

Counting independent sets in unbalanced bipartite graphs

Sarah Cannon ^{*}

Will Perkins [†]

Abstract

Understanding the complexity of approximately counting the number of weighted or unweighted independent sets in a bipartite graph ($\#$ BIS) is a central open problem in the field of approximate counting. Here we consider a subclass of this problem and give an FPTAS for approximating the partition function of the hard-core model for bipartite graphs when there is sufficient imbalance in the degrees or fugacities between the sides (L, R) of the bipartition. This includes, among others, the biregular case when $\lambda = 1$ (approximating the number of independent sets of G) and $\Delta_R \geq 7\Delta_L \log(\Delta_L)$. Our approximation algorithm is based on truncating the cluster expansion of a polymer model partition function that expresses the hard-core partition function in terms of deviations from independent sets that are empty on one side of the bipartition.

Further consequences of this method for unbalanced bipartite graphs include an efficient sampling algorithm for the hard-core model and zero-freeness results for the partition function with complex fugacities. By utilizing connections between the cluster expansion and joint cumulants of certain random variables, we go beyond previous algorithmic applications of the cluster expansion to prove that the hard-core model exhibits exponential decay of correlations for all graphs and fugacities satisfying our conditions. This illustrates the applicability of statistical mechanics tools to algorithmic problems and refines our understanding of the connections between different methods of approximate counting.

1 Introduction

The computational complexity of approximating the number of independent sets in a bipartite graph is a central open problem in the field of approximate counting: neither a general polynomial-time algorithm nor a proof of NP-hardness is known. The problem defines a complexity class $\#$ BIS (bipartite independent set) and many other important counting problems have been shown to be as hard to approximate as $\#$ BIS, including

counting weighted independent sets in bounded-degree bipartite graphs when the weighting factor λ is large enough. Counting weighted independent sets arises naturally in statistical physics as it is exactly the problem of computing the partition function of the hard-core model at *fugacity* λ .

We give an FPTAS for both the weighted and unweighted counting problems, provided that the bipartite graph has bounded degree and there is a sufficient imbalance in the degrees or fugacities between the sides of the bipartition.

More formally, the hard-core model on a graph G is a probability distribution over the independent sets of G given by

$$\mu_G(I) = \frac{\lambda^{|I|}}{Z(G, \lambda)}$$

where $\lambda > 0$ is the fugacity and the normalizing constant, or the partition function, is

$$Z(G, \lambda) = \sum_{I \in \mathcal{I}(G)} \lambda^{|I|},$$

where $\mathcal{I}(G)$ is the set of all independent sets of G .

More generally, one can consider the multivariate hard-core model, assigning a fugacity λ_v to each vertex $v \in V(G)$. The resulting partition function is

$$Z(G) = \sum_{I \in \mathcal{I}(G)} \prod_{v \in I} \lambda_v.$$

We focus here on a bivariate hard-core model, with a fugacities λ_L, λ_R assigned to vertices on the respective sides of a bipartite graph with bipartition (L, R) . We denote the partition function by $Z(G, \lambda_L, \lambda_R)$ or $Z(G)$ for short.

There are two natural computational problems associated to spin models such as the hard-core model: the approximate counting problem, where we wish to estimate $Z(G)$, and the approximate sampling problem, where we wish to sample from a distribution that is close to μ_G .

Formally, an ϵ -relative approximation to $Z(G)$ is a number $\hat{Z}$ so that

$$e^{-\epsilon} \hat{Z} \leq Z(G) \leq e^{\epsilon} \hat{Z}.$$

^{*}Claremont McKenna College, scannon@cmc.edu, supported in part by NSF award DMS-1803325.

[†]University of Illinois at Chicago, math@willperkins.org, supported in part by NSF Career award DMS-1847451.

A *fully polynomial-time approximation scheme* (FPTAS) for $Z(G)$ is an algorithm that given G and $\epsilon > 0$, outputs an ϵ -relative approximation to $Z(G)$ and runs in time polynomial in $|V(G)|$ and $1/\epsilon$. An *efficient sampling scheme* is a randomized algorithm that outputs an independent set from $\mathcal{I}(G)$ with distribution $\hat{\mu}$ so that $\|\mu_G - \hat{\mu}\|_{TV} < \epsilon$ and runs in time polynomial in $|V(G)|$ and $1/\epsilon$.

In general, the computational complexity of approximating $Z(G)$ is well understood. For graphs of maximum degree Δ , there is an FPTAS due to Weitz [29] when $\lambda < \lambda_c(\Delta)$, the uniqueness threshold for the infinite Δ -regular tree. On the other hand, Sly [27], Sly and Sun [28], and Galanis, Stefankovic, and Vigoda [11] showed that there is no polynomial-time approximation algorithm for $\lambda > \lambda_c(\Delta)$ unless $\text{NP}=\text{RP}$.

For bipartite graphs, however, the situation is far from clear. There is no hardness known, and one might expect that the problem becomes easier as λ gets large since finding a maximum size independent set in a bipartite graph is tractable, unlike in general graphs. In fact, Dyer, Greenhill, Goldberg, and Jerrum [9] defined the complexity class $\#\text{BIS}$ to capture the complexity of approximating the number of independent sets (i.e. approximating $Z(G)$ at $\lambda = 1$) in bipartite graphs. The approximation complexity of $\#\text{BIS}$ is still unresolved, but many other important approximate counting problems have been shown to be $\#\text{BIS}$ -hard, including approximating the partition function of the q -color ferromagnetic Potts model ($q \geq 3$) [13, 12]. Refined results of Cai, Galanis, Goldberg, Guo, Jerrum, Stefankovic, and Vigoda show that for bipartite graphs of maximum degree Δ , it is already $\#\text{BIS}$ -hard to approximate $Z(G)$ for any $\lambda > \lambda_c(\Delta)$ [5].

There are a handful of algorithmic results on approximating $Z(G)$ that exploit bipartite structure. Liu and Lu [23] gave an FPTAS for $Z(G)$ at $\lambda = 1$ on bipartite graphs when the degrees on one side of the bipartition are at most 5, with arbitrary degrees on the other side. Helmuth, Perkins, and Regts [15] used contour models from Pirogov-Sinai theory in statistical physics to give an FPTAS for the hard-core model on subsets of $\mathbb{Z}^d$ for sufficiently large λ . Building on this approach, Jenssen, Keevash, and Perkins [16] gave an FPTAS for the hard-core model on bipartite expander graphs at large λ , and these results were sharpened for random regular bipartite graphs [17, 22]. Most relevant for this paper, Barvinok and Regts [2] give an FPTAS for $Z(G)$ for biregular, bipartite graphs with unequal degrees when the fugacity is sufficiently large, as an application of a more general approximate counting result.

Here we give an FPTAS for the hard-core model on bipartite graphs of bounded degree whenever there is

sufficient asymmetry in the degrees on either side of the bipartition or in the fugacities assigned to the respective sides of the bipartition. In most biregular cases our results give significant improvement to the parameters from [2], as we discuss below, and our algorithm does not require biregularity.

More importantly, the method, based on the cluster expansion and the Kotecký-Preiss condition [20] and related to that of [15, 16], gives detailed probabilistic information about the hard-core model in addition to the algorithmic results. In particular, we show that on bipartite graphs with parameters satisfying our conditions, the correlation between the occupancies of two vertices in the hard-core model decays exponentially fast in their distance. Our results hold for a wide range of parameters for which the hard-core model on the appropriate infinite bi-regular tree exhibits phase coexistence, and so our results cannot be obtained by methods which exploit uniqueness on the tree. The connection between convergence of the cluster expansion and decay of correlation for lattice spin models is well studied in statistical physics, but has not been widely utilized in computer science. We hope that by illustrating the applicability of these techniques to general graphs and showing their connection to algorithmic results, we may encourage their adoption.

