Lecture 11: The Polymer Method and the Cluster Expansion

October 17, 2023

The primary theme of this lecture will be *embedding* other statistical mechanics models into general hardcore models. These representation results are achieved through the *(abstract) polymer method*. One beautiful aspect of this idea, in some sense, is it allows us to treat a large swathe of counting problems in a unified manner. While this often costs us a significant "blow-up in the size of the model representation", we'll nonetheless see how to develop efficient algorithms based on very similar ideas in Barvinok's polynomial interpolation algorithm.

In the specific setting of (multivariate) independence polynomials, this idea of using low-degree Taylor approximations to the log-partition function goes under the name *cluster expansion* (or *Mayer expansion*); see [SS05] and references therein. We'll see new applications to approximate counting in systems at *low temperatures*, i.e. with *strong interactions*. This is a very challenging setting which defeats standard local Markov chains like Glauber dynamics, and where standard notions of correlation decay (e.g. weak/strong spatial mixing) fail dramatically. The algorithmic results we will discuss are largely based on [JKP20] (see also [HPR20]).

1 Universality of the Hardcore Model

1.1 Abstract Polymer Models

We start by defining abstract polymer models at high generality. We emphasize that, at least for the purposes of this lecture, they are *exactly the same* as hardcore models but with possibly heterogeneous fugacities. However, we use the nomenclature of polymer models to be consistent with the literature.[1] One can incorporate softcore interactions but we won't do so here; see the Cluster Expansion chapter in [FV17] for a more general discussion.

Definition 1 ((Abstract) Polymer Model)**.** *An* (abstract) polymer model *is a triple* $(\mathcal{C}, \sim, \boldsymbol{w} : \mathcal{C} \to \mathbb{C})$*, where* $\mathcal{C}$ *is a finite set of elements called* polymers*,* $\sim$ *is a symmetric relation on polymers, and* $\boldsymbol{w}$ *is a collection of complex-valued weights. If two polymers* γ, γ' *are compatible, then we write* $\gamma \sim \gamma'$*; we write* $\gamma \not\sim \gamma'$ *otherwise. We further impose that all polymers are incompatible with themselves, i.e.* $\gamma \not\sim \gamma$ *for all* $\gamma \in \mathcal{C}$*. The associated* polymer partition function *is given by*

$$\Xi_{\mathcal{C}} = \Xi_{\mathcal{C}}(\boldsymbol{w}) \stackrel{\text{def}}{=} \sum_{\Gamma} \prod_{\gamma \in \Gamma} w_{\gamma}, \tag{1}$$

where the summation is over all collections of mutually compatible polymers.

Any abstract polymer model is a hardcore model w.r.t. the associated *incompatibility graph* on $(\mathcal{C}, \not\sim)$, where there is an edge between γ, γ' if and only if $\gamma \not\sim \gamma'$. Each Γ in Eq. (1) is just an independent set in this incompatibility graph, and the function $\Xi_{\mathcal{C}}(\boldsymbol{w})$ is then just the associated multivariate independence polynomial evaluated at $\boldsymbol{w}$.

However, unlike our previous discussion of the hardcore model on bounded-degree graphs, the polymers γ in this lecture will often have concrete combinatorial meaning as connected subgraphs (U, F) of some *host graph* $G = (V, E)$. We will then say two subgraphs $\gamma = (U, F), \gamma' = (U', F')$ are compatible if, for example, they are vertex-disjoint or $d(U, U') > 1$ where $d(\cdot, \cdot)$ denotes the shortest path distance in G. These types of compatibility relations preclude possibly nasty "double

[1] The term "polymer model" is a bit overloaded in statistical mechanics. Dobrushin [Dob96] tried to remedy this by renaming these as *animal models*, but the term polymer model stuck.

counting issues" where e.g. the union of two smaller compatible polymers is itself a polymer. These types of polymer models are sometimes called *subgraph polymer models*, and in general have size exponential in the size of the host graph $G = (V, E)$. Despite this blow-up in the representation size, they are nonetheless useful gadgets in both probabilistic analyses as well as approximate counting.

The rest of this section will be devoted to nontrivial examples which hopefully illustrate how they can arise in a multitude of settings seemingly unrelated to independent sets. We emphasize that any given "base model" can admit many genuinely distinct polymer representations, each of which may illuminate a different aspect of the "base model". In particular, there is some creative freedom in designing good polymers. For instance, we will see three very different polymer models which capture the ferromagnetic Ising model. See [HPR20] for more sophisticated examples called *contour models*, which take advantage of the additional geometric structure available when $G = \mathbb{Z}^d$. An example of bipartite independent sets is discussed in Section 3 in greater depth.

