundations of Computer Science (FOCS) (2015)
2. Anari, N., Gharan, S.O.: A generalization of permanent inequalities and applications in counting and optimization. In: Proceedings of the 49th ACM Symposium on Theory and Computation (STOC), pp. 384–396. [arXiv:1702.02937](https://arxiv.org/abs/1702.02937) (2017)
3. Asano, T.: Lee–Yang theorem and the Griffiths inequality for the anisotropic Heisenberg ferromagnet. *Phys. Rev. Lett.* **24**(25), 1409–1411 (1970). <https://doi.org/10.1103/PhysRevLett.24.1409>
4. Barata, J.C.A., Goldbaum, P.S.: 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). <https://doi.org/10.1023/A:101033250031>
5. Barata, J.C.A., Marchetti, D.H.U.: Griffiths’ singularities in diluted Ising models on the Cayley tree. *J. Stat. Phys.* **88**(1–2), 231–268 (1997). <https://doi.org/10.1007/BF02508471>
6. Barvinok, A.: Computing the partition function for cliques in a graph. *Theory Comput.* **11**(13), 339–355 (2015)
7. Barvinok, A.: Computing the permanent of (some) complex matrices. *Found. Comput. Math.* **16**(2), 329–342 (2015). <https://doi.org/10.1007/s10208-014-9243-7>
8. Barvinok, A.: Combinatorics and Complexity of Partition Functions. Springer, New York (2016)
9. Barvinok, A., Soberón, P.: Computing the partition function for graph homomorphisms with multiplicities. *J. Combin. Theory Ser. A* **137**, 1–26 (2016)
10. Barvinok, A., Soberón, P.: Computing the partition function for graph homomorphisms. *Combinatorica* **37**(4), 633–650 (2017)
11. Berger, N., Kenyon, C., Mossel, E., Peres, Y.: Glauber dynamics on trees and hyperbolic graphs. *Probab. Theory Relat. Fields* **131**(3), 311–340 (2005). <https://doi.org/10.1007/s00440-004-0369-4>
12. Borcea, J., Brändén, P.: The Lee–Yang and Pólya–Schur programs. I. Linear operators preserving stability. *Invent. Math.* **177**(3), 541–569 (2009)
13. Borcea, J., Brändén, P.: The Lee–Yang and Pólya–Schur programs. II. Theory of stable polynomials and applications. *Commun. Pure Appl. Math.* **62**(12), 1595–1631 (2009)
14. Borcea, J., Brändén, P., Liggett, T.: Negative dependence and the geometry of polynomials. *J. AMS* **22**(2), 521–567 (2009)
15. Borgs, C., Chayes, J., Kahn, J., Lovász, L.: Left and right convergence of graphs with bounded degree. *Random Struct. Algorithms* **42**(1), 1–28 (2013). <https://doi.org/10.1002/rsa.20414>