1.1 Algorithmic results Let $\mathcal{G}(\Delta_L, \delta_R, \Delta_R)$ be the family of bipartite graphs G with bipartition (L, R) so that each $v \in L$ has degree $d_v \leq \Delta_L$ and each $v \in R$ has degree d_v satisfying $\delta_R \leq d_v \leq \Delta_R$.

We consider the hard-core model on a graph $G \in \mathcal{G}(\Delta_L, \delta_R, \Delta_R)$ with fugacity λ_L assigned to each vertex $v \in L$ and fugacity λ_R assigned to each $v \in R$. Our main algorithmic result is a sufficient condition on the imbalance of the graph (in terms of $\Delta_L, \delta_R, \Delta_R, \lambda_L, \lambda_R$) to obtain an FPTAS for approximating $Z(G, \lambda_L, \lambda_R)$.

THEOREM 1.1. *Suppose*

$$(1.1) \quad 6\Delta_L\Delta_R\lambda_R \leq (1 + \lambda_L)^{\frac{\delta_R}{\Delta_L}}.$$

Then there is an FPTAS for approximating the hard-core partition function $Z(G, \lambda_L, \lambda_R)$ and an efficient sampling scheme for sampling from μ_G for all $G \in \mathcal{G}(\Delta_L, \delta_R, \Delta_R)$ at fugacities λ_L and λ_R .

Some special cases of Theorem 1.1 for the bivariate and univariate hard-core partition functions are given in the following corollary. Throughout, we assume all logarithms have base e .

COROLLARY 1.1. *There is an FPTAS for approximating $Z(G)$ and an efficient sampling scheme for μ_G when:*

1. G is a Δ -regular bipartite graph, and the fugacities λ_L and λ_R satisfy

$$\lambda_L \geq 6\Delta^2\lambda_R.$$

2. G is a biregular bipartite graph with degrees $\Delta_R > \Delta_L$, and the fugacity λ satisfies

$$\lambda > (6\Delta_L\Delta_R)^{\frac{\Delta_L}{\Delta_R - \Delta_L}}.$$

3. G is a biregular bipartite graph with degrees Δ_L and Δ_R satisfying

$$\Delta_R \geq 7\Delta_L \log(\Delta_L),$$

and the fugacity is $\lambda = 1$.

We can compare Corollary 1.1, part (2), to the condition given by Barvinok and Regts [2] for an FPTAS for biregular, bipartite graphs at fugacity λ : $\Delta_R > \Delta_L$ and

$$\lambda \geq \left(6.7\sqrt{\Delta_R}\right)^{\Delta_L + \Delta_R}.$$

Our bound of $\lambda > (6\Delta_L\Delta_R)^{\frac{\Delta_L}{\Delta_R - \Delta_L}}$ gives a significant improvement except in the case when $\Delta_R = \Delta_L + 1$. In fact, for Δ_L fixed, our bound on λ decreases as Δ_R grows.

1.2 Correlation decay We next show that for graphs satisfying condition (1.1), vertex-to-vertex correlations decay exponentially fast in their distance. We are able to do this because the convergence of the cluster expansion, which we use to obtain the FPTAS described above, gives detailed probabilistic information about the model.

Let $\mathbf{I}$ be a random independent set drawn from the hard-core model μ_G . For $v \in V(G)$, let $X_v = \mathbf{1}_{v \in \mathbf{I}}$, and $\mu_v = \mathbb{E}X_v = \Pr[v \in \mathbf{I}]$. For $u, v \in V(G)$ let $X_{uv} = \mathbf{1}_{u, v \in \mathbf{I}}$, and $\mu_{uv} = \mathbb{E}X_{uv} = \Pr[u \in \mathbf{I} \wedge v \in \mathbf{I}]$. For $A \subseteq V(G)$, let $X_A = \mathbf{1}_{A \subseteq \mathbf{I}}$, and $\mu_A = \Pr[A \subseteq \mathbf{I}]$.

THEOREM 1.2. *For $\Delta_L, \Delta_R, \delta_R, \lambda_L, \lambda_R$ satisfying (1.1), there exists constants $\epsilon > 0, C > 0$ so that the following holds.*

For all $G \in \mathcal{G}(\Delta_L, \delta_R, \Delta_R)$, and for all $u, v \in V(G)$, we have

$$(1.2) \quad |\mu_{uv} - \mu_u\mu_v| \leq Ce^{-\epsilon D(u, v)},$$

where $D(\cdot, \cdot)$ is the graph distance in G , and μ_G is the hard-core model on G with fugacities λ_L, λ_R . More generally, suppose $A \subset V(G), B \subset V(G)$. Then

$$(1.3) \quad |\mu_{A \cup B} - \mu_A\mu_B| \leq C'e^{-\epsilon D(A, B)},$$

where the constant C' depends on ϵ and $|A|, |B|, |N(A)|$, and $|N(B)|$.

In fact, Theorem 1.2 follows from a bound on the *truncated m -point correlation functions* (or *semi-invariants*, or *joint cumulants*), a measure of correlation decay that arises naturally in statistical mechanics (see e.g. [25, 7]).

For $A = \{v_1, \dots, v_k\} \subseteq V(G)$, the *truncated correlation function* of A is defined as

$$\kappa(A) = \frac{\partial}{\partial t_{v_1}} \cdots \frac{\partial}{\partial t_{v_k}} \log \mathbb{E} \left[e^{\sum_v t_v X_v} \right] \Big|_{t_v=0, v \in V(G)},$$

where the expectation is over the choice of random independent set from μ_G .

In particular, we can recover the vertex marginals and the correlation between two vertices, as one can show:

$$\kappa(\{v\}) = \mu_v,$$

and

$$\kappa(\{u, v\}) = \mu_{uv} - \mu_u\mu_v.$$

For a set of vertices $A \subseteq R$, let $\text{MST}(A)$ (minimum size Steiner tree) be the smallest k so that G contains a connected subgraph H with k edges containing all vertices of A .

THEOREM 1.3. *If $\Delta_L, \delta_R, \Delta_R, \lambda_L, \lambda_R$ satisfy (1.1), there exists $\epsilon > 0$ so that for any $A \subseteq R$,*

$$|\kappa(A)| \leq Ce^{-\epsilon \text{MST}(A)},$$

where C depends only on ϵ and $|A|$.

Theorem 1.3 immediately implies the bound (1.2) in Theorem 1.2 for $u, v \in R$ since $\text{MST}(\{u, v\}) = D(u, v)$. We show how to derive the full conclusion of Theorem 1.2 from Theorem 1.3 in Section 3.