1.2 Even Subgraphs Representation of the Ferromagnetic Ising Model

Let $G = (V, E)$ be a host graph, and consider the ferromagnetic Ising model on G with inverse temperature $\beta \geq 0$ and uniform external field $h \in \mathbb{R}$. Now define an abstract polymer model $(\mathcal{C}, \sim, \boldsymbol{w} : \mathcal{C} \to \mathbb{R}_{\geq 0})$ via the following data:

- For a subset of edges $F \subseteq E$, let $V_F = \{v : \deg_F(v) > 0\}$ denote the set of vertices participating in some edge of F. A polymer γ is a subgraph (V, F) such that (V_F, F) is connected.
- Two polymers $\gamma = (V, F), \gamma' = (V, F')$ are compatible if $V_F \cap V_{F'} = \emptyset$.
- The weight of a polymer is given by $w_\gamma \stackrel{\text{def}}{=} \rho^{|\mathsf{odd}(F)|}\lambda^{|F|}$ where $\rho = \tanh(h), \lambda = \tanh(\beta)$.

Since $w_{\gamma \sqcup \gamma'} = w_\gamma \cdot w_{\gamma'}$ for any pair of incompatible polymers (where the disjoint union is applied to the edge sets of γ, γ'), and every subgraph (V, F) of G can be uniquely decomposed as a disjoint union over its maximal connected components, the polymer partition function $\Xi_{\mathcal{C}}(\boldsymbol{w})$ is precisely

$$\widehat{Z}_G^{\mathsf{even}}(\rho, \lambda) = \sum_{F \subseteq E} \rho^{|\mathsf{odd}(F)|}\lambda^{|F|},$$

the even subgraphs partition function. We already previously saw that $\widehat{Z}_G^{\mathsf{even}}(\rho, \lambda) = C(\beta, h) \cdot Z_G^{\mathsf{Ising}}$, where $C(\beta, h) = 2^{|V|}\cosh(h)^{|V|}\cosh(\beta)^{|E|}$ is an easy-to-compute constant.

1.3 Low-Temperature Ferromagnetic Potts Model

Let $G = (V, E)$ be a graph, $\beta \geq 0$ be an inverse temperature parameter, and $q \in \mathbb{N}$ be a number of colors. For an assignment $\sigma : V \to [q]$, write $m(\sigma) = \#\{e = uv \in E : \sigma(u) = \sigma(v)\}$ for the number of monochromatic edges in G under σ. The Gibbs distribution of the *(q-state) ferromagnetic Potts model* is given by

$$\mu_{G,q,\beta}^{\mathsf{Potts}}(\sigma) \propto \exp\left(\beta \cdot m(\sigma)\right), \qquad \forall \sigma \in [q]^V, \tag{2}$$

with partition function

$$Z_{G,q,\beta}^{\mathsf{Potts}} \stackrel{\text{def}}{=} \sum_{\sigma : V \to [q]} \exp\left(\beta \cdot m(\sigma)\right).$$

The special case $q = 2$ recovers the ferromagnetic Ising model. $\mu_{G,q,\beta}^{\mathsf{Potts}}$ is *ferromagnetic* in the sense that neighboring vertices are encouraged to take on the same color. In particular, when β is large, the distribution $\mu_{G,q,\beta}^{\mathsf{Potts}}$ concentrates on the q trivial *ground states* given by assigning all vertices the same color.

Following [JKP20], we build a polymer model with captures the *deviations* from an arbitrarily chosen ground state. Define an abstract polymer model $(\mathcal{C}, \sim, \boldsymbol{w} : \mathcal{C} \to \mathbb{R}_{\geq 0})$ via the following data:

- A polymer γ is a subset of vertices V such that the *induced* subgraph $G[\gamma]$ is connected.

- We say two polymers γ, γ' are compatible if $d(\gamma, \gamma') > 1$ w.r.t. the shortest path metric.
- The weight of a polymer $\gamma \subseteq V$ is given by

$$w_\gamma \stackrel{\text{def}}{=} \exp\left(-\beta \cdot |E(\gamma, V \setminus \gamma)|\right) \cdot Z^{\mathsf{Potts}}_{G[\gamma], q-1, \beta}.$$

The associated polymer partition function is then exactly

$$\Xi_{\mathcal{C}}(\boldsymbol{w}) = \exp(-\beta \cdot |E|) \cdot Z^{\mathsf{Potts}}_{G,q,\beta}.$$

The intuitive picture is the following: If we fix an arbitrarily chosen color $\mathfrak{c}^* \in [q]$, then each γ represents a connected component of vertices who are colored anything but $\mathfrak{c}^*$. In other words, the polymers capture "defects" relative to the ground state where all vertices are assigned $\mathfrak{c}^*$. Specializing to $q = 2$, this gives another way of encoding the ferromagnetic Ising model as an abstract polymer model, which is very different from the one we say in Section 1.2.