16. Cai, J.Y., Chen, X., Lu, P.: Graph homomorphisms with complex values: a dichotomy theorem. In: Proceedings of the ICALP, Lecture Notes in Computer Science, vol. **6198**, pp. 275–286. Springer. <http://www.springerlink.com/content/46275700132p5250/abstract/> (2010)
17. Csikvári, P., Frenkel, P.E.: Benjamini–Schramm continuity of root moments of graph polynomials. *Eur. J. Combin.* **52**, 302–320 (2016)
18. Efthymiou, C., Hayes, T.P., Štefankovič, D., Vigoda, E., Yin, Y.: Convergence of MCMC and loopy BP in the tree uniqueness region for the hard-core model. In: Proceedings of the 57th IEEE Symposium on Foundations of Computer Science (FOCS), pp. 704–713 (2016)
19. Galanis, A., Goldberg, L.A.: The complexity of approximately counting in 2-spin systems on k -uniform bounded-degree hypergraphs. *Inf. Comput.* **251**, 36–66 (2016). <https://doi.org/10.1016/j.ic.2016.07.003>. <http://www.sciencedirect.com/science/article/pii/S0890540116300426>
20. Galanis, A., Ge, Q., Štefankovič, D., Vigoda, E., Yang, L.: Improved inapproximability results for counting independent sets in the hard-core model. *Random Struct. Algorithms* **45**(1), 78–110 (2014). <https://doi.org/10.1002/rsa.20479>
21. Goldberg, L.A., Jerrum, M.: Inapproximability of the Tutte polynomial. *Inf. Comput.* **206**(7), 908–929 (2008). <https://doi.org/10.1016/j.ic.2008.04.003>. <http://www.sciencedirect.com/science/article/pii/S089054010800031X>
22. Goldberg, L.A., Grohe, M., Jerrum, M., Thurley, M.: A complexity dichotomy for partition functions with mixed signs. *SIAM J. Comput.* **39**(7), 3336–3402 (2010). <https://doi.org/10.1137/090757496>
23. Goldberg, L.A., Jerrum, M., Paterson, M.: The computational complexity of two-state spin systems. *Random Struct. Algorithms* **23**(2), 133–154 (2003)
24. Guo, H., Jerrum, M.: Random cluster dynamics for the Ising model is rapidly mixing. In: Proceedings of the 28th ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 1818–1827 (2017)
25. Guo, H., Lu, P.: Uniqueness, spatial mixing, and approximation for ferromagnetic 2-spin systems. In: Proceedings of the APPROX/RANDOM (2016)
26. Ising, E.: Beitrag zur Theorie des Ferromagnetismus. *Z. Phys.* **31**, 253–258 (1925). <https://doi.org/10.1007/BF02980577>
27. Jerrum, M., Sinclair, A.: Approximating the permanent. *SIAM J. Comput.* **18**(6), 1149–1178 (1989). <https://doi.org/10.1137/0218077>
28. Jerrum, M., Sinclair, A.: Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.* **22**(5), 1087–1116 (1993)
29. Jerrum, M., Valiant, L.G., Vazirani, V.V.: Random generation of combinatorial structures from a uniform distribution. *Theor. Comput. Sci.* **43**, 169–188 (1986)
30. Lee, T.D., Yang, C.N.: Statistical theory of equations of state and phase transitions. II. Lattice gas and Ising model. *Phys. Rev.* **87**(3), 410–419 (1952). <https://doi.org/10.1103/PhysRev.87.410>
31. Li, L., Lu, P., Yin, Y.: Correlation decay up to uniqueness in spin systems. In: Proceedings of the 24th ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 67–84 (2013)
32. Long, Y., Nachmias, A., Ning, W., Peres, Y.: A power law of order 1/4 for critical mean-field Swendsen–Wang dynamics. *AMS MEMO/232/1092* (2014)
33. Lu, P., Yang, K., Zhang, C.: FPTAS for hardcore and Ising models on hypergraphs. In: Proceedings of the 33rd Symposium on Theoretical Aspects of Computer Science (STACS), pp. 51:1–51:14 (2016)
34. Lubetzky, E., Sly, A.: Critical Ising on the square lattice mixes in polynomial time. *Commun. Math. Phys.* **313**(3), 815–836 (2012). <https://doi.org/10.1007/s00220-012-1460-9>
35. Lubly, M., Vigoda, E.: Approximately counting up to four. In: Proceedings of the 29th ACM Symposium on Theory of Computing, pp. 682–687. <https://doi.org/10.1145/258533.258663> (1997)
36. Marcus, A., Spielman, D., Srivastava, N.: Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Ann. Math.* (2015) <https://doi.org/10.4007/annals.2015.182.1.7>. <http://annals.math.princeton.edu/2015/182-1/p07>
37. Marcus, A.W., Spielman, D.A., Srivastava, N.: Interlacing families II: Mixed characteristic polynomials and the Kadison–Singer problem. *Ann. Math.* **182**, 327–350 (2015)
38. Martinelli, F., Olivieri, E.: Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case. *Commun. Math. Phys.* **161**(3), 447–486 (1994)
39. Martinelli, F., Olivieri, E.: Approach to equilibrium of Glauber dynamics in the one phase region: II. The general case. *Commun. Math. Phys.* **161**, 487–514 (1994)
40. Martinelli, F., Sinclair, A., Weitz, D.: Glauber dynamics on trees: Boundary conditions and mixing time. *Commun. Math. Phys.* **250**(2), 301–334 (2004). <https://doi.org/10.1007/s00220-004-1147-y>
41. Mossel, E., Sly, A.: Exact thresholds for Ising–Gibbs samplers on general graphs. *Ann. Probab.* **41**(1), 294–328 (2013)
42. Patel, V., Regts, G.: Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM J. Comput.* **46**(6), 1893–1919 (2017). [arXiv:1607.01167](https://arxiv.org/abs/1607.01167)