We remark that these correlation decay results (and the preceding algorithmic results) cannot be obtained by showing that the hard-core model on the infinite bi-regular tree has a unique infinite volume Gibbs measure: for a wide range of parameters, the infinite tree exhibits phase coexistence while we are still able to obtain unconditional correlation decay on all finite graphs. For example, consider the case of biregular graphs with $\lambda = 1$. Our results give exponential decay of correlations and efficient counting and sampling algorithms when $\Delta_R \geq 7\Delta_L \log \Delta_L$. In contrast, a simple fixed point calculation (i.e. the method of [19]) shows that on the infinite biregular tree for large Δ_L the uniqueness threshold is $\Delta_R = (1 + o(1)) \frac{2}{e} 2^{\Delta_L}$ (with phase coexistence below this threshold and uniqueness above the threshold). Similarly, for the Δ -regular case, we prove correlation decay and give efficient algorithms

for $\lambda_L \geq 6\Delta^2\lambda_R$; on the infinite Δ -regular tree, the uniqueness threshold is at

$$\lambda_L = (1 + o(1)) \frac{1 + \lambda_R}{e\lambda_R\Delta^2} (1 + \lambda_R)^\Delta.$$

In both cases there is an exponential gap between the threshold at which we obtain correlation decay and the tree uniqueness threshold. This gap is due to the fact that in constructing a Gibbs measure on the infinite tree we can use arbitrary boundary conditions, while the boundary conditions typically encountered on a subset of a finite graph are nicely behaved (see also the recent work [24] exploiting the difference between arbitrary and nice boundary conditions in the context of colorings).

1.3 Complex zeroes of the partition function

The polynomial interpolation method, Barvinok's approach to approximate counting [1, 2], involves deducing convergence of the Taylor series for the logarithm of a univariate partition function from the existence of a zero-free region of the partition function in the complex plane. This convergence, along with the algorithm of Patel and Regts [26] for efficient computation of low-order coefficients of a partition function, leads to an FPTAS in a wide variety of approximate counting problems.

Our approach of truncating the cluster expansion was inspired by and is related to this approach. In fact, convergence of the cluster expansion implies that the partition function does not vanish, and while our algorithmic approach does not require it (and the preceding algorithmic and probabilistic theorems are stated for positive fugacities), the Kotecký-Preiss convergence criteria works naturally with complex fugacities. Thus we can deduce the following result on zeroes of the bivariate hard-core partition function.

THEOREM 1.4. *Suppose $\Lambda_L, \Lambda_R > 0$ and*

$$(1.4) \quad 6\Delta_L\Delta_R\Lambda_R \leq (1 + \Lambda_L)^{\frac{\delta_R}{\Delta_L}}.$$

Then for all $G \in \mathcal{G}(\Delta_L, \delta_R, \Delta_R)$ and all $\lambda_L, \lambda_R \in \mathbb{C}$ satisfying

$$\begin{aligned} |\lambda_R| &\leq \Lambda_R \text{ and} \\ |1 + \lambda_L| &\geq 1 + \Lambda_L, \end{aligned}$$

the bivariate hard-core partition function satisfies

$$Z(G, \lambda_L, \lambda_R) \neq 0.$$

Theorem 1.4 follows from the proof of Theorem 1.1 with λ_L, λ_R replaced with Λ_L, Λ_R ; we give the details in Section 2.

1.4 Discussion The cluster expansion is a *perturbative* technique, based on expressing a partition function in terms of deviations from a simple, easy to understand *ground state*. In high temperature (small λ) regimes, this ground state is the empty independent set. Barvinok's polynomial interpolation method is also a perturbative technique, and interpolating from $\lambda = 0$ is akin to measuring deviations from the empty independent set.

Previous algorithmic applications of the cluster expansion in low temperature (large λ) regimes [15, 16] considered systems with multiple ground states, e.g. the all L or all R occupied independent sets for the hard-core model on a bipartite graph. A necessary first step in these cases is to show that the partition function of the entire system is well approximated by the sum of partition functions representing deviations from each ground state. The asymmetric setting of this paper is in fact a simpler low-temperature case than those previously considered cases. Condition (1.1) ensures sufficient asymmetry in a bipartite graph that the hard-core model can be expressed in terms of deviations from a single ground state; this ground state is dominant enough that $Z(G)$ is well approximated by small deviations from it. This simplifies the general low-temperature argument, as the first approximation step showing $Z(G)$ is well-approximated by the sum of multiple partition functions is not necessary. As such, our algorithms and analysis are perhaps the most pure application of the algorithmic cluster expansion method and present the best opportunity for a deeper understanding of the power and the limits of this method. Our study of correlation decay in terms of joint cumulants is one example of this, where we were able to leverage the additional probabilistic information we get from a convergent cluster expansion to prove correlation decay for a wide range of parameters for which correlation decay methods based on showing uniqueness on an infinite tree cannot succeed.

As another example, it would be interesting to see if Markov chains can be used to efficiently sample for a similar range of parameters as our algorithms do here. While we do not expect that the Glauber dynamics for sampling from the hard-core model are rapidly mixing for all graphs and parameters satisfying (1.1) (in particular, when the infinite tree has phase coexistence Glauber dynamics on a suitably chosen random graph should be slow mixing), nevertheless we do expect that Glauber dynamics can be used to sample efficiently from the model if we carefully choose our starting state.

CONJECTURE 1. *Suppose $G \in \mathcal{G}(\Delta_L, \delta_R, \Delta_R)$ and λ_L, λ_R satisfy (1.1). Let $n = |V(G)|$. Then if we start the Glauber dynamics at the all L occupied independent*

set and run it for $O(n/\epsilon \cdot \log(n/\epsilon))$ steps, the resulting distribution is within ϵ total variation distance of μ_G .

The intuition behind Conjecture 1 is that mixing time can be too strict a notion for the successful use of a Markov chain for sampling – there may be ‘good’ and ‘bad’ starting configurations, and if we know these ahead of time we can start our chain in a good configuration and avoid the bottleneck responsible for a slow mixing time.

2 Convergence of the cluster expansion

Following [15, 16], our algorithms will be based on approximating the partition function of a *polymer model* [14, 20] using the cluster expansion (for a textbook introduction to both polymer models and the cluster expansion see Chapter 5 of [10]).

A polymer model consists of a set $\mathcal{P}$ of abstract objects we call *polymers*. Each polymer γ is equipped with a real or complex-valued weight w_γ , and there is a symmetric compatibility relation on $\mathcal{P}$; we write $\gamma \sim \gamma'$ if γ and γ' are compatible, and $\gamma \approx \gamma'$ if they are incompatible. We require that $\gamma \approx \gamma$ for all $\gamma \in \mathcal{P}$.

The partition function of a polymer model is

$$\Xi(\mathcal{P}) = \sum_{\substack{\Gamma \subseteq \mathcal{P} \\ \text{compatible}}} \prod_{\gamma \in \Gamma} w_\gamma$$

where the sum is over all pairwise compatible collections of polymers (the empty collection contributes 1 to the sum). If the weights w_γ are real and non-negative then we can define a probability measure ν on Ω , with

$$(2.5) \quad \nu(\Gamma) = \frac{\prod_{\gamma \in \Gamma} w_\gamma}{\Xi(\mathcal{P})}.$$

The cluster expansion is an infinite series representation of $\log \Xi(\mathcal{P})$. A *cluster* Γ is an ordered multiset of polymers from $\mathcal{P}$ whose *incompatibility graph* $H(\Gamma)$ (a vertex for every polymer and an edge between each pair of incompatible polymers) is connected. Denote by $\mathcal{C}$ the set of all clusters from $\mathcal{P}$. As a formal power series in the polymer weights, the cluster expansion is

$$\log \Xi(\mathcal{P}) = \sum_{\Gamma \in \mathcal{C}} w(\Gamma),$$

where

$$w(\Gamma) = \phi(H(\Gamma)) \prod_{\gamma \in \Gamma} w_\gamma,$$

and the function $\phi(H)$ is the *Ursell function*

$$\phi(H) = \frac{1}{|V(H)|!} \sum_{\substack{A \subseteq E(H) \\ \text{spanning, connected}}} (-1)^{|A|}.$$

A sufficient condition for the convergence of the cluster expansion is given by a theorem of Kotecký and Preiss [20].