By itself, this single polymer model isn't so useful since the weights $Z^{\mathsf{Potts}}_{G[\gamma],q-1,\beta}$ are almost hopelessly complicated, especially if $|\gamma|$ is large. However, [JKP20] proved the following interesting result: Suppose G is an *expander* graph and β is sufficiently large as a function of q, the expansion factor, and the maximum degree of G. Then, we can *approximate* the Potts model on G as a q-component mixture of these polymer models *truncated* only to polymers of size $\leq |V|/2$. Each such truncated polymer model captures deviations from each of the q ground states, and has a convergent cluster expansion which we can approximate via Barvinok's algorithm. We fully implement this strategy for a simpler model in Section 3.

This approximation result formalizes the intuition that the main contributions to $Z^{\mathsf{Potts}}_{G,q,\beta}$ come from configurations which are predominantly one color. Notably, it also implies *torpid* mixing of local Markov chains like Glauber dynamics. We refer interested readers to [JKP20] for more details.

1.4 The Random Cluster Model

Let $G = (V, E)$ be a graph, and let $q \in \mathbb{R}_{\geq 0}, 0 \leq p \leq 1$ be parameters. Define a subgraph polymer model as follows:

- A polymer γ is a subset of edges $F \subseteq E$ such that the subgraph (V_F, F) is connected, where as before, we write $V_F = \{v : \deg_F(v) > 0\}$.
- Two polymers $\gamma = (V, F), \gamma' = (V, F')$ are compatible if $V_F \cap V_{F'} = \emptyset$.
- The weight of a polymer $\gamma = (V, F)$ is given by

$$w_\gamma \stackrel{\text{def}}{=} q^{-(|V_F|-1)} \cdot \left(\frac{p}{1-p}\right)^{|F|}.$$

The associated polymer partition function is given by

$$\Xi_{\mathcal{C}}(\boldsymbol{w}) = q^{-|V|}(1-p)^{-|E|} \cdot Z^{\mathsf{RC}}_G(p, q),$$

where

$$Z^{\mathsf{RC}}_G(p, q) \stackrel{\text{def}}{=} \sum_{F \subseteq E} q^{k(F)} \cdot p^{|F|}(1-p)^{|E \setminus F|}$$

is the partition function of the *Fortuin–Kasteleyn random cluster model* with parameters p, q [FK72]. We refer interested readers to [HJP23] for further connections between the random cluster model and abstract polymer models. The following lemma shows that this gives an alternative encoding of the ferromagnetic Potts model compared to the one in Section 1.3.

Lemma 1.1. *Let $G = (V, E)$ be a graph. Then for every positive integer $q \in \mathbb{N}$ and every $\beta \geq 0$,*

$$Z^{\mathsf{Potts}}_{G,q,\beta} = e^{-\beta \cdot |E|} \cdot Z^{\mathsf{RC}}_G\left(1 - e^{-\beta}, q\right).$$

Proof. We expand the Potts partition function as

$$\begin{aligned}
Z^{\mathsf{Potts}}_{G,q,\beta} &= \sum_{\sigma:V\to[q]} \prod_{uv\in E} \exp\left(\beta \cdot \mathbf{1}[\sigma(u)=\sigma(v)]\right) \\
&= \sum_{\sigma:V\to[q]} \prod_{uv\in E} \left(1 + \left(e^{\beta}-1\right)\cdot \mathbf{1}[\sigma(u)=\sigma(v)]\right) \\
&= \sum_{\sigma:V\to[q]} \sum_{F\subseteq E} \left(e^{\beta}-1\right)^{|F|} \prod_{uv\in F} \mathbf{1}[\sigma(u)=\sigma(v)] \\
&= \sum_{F\subseteq E} \left(e^{\beta}-1\right)^{|F|} \sum_{\sigma:V\to[q]} \prod_{uv\in F} \mathbf{1}[\sigma(u)=\sigma(v)] \\
&= \sum_{F\subseteq E} q^{k(F)} \left(e^{\beta}-1\right)^{|F|} \\
&= e^{\beta\cdot|E|} \cdot Z^{\mathsf{RC}}_{G}\left(1-e^{-\beta}, q\right).
\end{aligned}$$