43. Randall, D., Wilson, D.: Sampling spin configurations of an Ising system. In: Proceedings of the 10th ACM-SIAM Symposium on Discrete Algorithms (SODA), pp. 959–960 (1999)
44. Ruelle, D.: Characterization of Lee-Yang polynomials. *Ann. Math.* **171**(1), 589–603 (2010). <https://doi.org/10.4007/annals.2010.171.589>. <http://annals.math.princeton.edu/2010/171-1/p16>
45. Scott, A., Sokal, A.: The repulsive lattice gas, the independent-set polynomial, and the Lovász local lemma. *J. Stat. Phys.* **118**(5–6), 1151–1261 (2004)
46. Shearer, J.B.: On a problem of Spencer. *Combinatorica* **5**(3), 241–245 (1985)
47. Sinclair, A., Srivastava, P.: Lee–Yang theorems and the complexity of computing averages. *Commun. Math. Phys.* **329**(3), 827–858 (2014)
48. Sinclair, A., Srivastava, P., Thurley, M.: Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs. *J. Stat. Phys.* **155**(4), 666–686 (2014)
49. Sinclair, A., Srivastava, P., Štefankovič, D., Yin, Y.: Spatial mixing and the connective constant: optimal bounds. *Probab. Theory Relat. Fields* **168**, 153–197 (2016)
50. Sly, A., Sun, N.: Counting in two-spin models on d -regular graphs. *Ann. Probab.* **42**(6), 2383–2416 (2014). <https://doi.org/10.1214/13-AOP888>. <http://projecteuclid.org/euclid.aop/1412083628>
51. Song, R., Yin, Y., Zhao, J.: Counting hypergraph matchings up to uniqueness threshold. In: Proceedings of the APPROX/RANDOM, pp. 46:1–46:29 (2016)
52. Stanley, R., Fomin, S.: *Enumerative Combinatorics*. Cambridge University Press, Cambridge (1999)
53. Straszak, D., Vishnoi, N.K.: Real stable polynomials and matroids: optimization and counting. In: Proceedings of the 49th ACM Symposium on Theory of Computing (STOC), pp. 370–383. [arXiv:1611.04548](https://arxiv.org/abs/1611.04548) (2017)
54. Suzuki, M., Fisher, M.E.: Zeros of the partition function for the Heisenberg, Ferroelectric, and general Ising models. *J. Math. Phys.* **12**(2), 235–246 (1971). <https://doi.org/10.1063/1.1665583>. <http://scitation.aip.org/content/aip/journal/jmp/12/2/10.1063/1.1665583;jsessionid=tsaFQsrLOe4npdBWQR-8iADE.x-aip-live-06>
55. Weitz, D.: Counting independent sets up to the tree threshold. In: Proceedings of the 38th ACM Symposium on Theory of Computing (STOC), pp. 140–149 (2006)
56. Yang, C.N., Lee, T.D.: Statistical theory of equations of state and phase transitions. I. Theory of condensation. *Phys. Rev.* **87**(3), 404–409 (1952). <https://doi.org/10.1103/PhysRev.87.404>
57. Zhang, J., Liang, H., Bai, F.: Approximating partition functions of the two-state spin system. *Inf. Process. Lett.* **111**(14), 702–710 (2011). <https://doi.org/10.1016/j.ipl.2011.04.012>

Affiliations

Jingcheng Liu¹ · Alistair Sinclair¹ · Piyush Srivastava²

Jingcheng Liu
liuexp@berkeley.edu

Piyush Srivastava
piyush.srivastava@tifr.res.in

¹ Computer Science Division, UC Berkeley, Berkeley, USA

² Tata Institute of Fundamental Research, Mumbai, India