THEOREM 2.1. *Let $a : \mathcal{P} \rightarrow [0, \infty)$ and $b : \mathcal{P} \rightarrow [0, \infty)$ be two given functions and suppose that for all $\gamma \in \mathcal{P}$,*

$$(2.6) \quad \sum_{\gamma' \approx \gamma} |w_{\gamma'}| e^{a(\gamma') + b(\gamma')} \leq a(\gamma),$$

then the cluster expansion converges absolutely, and, moreover, for all $\gamma \in \mathcal{C}$ we have

$$(2.7) \quad \sum_{\substack{\Gamma \in \mathcal{C}: \\ \exists \gamma' \in \Gamma, \gamma' \approx \gamma}} |w(\Gamma)| e^{b(\Gamma)} \leq a(\gamma),$$

where

$$b(\Gamma) = \sum_{\gamma \in \Gamma} b(\gamma).$$

Now we define a polymer model representation of the hard-core model on a bipartite graph and use the Kotecký-Preiss condition to obtain our algorithmic and correlation decay results. We fix bipartite graph G with bipartition (L, R) and respective minimum and maximum degrees δ_L, Δ_L and δ_R, Δ_R . Fix also activities λ_L and λ_R for vertices in L and R respectively.

We define a polymer γ to be a 2-linked subset of R ; that is, $\gamma \subseteq R$ is connected in the graph G^2 in which each vertex in $V(G)$ is joined to all vertices within distance 2 of it in G . Let $\mathcal{P} = \mathcal{P}(G)$ be the set of all polymers of G . Two polymers γ, γ' are compatible if $\gamma \cup \gamma'$ is not 2-linked and incompatible otherwise. In particular, as required, $\gamma \approx \gamma$ for every polymer γ . For each polymer γ , we let $N(\gamma)$ be all vertices in L adjacent to a vertex in γ . We define the *weight* of a polymer γ to be

$$w_\gamma = \frac{\lambda_R^{|\gamma|}}{(1 + \lambda_L)^{|N(\gamma)|}}.$$

Let $\Xi(\mathcal{P})$ be the corresponding polymer model partition function. This polymer model is almost the same as the polymer models used in [17, 22] to approximate the hard-core partition function on random regular bipartite graphs; the difference being that in those cases, two polymer models were defined, one representing L -dominant independent sets and one representing R -dominant independent sets, and the sum of their partition functions was shown to be a good approximation of $Z(G)$. Here in our asymmetric setting, one polymer model suffices, and in fact the polymer model partition function $\Xi(\mathcal{P})$ is, up to scaling, exactly the hard-core partition function.

LEMMA 2.1.

$$(2.8) \quad Z(G) = (1 + \lambda_L)^{|L|} \Xi(\mathcal{P}).$$

Proof. To see this, consider an independent set $I \in \mathcal{I}(G)$, and decompose $I \cap R$ into its maximal 2-linked components, call these $\gamma_1, \dots, \gamma_k$. Each of these is a polymer since it is 2-linked, and they are pairwise compatible since each is maximal. Moreover, the number of vertices in L that are blocked from being in an independent set by a vertex in $I \cap R$ is $\sum_{i=1}^k |N(\gamma_i)|$, and so if we sum over all independent sets I' so that $I' \cap R = I \cap R$, the contribution to $Z(G)$ is $(1 + \lambda_L)^{|L|} \prod_{i=1}^k \frac{\lambda_R^{|\gamma_i|}}{(1 + \lambda_L)^{|N(\gamma_i)|}}$. Summing over all possible sets $I \cap R$, we obtain (2.8). $\square$

Moreover, as the proof of Lemma 2.1 indicates, if we let ν be the probability measure on sets of compatible polymers defined by (2.5), then it is easy to recover a sample $\mathbf{I}$ from μ_G given a sample $\mathbf{\Gamma}$ from ν : include all vertices of the polymers of $\mathbf{\Gamma}$ in $\mathbf{I}$ and in addition, for each vertex $v \in L$ that is not blocked from being in $\mathbf{I}$ by one of these vertices, include it in $\mathbf{I}$ independently with probability $\frac{\lambda}{1 + \lambda}$.

Next we turn to using the cluster expansion algorithmically. We define the size of a cluster Γ of polymers as the sum of their sizes:

$$|\Gamma| = \sum_{\gamma \in \Gamma} |\gamma|.$$

With this definition, we can define a truncated cluster expansion that only sums over clusters up to a certain size which we will use for the approximate counting algorithm.

$$T_m(\mathcal{P}) := \sum_{\substack{\Gamma \in \mathcal{C} \\ |\Gamma| \leq m}} w(\Gamma).$$

We can then ask when $T_m(\mathcal{P})$ is a good additive approximation of $\log \Xi(\mathcal{P})$.

LEMMA 2.2. *Suppose that for some $\eta > 0$ and every $v \in R$,*

$$(2.9) \quad \sum_{\gamma \ni v} |w_\gamma| e^{(1/2 + \eta)|\gamma|} \leq \frac{1}{2(\Delta_R(\Delta_L - 1) + 1)},$$

then the cluster expansion converges absolutely, the cluster weights satisfy

$$(2.10) \quad \sum_{\substack{\Gamma \in \mathcal{C} \\ \Gamma \ni v}} |w(\Gamma)| e^{\eta|\Gamma|} \leq 1,$$

and, in particular,

$$(2.11) \quad |T_m(\mathcal{P}) - \log \Xi(\mathcal{P})| \leq |R| e^{-m\eta}.$$

Proof. We will apply Theorem 2.1 with the functions $a(\gamma) = |\gamma|/2$ and $b(\gamma) = \eta|\gamma|$. Because any polymer

incompatible with γ must include a vertex of γ or a vertex at distance 2 from γ , for our polymer model we see that

$$\sum_{\gamma' \not\sim \gamma} |w_{\gamma'}| e^{a(\gamma') + b(\gamma')} \leq \sum_{\substack{v \in R \\ D(v, \gamma) \leq 2}} \sum_{\gamma' \ni v} |w_{\gamma'}| e^{(1/2 + \eta)|\gamma'|}.$$

The number of vertices in R with $D(v, \gamma) \leq 2$ is at most $(\Delta_R(\Delta_L - 1) + 1)|\gamma|$, and so

$$\begin{aligned} \sum_{\gamma' \not\sim \gamma} |w_{\gamma'}| e^{a(\gamma') + b(\gamma')} &\leq (\Delta_R(\Delta_L - 1) + 1)|\gamma| \max_{v \in R} \sum_{\gamma' \ni v} |w_{\gamma'}| e^{(1/2 + \eta)|\gamma'|} \\ &\leq |\gamma|/2 = a(\gamma). \end{aligned}$$

by (2.9), giving (2.6). Thus by Theorem 2.1 the cluster expansion for $\log \Xi(\mathcal{P})$ converges absolutely.