In the penultimate step, we used the fact that every connected component of (V,F) must be monochromatic in order to have nonzero contribution, and so we have exactly q choices for each such component. □

Remark 1. By sending β to $-\infty$, these manipulations also establish the following interesting formula for the number of proper q-colorings in a graph:

$$\sum_{F\subseteq E} (-1)^{|F|} \cdot q^{k(F)}.$$

2 The Cluster Expansion

Now suppose we view each weight w_γ not as a single complex number, but as a function $w_\gamma : \mathbb{C} \to \mathbb{C}$ (e.g. we could replace w_γ by $w_\gamma \cdot e^z$ for a variable $z \in \mathbb{C}$). For instance, in the Ising model example Section 1.2, w_γ was naturally expressed as a function of the external field. Then the polymer partition function becomes a function

$$\Xi_G(z) = \sum_{\Gamma} \prod_{\gamma\in\Gamma} w_\gamma(z).$$

We can then ask expand the Taylor series of its logarithm, and ask where this series converges. We'll eventually study the region of validity of this series, but for now, we state a beautiful combinatorial formula for the coefficients of the expansion. This is essentially the same idea as what we did previously for the matching polynomial. We refer interested readers to [Dob96; SS05; FV17; PR17] for more general statements.

Throughout, let $(\mathcal{C}, \boldsymbol{w} : \mathcal{C} \to \mathbb{C})$ be an abstract polymer model.

Definition 2 (Cluster)**.** *For a multiset of polymers Γ (with repetitions allowed), let H_Γ denote its* incompatibility graph*, where each vertex of H_Γ corresponds to a polymer in Γ and two such polymers are connected if and only if they are incompatible. A* cluster *is a multiset of polymers such that H_Γ is connected.*

Definition 3 (Ursell Function)**.** *Define the* Ursell function *$\varphi : \{\mathsf{Graphs}\} \to \mathbb{R}$ mapping graphs to reals by*

$$\varphi(H) = \sum_{\substack{A\subseteq E(H) \\ \text{spanning and} \\ \text{connected}}} (-1)^{|A|}.$$

Later on when we consider the issue of computing the Ursell Function for small graphs H, we will need the fact that this is an evaluation of the *Tutte polynomial* of H at $(x,y) = (1,0)$; see Eq. (6).

Theorem 2.1 (Cluster Expansion as Taylor Series; Dobrushin [Dob96]). *For every $k \in \mathbb{N}$, we have the formula*

$$\frac{d^k}{dz^k} \log \Xi_G(\boldsymbol{w}e^z)\Big|_{z=0} = \sum_{j=1}^{k} \frac{1}{j!} \sum_{\substack{\Gamma=(\gamma_1,\dots,\gamma_j) \\ cluster}} \varphi(H_\Gamma) \prod_{i=1}^{j} w_{\gamma_j}.$$

Furthermore, in a neighborhood of $z = 0$,

$$\log \Xi_G(\boldsymbol{w}e^z) = \sum_{k=0}^{\infty} \frac{e^{kz}}{k!} \sum_{\substack{\Gamma=(\gamma_1,\dots,\gamma_k) \\ cluster}} \varphi(H_\Gamma) \prod_{\gamma\in\Gamma} w_\gamma,$$

and absolute convergence holds in the right-hand side.

Remark 2. One can generalize this to where $\boldsymbol{w} : \mathcal{C} \to \mathbb{C}$ is some other function of z, rather than simply scaling everything by e^z. In this case, one replaces $e^{kz}\prod_{i=1}^{k} w_{\gamma_i}$ by $\frac{d^k}{dz^k}\prod_{i=1}^{k} w_{\gamma_i}(z)$.

Since our focus in this lecture is on applications of the cluster expansion, we omit the proof and refer interested readers to [Dob96]. For a derivation of the cluster expansion at a purely formal level, see [SS05; FV17].

2.1 Convergence Criteria for the Cluster Expansion

We now recall the Kotecký–Preiss Condition for convergence. We previously stated it for completely generic hardcore models. Here, we restate it in the language of abstract polymer models, along with quantitative control on the rate of convergence.