To obtain the bound (2.10), we augment the set of polymer we consider and use Theorem 2.1 again. We adjoin to $\mathcal{P}$ a polymer γ_v for each $v \in R$, with $w_{\gamma_v} = 0$ and $\gamma_v \approx \gamma$ for each $\gamma \ni v$. We set $a(\gamma_v) = 1$ and $b(\gamma_v) = 0$. To verify the Kotecký-Preiss condition (2.6) still holds for $\mathcal{P}$ with these new polymers added, we must also verify that for all $v \in R$,

$$\sum_{\gamma \ni v} |w_\gamma| e^{(1/2 + \eta)|\gamma|} \leq 1.$$

This is implied by (2.9). Using conclusion (2.7) of Theorem 2.1 for the adjoined polymers, we see that for all $v \in R$,

$$\begin{aligned} \sum_{\substack{\Gamma \in \mathcal{C} \\ \Gamma \ni v}} |w(\Gamma)| e^{\eta|\Gamma|} &\leq \sum_{\substack{\Gamma \in \mathcal{C}; \\ \exists \gamma' \in \Gamma, \gamma' \approx \gamma_v}} |w(\Gamma)| e^{b(\Gamma)} \\ &\leq a(\gamma_v) = 1. \end{aligned}$$

The bound (2.11) follows from (2.10) by observing that

$$(2.12) \quad \sum_{\substack{\Gamma \in \mathcal{C} \\ \Gamma \ni v \\ |\Gamma| \geq m}} |w(\Gamma)| \leq e^{-\eta m} \sum_{\substack{\Gamma \in \mathcal{C} \\ \Gamma \ni v \\ |\Gamma| \geq m}} |w(\Gamma)| e^{\eta|\Gamma|} \leq e^{-\eta m},$$

and then summing over all $v \in R$. $\square$

With this sufficient condition for the convergence of the cluster expansion for $\log \Xi(\mathcal{P})$ we can prove Theorem 1.1.

Proof. [Proof of Theorem 1.1] Let $G \in \mathcal{G}(\Delta_L, \delta_R, \Delta_R)$, and let $n = |R|$. Let λ_L and λ_R be such that (1.1) is satisfied. Let $\mathcal{P} = \mathcal{P}(G)$ be the set of polymers defined

above (without the additional polymers γ_v that were adjoined in the previous proof).

We first consider the approximate counting algorithm. By Lemma 2.1, to approximate $Z(G)$ it suffices to obtain an ϵ -relative approximation to $\Xi(\mathcal{P})$; we can then get an ϵ -relative approximation to $Z(G)$ by multiplying by $(1 + \lambda_L)^{|L|}$.

The bound (2.11) suggests an algorithm for approximating $\Xi(\mathcal{P})$ when (2.9) holds by calculating $T_m(\mathcal{P})$ for $m = \log(n/\epsilon)/\eta$ and exponentiating. This can be done as follows:

1. Enumerate all clusters $\Gamma \in \mathcal{C}$ with $|\Gamma| < m$; call the list of such clusters $\mathcal{C}_m$.
2. For each cluster $\Gamma \in \mathcal{C}_m$, compute $\phi(H(\Gamma))$ and $\prod_{\gamma \in \Gamma} w_\gamma$.
3. Compute $T_m(\mathcal{P})$ by summing:

$$T_m(\mathcal{P}) = \sum_{\Gamma \in \mathcal{C}_m} \phi(H(\Gamma)) \prod_{\gamma \in \Gamma} w_\gamma.$$

4. Output $\exp(T_m(\mathcal{P}))$.

It is shown in [17], using ideas and tools from [15, 26, 3], that this algorithm can be implemented with running time $O(n \cdot (n/\epsilon)^{O(\log(\Delta_L \Delta_R)/\eta)})$, which for Δ_L, Δ_R fixed is polynomial in n and $1/\epsilon$.

What remains is to show that (2.9) holds for some $\eta > 0$. By double counting the edge boundary of a polymer γ , we have that $|N(\gamma)| \geq \frac{\delta_R}{\Delta_L} |\gamma|$, and so

$$|w_\gamma| \leq \left(\frac{\lambda_R}{(1 + \lambda_L)^{\frac{\delta_R}{\Delta_L}}} \right)^{|\gamma|}.$$

Then using the fact from [4] that the number of 2-linked subsets of R containing a fixed vertex v of size k is at most $\frac{(e\Delta_R(\Delta_L - 1))^{k-1}}{k^{3/2}}$, we have

$$\begin{aligned} & \sum_{\gamma \ni v} |w_\gamma| e^{(1/2+\eta)|\gamma|} \\ & \leq \sum_{k=1}^{\infty} \frac{(e\Delta_R(\Delta_L - 1))^{k-1}}{k^{3/2}} \left(\frac{\lambda_R}{(1 + \lambda_L)^{\frac{\delta_R}{\Delta_L}}} \right)^k e^{(1/2+\eta)k} \\ & \leq \sum_{k=1}^{\infty} \frac{(e(\Delta_R(\Delta_L - 1) + 1))^{k-1}}{k^{3/2}} \cdot \left(\frac{\lambda_R}{(1 + \lambda_L)^{\frac{\delta_R}{\Delta_L}}} \right)^k e^{(1/2+\eta)k} \\ & = \frac{1}{e(\Delta_R(\Delta_L - 1) + 1)} \\ & \quad \cdot \sum_{k=1}^{\infty} \frac{1}{k^{3/2}} \left(\frac{(\Delta_R(\Delta_L - 1) + 1)\lambda_R e^{3/2+\eta}}{(1 + \lambda_L)^{\frac{\delta_R}{\Delta_L}}} \right)^k. \end{aligned}$$

Thus it suffices to show that

$$\sum_{k=1}^{\infty} \frac{1}{k^{3/2}} \left(\frac{(\Delta_R(\Delta_L - 1) + 1)\lambda_R e^{3/2+\eta}}{(1 + \lambda_L)^{\frac{\delta_R}{\Delta_L}}} \right)^k \leq \frac{e}{2}.$$

Since $\sum_{k \geq 1} s^k/k^{3/2} < e/2$ for $0 \leq s \leq .832$, it is enough to show that

$$\frac{(\Delta_R(\Delta_L - 1) + 1)\lambda_R e^{3/2+\eta}}{(1 + \lambda_L)^{\frac{\delta_R}{\Delta_L}}} \leq .832,$$

or

$$\frac{(\Delta_R(\Delta_L - 1) + 1)\lambda_R e^\eta}{(1 + \lambda_L)^{\frac{\delta_R}{\Delta_L}}} \leq \frac{.832}{e^{3/2}} \approx .1856.$$

In particular, if (1.1) holds then

$$\frac{(\Delta_R(\Delta_L - 1) + 1)\lambda_R}{(1 + \lambda_L)^{\frac{\delta_R}{\Delta_L}}} \leq \frac{1}{6},$$

and so we can obtain (2.9) with $\eta = .1 < \log(6 \cdot .1856)$. Applying the algorithm of [17] gives the FPTAS for $Z(G)$.

Next we turn to approximate sampling. The usual approach to using an approximate counting algorithm to obtain an approximate sampling algorithm is via self-reducibility [18]. We cannot directly apply this in our setting, however, as reducing a graph $G \in \mathcal{G}(\Delta_L, \delta_R, \Delta_R)$ might result in a graph $G' \notin \mathcal{G}(\Delta_L, \delta_R, \Delta_R)$ because of the minimum degree condition. Instead we can apply self-reducibility on the level of polymers, using [15, Theorem 5.1], which provides a general reduction of approximate sampling to approximate counting for polymer models. The high-level idea of the algorithm, applied to our setting, is to create a compatible polymer configuration Γ one polymer at a time, with a final distribution close to ν given in (2.5). Once we have such a configuration Γ we can extend this to an independent set $I \in \mathcal{I}(G)$ as described in the remark after Lemma 2.1.

To approximately sample from ν , we order the vertices in R arbitrarily, and then one-by-one we determine which, if any, polymer containing a given vertex v is present in Γ . Based on the previous choices, some polymers in $\mathcal{P}$ are excluded from future choices as their addition would form an incompatible pair of polymers, and so the calculation of marginal probabilities involves approximating $\Xi(\mathcal{P}')$, the partition function of a set of polymers $\mathcal{P}' \subset \mathcal{P}$. While the polymer model associated to an arbitrary set $\mathcal{P}' \subset \mathcal{P}$ may not map to the hard-core model on any graph, this polymer model automatically satisfies (2.6) if the original polymer model does, since the condition only becomes weaker on removing polymers. Thus we can efficiently approximate

$\Xi(\mathcal{P}')$ using the algorithm from [17], and thus sample efficiently from ν . $\square$

Next we prove Corollary 1.1.

Proof. [Proof of Corollary 1.1] Parts (1) and (2) follow from simple substitutions into (1.1).

For (3), we set $\Delta_L = \Delta$ and $\Delta_R = c\Delta \log(\Delta)$, and ask for which c we have

$$(2.13) \quad c \log 2 \log \Delta - 2 \log \Delta - \log \log \Delta - \log c - \log 6 \geq 0,$$

which implies (1.1) is satisfied. Some simple calculus shows that (2.13) is satisfied for $\Delta \geq 2$ when $c \geq 11$. But if we use the fact that for $\Delta \leq 5$, an FPTAS is given by Liu and Lu [23], then we see that (2.13) is satisfied for $\Delta \geq 6$ when $c \geq 7$. $\square$

Finally we prove Theorem 1.4, which follows almost immediately from the proof of Theorem 1.1.

Proof. [Proof of Theorem 1.4] Condition (1.4) is the same as Condition (1.1) with λ_L and λ_R replaced with Λ_L and Λ_R . If λ_L and λ_R are complex such that $|\lambda_R| \leq \Lambda_R$ and $|1 + \lambda_L| \geq 1 + \Lambda_L$, then for any polymer γ ,

$$|w_\gamma| = \left| \frac{\lambda_R^{|\gamma|}}{(1 + \lambda_L)^{|N(\gamma)|}} \right| \leq \left(\frac{\Lambda_R}{(1 + \Lambda_L)^{\frac{\delta_R}{\Delta_L}}} \right)^{|\gamma|}.$$

Condition (1.4) for Λ_L and Λ_R then suffices for the analysis in the proof of Theorem 1.1 to show Condition (2.9) holds for λ_L and λ_R . Lemma 2.2 then implies the cluster expansion for λ_L and λ_R converges absolutely. Because the cluster expansion is a convergent power series for $\log \Xi(\mathcal{P})$, for these λ_L and λ_R the polymer partition function $\Xi(\mathcal{P})$ cannot be zero and so $Z(G) = (1 + \lambda_L)^{|L|} \Xi(\mathcal{P})$ is also nonzero. $\square$

3 Correlation decay

As Dobrushin showed [6, 8] the cluster expansion and the Kotecký-Preiss convergence condition are very well suited to studying the cumulants and joint cumulants of random variables that are functions of polymer configurations.

It follows almost directly from estimates such as (2.10) that the magnitudes of joint cumulants of random variables that depend on disjoint polymers decay exponentially in the distance between these polymers (Theorem 1.3). The joint cumulant of a set of random variables vanishes if there is a non-trivial partition of the set into two sets of random variables that are independent of each other, and this holds approximately

as well. Showing joint cumulants decay exponentially in the distance between polymers shows a form of decay of correlations, and in fact implies the perhaps more familiar form of (1.3).

While the techniques and calculations described here are very similar to those in [6], and the properties of joint cumulants are standard facts in the study of statistical mechanics on lattices, we present the proofs in a self-contained way in order to emphasize their elementary nature and to encourage the use of these methods in computer science and in non-lattice settings.

We again fix a graph G and consider the associated polymer model with polymers $\mathcal{P}$ and partition function Ξ . We suppose that (2.9) holds for some $\eta > 0$ and calculate exact expressions for the joint cumulants in our setting using the cluster expansion. We begin by introducing auxiliary polymer weights as follows. Given variables t_v , $v \in R$, let

$$\tilde{w}_\gamma = w_\gamma e^{\sum_{v \in \gamma} t_v},$$

and let $\tilde{\Xi}$ be the polymer model partition function derived from these weights. Let $X_v = \mathbf{1}_{v \in \mathbf{I}}$ for $\mathbf{I}$ drawn from μ_G and recall that the distribution of $\mathbf{I} \cap R$ (the set of occupied vertices in R) is identical to the distribution of $\bigcup_{\gamma \in \Gamma} \gamma$ (the union of polymers) for a polymer configuration Γ drawn from ν . Using this we can write

$$\begin{aligned} \mathbb{E} e^{\sum_{v \in R} t_v X_v} &= \sum_{\substack{\Gamma \subseteq \mathcal{P} \\ \text{compatible}}} \nu(\Gamma) \prod_{\gamma \in \Gamma} e^{\sum_{v \in \gamma} t_v} \\ &= \sum_{\substack{\Gamma \subseteq \mathcal{P} \\ \text{compatible}}} \frac{1}{\tilde{\Xi}} \prod_{\gamma \in \Gamma} w_\gamma e^{\sum_{v \in \gamma} t_v} \\ &= \frac{\tilde{\Xi}}{\Xi}. \end{aligned}$$

For a set $A \subseteq R$ we can then write the joint cumulant of A as

$$\begin{aligned} \kappa(A) &:= \frac{\partial^{|A|} \log \mathbb{E} e^{\sum_{v \in R} t_v X_v}}{\prod_{u \in A} \partial t_u} \Bigg|_{t_v=0, v \in V(G)} \\ &= \frac{\partial^{|A|} \log \tilde{\Xi}}{\prod_{u \in A} \partial t_u} \Bigg|_{t_v=0, v \in V(G)}. \end{aligned}$$

For a vertex $v \in R$, and a cluster Γ , define $Y_v(\Gamma)$ to be the number of polymers in Γ containing v :

$$Y_v(\Gamma) = \sum_{\gamma \in \Gamma} \mathbf{1}_{v \in \gamma}.$$

Then using the cluster expansion for $\log \tilde{\Xi}$ (and the fact

that it converges absolutely if $t_v \leq \eta$ for all v), we have

$$(3.14) \quad \begin{aligned} \kappa(A) &= \frac{\partial^{|A|}}{\prod_{u \in A} \partial t_u} \sum_{\Gamma \in \mathcal{C}} w(\Gamma) \Big|_{t_v=0, v \in V(G)} \\ &= \sum_{\Gamma \in \mathcal{C}} w(\Gamma) \prod_{v \in A} Y_v(\Gamma). \end{aligned}$$

From this expression we can immediately derive Theorem 1.3.

Proof. [Proof of Theorem 1.3] We will show that for $A \subseteq R$,

$$(3.15) \quad \sum_{\Gamma \in \mathcal{C}} |w(\Gamma)| \prod_{v \in A} Y_v(\Gamma) \leq C e^{-\eta \text{MST}(A)/2}$$

where C depends on η and $|A|$. The size of a cluster Γ containing all vertices of A is at least $\text{MST}(A)/2$, where $\text{MST}(A)$ is the number of edges of the minimum Steiner tree in G containing A ; this Steiner tree will contain $\text{MST}(A) + 1$ vertices, and at least half of these vertices will be in R . More generally suppose a cluster Γ contains $Y_v(\Gamma) \geq 1$ copies of v for each $v \in A$. Then

$$(3.16) \quad |\Gamma| \geq \text{MST}(A)/2 + \sum_{v \in A} (Y_v(\Gamma) - 1)$$

since each additional copy of a vertex contributes 1 to the size of Γ .

We arrange the sum on the LHS of (3.15) based on the number of copies of each $v \in A$:

$$\begin{aligned} &\sum_{\Gamma \in \mathcal{C}} |w(\Gamma)| \prod_{v \in A} Y_v(\Gamma) \\ &= \sum_{\{y_v \geq 1\}_{v \in A}} \sum_{\substack{\Gamma \in \mathcal{C} \\ Y_v(\Gamma) = y_v \forall v \in A}} |w(\Gamma)| \prod_{v \in A} y_v. \end{aligned}$$

Now since Condition (2.9) holds with constant $\eta > 0$, we can use (2.12) and (3.16):

$$\begin{aligned} &\sum_{\Gamma \in \mathcal{C}} |w(\Gamma)| \prod_{v \in A} Y_v(\Gamma) \\ &\leq \sum_{\{y_v \geq 1\}_{v \in A}} e^{-\eta(\text{MST}(A)/2 + \sum_{v \in A} (y_v - 1))} \prod_{v \in A} y_v \\ &\leq e^{-\eta \text{MST}(A)/2} \sum_{\{y_v \geq 1\}_{v \in A}} e^{-\eta \sum_{v \in A} (y_v - 1)} \prod_{v \in A} y_v \\ &\leq C e^{-\eta \text{MST}(A)/2}, \end{aligned}$$

and so the conclusion of Theorem 1.3 holds with $\epsilon = \eta/2$ and

$$C = \sum_{\{y_v \geq 1\}_{v \in A}} e^{-\eta \sum_{v \in A} (y_v - 1)} \prod_{v \in A} y_v < \infty.$$

□

To prove Theorem 1.2 we need to do two things. First we need to show that the exponential decay of joint cumulants in the minimum Steiner tree size implies that $|\mu_{A \cup B} - \mu_A \mu_B|$ decays exponentially in the distance from A to B ; second we need to extend these results from sets of vertices contained solely in R to general sets of vertices in $L \cup R$.

LEMMA 3.1. *Suppose $A \subset R$ and $B \subset R$ are disjoint. Suppose also that for any $S \subseteq A \cup B$ so that $S \cap A \neq \emptyset$ and $S \cap B \neq \emptyset$, we have*

$$|\kappa(S)| \leq \epsilon.$$

Then

$$|\mu_{A \cup B} - \mu_A \mu_B| \leq C \epsilon$$

where C depends only on $|A|$ and $|B|$.

Proof. We will use the following formula (see e.g. [21]):

$$\mu_A = \sum_{\pi \in \Sigma(A)} \prod_{S \in \pi} \kappa(S),$$

where $\Sigma(A)$ is the family of all set partitions of A . This gives

$$\begin{aligned} \mu_{A \cup B} - \mu_A \mu_B &= \left(\sum_{\pi \in \Sigma(A \cup B)} \prod_{S \in \pi} \kappa(S) \right) \\ &\quad - \left(\sum_{\pi \in \Sigma(A)} \prod_{S \in \pi} \kappa(S) \right) \left(\sum_{\pi \in \Sigma(B)} \prod_{S \in \pi} \kappa(S) \right). \end{aligned}$$

Now if we restrict the first sum on the RHS to those partitions whose parts are either entirely within A or entirely within B , then we obtain exactly $\mu_A \mu_B$ which cancels the second term on the RHS, and so we obtain (for $|\pi|$ the number of sets in partition π):

$$\mu_{A \cup B} - \mu_A \mu_B = \sum_{\substack{\pi \in \Sigma(A \cup B) \\ \exists S' \in \pi: S' \cap A \neq \emptyset, S' \cap B \neq \emptyset}} \prod_{S \in \pi} \kappa(S).$$

Now since $|\kappa(S')| \leq \epsilon$ for S' that intersects both A and B , and the partitions in the above sum have at least one such S' , we have

$$|\mu_{A \cup B} - \mu_A \mu_B| \leq C \epsilon$$

where C is the number of set partitions of $A \cup B$, times the maximum size of a joint cumulant of at most $|A| + |B|$ indicator random variables raised to the $|A| + |B|$ power, which depends only on $|A|$ and $|B|$ [21]. □

With these ingredients we prove Theorem 1.2.

Proof. [Proof of Theorem 1.2] It suffices to prove the second statement in the theorem as the first follows by taking $A = \{u\}, B = \{v\}$. First we consider the special case $A \subset R, B \subset R$.

For any $S \subseteq R$ that intersects both A and B , we have $\text{MST}(S) \geq D(A, B)$. Then from Theorem 1.3 we have $|\kappa(S)| \leq Ce^{-\eta D(A, B)/2}$ for such S . Then applying Lemma 3.1, we obtain

$$|\mu_{A \cup B} - \mu_A \mu_B| \leq C' e^{-\eta D(A, B)/2}.$$

Now we consider the general case $A \subset V(G), B \subset V(G)$. We assume that $D(A, B) > 2$ as otherwise the statement is trivial. We can also assume that both A and B are independent sets in G , as otherwise both $\mu_{A \cup B}$ and $\mu_A \mu_B$ are 0. We say a vertex v is *unblocked* by the independent set I if $N(v) \cap I = \emptyset$. Let $\hat{\mu}_v = \Pr[v \text{ unblocked}]$, and for a set of vertices S , $\hat{\mu}_S = \Pr[S \text{ unblocked}]$. Then if S is an independent set itself,

$$\mu_S = \left(\frac{\lambda}{1 + \lambda} \right)^{|S|} \hat{\mu}_S.$$

To see this, note that S must be unblocked to be in the independent set. Given that S is unblocked, since S is independent itself, the probability that each vertex $v \in S$ is in the independent set is $\frac{\lambda}{1 + \lambda}$ and these events are conditionally independent over the vertices in S .

For $S \subset V(G)$, we introduce the notation $\rho_S = \Pr[S \cap \mathbf{I} = \emptyset]$. If $S \subset L$, by definition, $\hat{\mu}_S = \rho_{N(S)}$. Using inclusion exclusion, we have, for $T \subset R$,

$$\rho_T = \sum_{Q \subseteq T} (-1)^{|Q|} \mu_Q,$$

where the sum includes the empty set and $\mu_\emptyset = 1$.

Now suppose $A \subset V(G), B \subset V(G)$, and let $A_L = A \cap L, A_R = A \cap R, B_L = B \cap L, B_R = B \cap R$. Then, since A is an independent set in G , we can write

$$\begin{aligned} \mu_A &= \left(\frac{\lambda}{1 + \lambda} \right)^{|A_L|} \Pr[A_R \subseteq \mathbf{I} \wedge N(A_L) \cap \mathbf{I} = \emptyset] \\ &= \left(\frac{\lambda}{1 + \lambda} \right)^{|A_L|} \sum_{Q' \subseteq N(A_L)} (-1)^{|Q'|} \mu_{Q' \cup A_R}. \end{aligned}$$

We can write similar formulae for μ_B and $\mu_{A \cup B}$. Considering $\mu_A \mu_B$, and using the fact that $N(A_L) \cap N(B_L) = \emptyset$ since $D(A, B) > 2$, we can simplify

$$\begin{aligned} &\left(\sum_{Q \subseteq N(A_L)} (-1)^{|Q|} \mu_{Q \cup A_R} \right) \left(\sum_{Q' \subseteq N(B_L)} (-1)^{|Q'|} \mu_{Q' \cup B_R} \right) \\ &= \sum_{Q \subseteq N(A_L), Q' \subseteq N(B_L)} (-1)^{|Q \cup Q'|} \mu_{Q \cup A_R} \mu_{Q' \cup B_R}. \end{aligned}$$

Therefore we have

$$\begin{aligned} \mu_{A \cup B} - \mu_A \mu_B &= \left(\frac{\lambda}{1 + \lambda} \right)^{|A_L| + |B_L|} \sum_{\substack{Q \subseteq N(A_L) \\ Q' \subseteq N(B_L)}} [\mu_{Q \cup A_R \cup Q' \cup B_R} - \mu_{Q \cup A_R} \mu_{Q' \cup B_R}], \end{aligned}$$

and so, as $\lambda/(1 + \lambda) < 1$,

$$|\mu_{A \cup B} - \mu_A \mu_B| \leq 2^{|N(A_L)| + |N(B_L)|} \max_{\substack{Q \subseteq N(A_L) \\ Q' \subseteq N(B_L)}} |\mu_{Q \cup A_R \cup Q' \cup B_R} - \mu_{Q \cup A_R} \mu_{Q' \cup B_R}|.$$

The first factor is bounded by a constant that depends only on $|N(A)|, |N(B)|$. Now let $\hat{A} = Q \cup A_R$ and $\hat{B} = Q' \cup B_R$. Then $\hat{A} \subset R$ and $\hat{B} \subset R$, and moreover $D(\hat{A}, \hat{B}) \geq D(A, B) - 2$, and so we can apply the special case above to obtain

$$|\mu_{Q \cup A_R \cup Q' \cup B_R} - \mu_{Q \cup A_R} \mu_{Q' \cup B_R}| \leq C' e^{-\eta D(A, B)/2 + 1},$$

and combining these bounds we get

$$|\mu_{A \cup B} - \mu_A \mu_B| \leq C e^{-\eta D(A, B)/2}$$

for a constant C that depends only on $\eta, |A|, |B|, |N(A)|, |N(B)|$. Setting $\epsilon = \eta/2$ we obtain Theorem 1.2. $\square$

Acknowledgements

The authors thank Guus Regts, Piyush Srivastava, Prasad Tetali, and Matthew Jenssen for many helpful conversations. This work was done while the authors were participating in the Geometry of Polynomials program at the Simons Institute for the Theory of Computing.

References

- [1] A. Barvinok. Combinatorics and complexity of partition functions. *Algorithms and Combinatorics*, 30, 2017.
- [2] A. Barvinok and G. Regts. Weighted counting of solutions to sparse systems of equations. *Combinatorics, Probability and Computing*, 28(5):696–719, 2019.
- [3] A. Björklund, T. Husfeldt, P. Kaski, and M. Koivisto. Computing the Tutte polynomial in vertex-exponential time. In *Proceedings of the Forty-ninth Annual Symposium on Foundations of Computer Science*, FOCS 2008, pages 677–686. IEEE, 2008.
- [4] C. Borgs, J. Chayes, J. Kahn, and L. Lovász. Left and right convergence of graphs with bounded degree. *Random Structures & Algorithms*, 42(1):1–28, 2013.

- [5] J.-Y. Cai, A. Galanis, L. A. Goldberg, H. Guo, M. Jerrum, D. Štefankovič, and E. Vigoda. #BIS-hardness for 2-spin systems on bipartite bounded degree graphs in the tree non-uniqueness region. *Journal of Computer and System Sciences*, 82(5):690–711, 2016.
- [6] R. Dobrushin. Estimates of semi-invariants for the Ising model at low temperatures. *Translations of the American Mathematical Society-Series 2*, 177:59–82, 1996.
- [7] R. Dobrushin and S. Shlosman. Completely analytical interactions: constructive description. *Journal of Statistical Physics*, 46(5-6):983–1014, 1987.
- [8] R. L. Dobrushin. Perturbation methods of the theory of Gibbsian fields. In *Lectures on probability theory and statistics*, pages 1–66. Springer, 1996.
- [9] M. Dyer, L. A. Goldberg, C. Greenhill, and M. Jerrum. The relative complexity of approximate counting problems. *Algorithmica*, 38(3):471–500, 2004.
- [10] S. Friedli and Y. Velenik. *Statistical mechanics of lattice systems: a concrete mathematical introduction*. Cambridge University Press, 2017.
- [11] A. Galanis, Q. Ge, D. Štefankovič, E. Vigoda, and L. Yang. Improved inapproximability results for counting independent sets in the hard-core model. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 567–578. Springer, 2011.
- [12] A. Galanis, D. Stefankovic, E. Vigoda, and L. Yang. Ferromagnetic Potts model: Refined #BIS-hardness and related results. *SIAM Journal on Computing*, 45(6):2004–2065, 2016.
- [13] L. A. Goldberg and M. Jerrum. Approximating the partition function of the ferromagnetic Potts model. *Journal of the ACM*, 59(5):25, 2012.
- [14] C. Gruber and H. Kunz. General properties of polymer systems. *Communications in Mathematical Physics*, 22(2):133–161, 1971.
- [15] T. Helmuth, W. Perkins, and G. Regts. Algorithmic Pirogov-Sinai theory. *Probability Theory and Related Fields*, 2019.
- [16] M. Jenssen, P. Keevash, and W. Perkins. Algorithms for #BIS-hard problems on expander graphs. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2235–2247. SIAM, 2019.
- [17] M. Jenssen, P. Keevash, and W. Perkins. Algorithms for #BIS-hard problems on expander graphs. *arXiv preprint*, 1807.04804v2, 2019.
- [18] M. R. Jerrum, L. G. Valiant, and V. V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43:169–188, 1986.
- [19] F. Kelly. Loss networks. *Annals of Applied Probability*, 1:319–378, 1991.
- [20] R. Kotecký and D. Preiss. Cluster expansion for abstract polymer models. *Communications in Mathematical Physics*, 103(3):491–498, 1986.
- [21] V. Leonov and A. N. Shiryaev. On a method of calculation of semi-invariants. *Theory of Probability & its applications*, 4(3):319–329, 1959.
- [22] C. Liao, J. Lin, P. Lu, and Z. Mao. Counting independent sets and colorings on random regular bipartite graphs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, RANDOM 2019, pages 34:1–34:12. LIPICS, 2019.
- [23] J. Liu and P. Lu. FPTAS for #BIS with degree bounds on one side. In *Proceedings of the Forty-seventh Annual ACM Symposium on Theory of Computing*, STOC 2015, pages 549–556. ACM, 2015.
- [24] J. Liu, A. Sinclair, and P. Srivastava. A deterministic algorithm for counting colorings with 2Δ colors. In *Proceedings of the Sixtieth Annual IEEE Symposium on Foundations of Computer Science*, FOCS 2019. IEEE, 2019.
- [25] V. A. Malyshev. Cluster expansions in lattice models of statistical physics and the quantum theory of fields. *Russian Mathematical Surveys*, 35(2):1, 1980.
- [26] V. Patel and G. Regts. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM Journal on Computing*, 46(6):1893–1919, 2017.
- [27] A. Sly. Computational transition at the uniqueness threshold. In *Proceedings of the Fifty-first Annual IEEE Symposium on Foundations of Computer Science*, FOCS 2010, pages 287–296. IEEE, 2010.
- [28] A. Sly and N. Sun. Counting in two-spin models on d-regular graphs. *The Annals of Probability*, 42(6):2383–2416, 2014.
- [29] D. Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing*, STOC 2006, pages 140–149. ACM, 2006.