Theorem 2.2 (Kotecký–Preiss Condition; [KP86]). *Suppose there exists nonnegative functions $a, b : \mathcal{C} \to \mathbb{R}_{\geq 0}$ such that*

$$\sum_{\gamma' \not\sim \gamma} |w_{\gamma'}| \cdot e^{a(\gamma')+b(\gamma')} \leq a(\gamma), \qquad \forall \gamma \in \mathcal{C}. \tag{3}$$

Then the cluster expansion for $\log \Xi_{\mathcal{C}}$ converges absolutely. Furthermore,

$$\sum_{\substack{\Gamma=(\gamma_1,\dots,\gamma_k)\ cluster \\ \gamma' \not\sim \gamma \text{ for some } \gamma'\in\Gamma}} \frac{1}{k!} \left| \varphi(H_\Gamma) \prod_{\gamma'\in\Gamma} w_{\gamma'} e^{b(\gamma')} \right| \leq a(\gamma), \qquad \forall \text{ polymers } \gamma \in \mathcal{C}.$$

Remark 3 (Markov Chains on Polymers). As we noted in the previous lecture, this criterion is essentially the same as classical path coupling proofs of rapid mixing for Glauber dynamics (e.g. Dobrushin/Dobrushin–Shlosman [Dob70; DS85], ℓ_2-Dobrushin [Hay06; DGJ09], etc.). In the context of polymer models, this path coupling analysis was carried out in [Che+21], where they studied what is essentially a (variant of) Glauber dynamics run on the polymer model; they call this Markov chain the *polymer dynamics*. The goal here is to develop faster algorithms for counting and sampling in polymer models, particularly for applications like the bipartite hardcore model (see Section 3) and the ferromagnetic Potts model (see Section 1.3). They show that the following *polymer mixing condition*

$$\sum_{\gamma' \not\sim \gamma} w_{\gamma'} \cdot |\gamma'| \leq (1-\delta) \cdot |\gamma|, \qquad \forall \gamma \in \mathcal{C}$$

implies contraction of path coupling w.r.t. an appropriate Hamming metric weighted via $|\gamma|$; in particular, polymer dynamics mixes in $O(n \log n)$ steps. However, the main challenge for the polymer dynamics isn't necessarily fast mixing, but actually implementing a single step of the Markov chain. This is because there can be exponentially many polymers in a given polymer model. Hence, to obtain fast algorithms, we actually need the weights w_γ to decay sufficiently quickly w.r.t. $|\gamma|$. We refer interested readers to [Che+21] for more details.

In the setting of subgraph polymer models, we truly need to study multivariate zero-freeness. Indeed, since polymers are connected subgraphs of G, large polymers γ will have much larger degree just by virtue of polymers γ' contained within γ. To accommodate this, in most applications, we will typically take the functions $a(\gamma), b(\gamma)$ to grow linearly in $|\gamma|$.

3 Application: Bipartite Hardcore Model on Expanders

Throughout this section, let $G = (V, E)$ be a *bipartite* graph, with bipartition $V = L \sqcup R$ and maximum degree Δ. We also fix a fugacity $\lambda \geq 0$ and consider the problem of estimating the partition function of the hardcore model on G, i.e. the independence polynomial $Z_G(\lambda)$. Recall that if λ is any number beyond the uniqueness threshold $\lambda_c(\Delta) \approx \frac{e}{\Delta - 1}$, then there is no FPRAS for this problem on *general* graphs of maximum degree Δ unless NP = RP. Furthermore, if $\lambda > \lambda_c(\Delta)$, then with high probability over a uniformly random Δ-regular bipartite graph, local Markov chains like Glauber dynamics require exponentially many steps to mix [MWW07]. Following [JKP20], we show that despite these negative results, we can design FPTAS in the bipartite setting via polymer models when λ is *much larger* than $\lambda_c(\Delta)$ and G is an expander. Throughout, for $S \subseteq V$, we write $N(S) \stackrel{\text{def}}{=} \{v \in V \setminus S : v \sim u \text{ for some } u \in S\}$ for the open neighborhood of S.

Definition 4 (Bipartite α-Expander)**.** *Let $\alpha > 0$ be a constant. We say a bipartite graph $G = (L \sqcup R, E)$ is a* bipartite α-expander *if $|N(S)| \geq (1 + \alpha) \cdot |S|$ for all $S \subseteq L$ with $|S| \leq |L|/2$ and all $S \subseteq R$ with $|S| \leq |R|/2$. Note that if $S \subseteq L$, then $N(S) \subseteq R$ and vice versa.*

Theorem 3.1 ([JKP20])**.** *There exists a universal constant $C > 0$ such that the following holds: For every $\alpha > 0$, $\Delta \in \mathbb{N}$ and $\lambda > C\Delta^{4/\alpha}$, there exists an FPTAS for approximating $Z_G(\lambda)$ any bipartite α-expander graph G of maximum degree Δ.*

Our focus will be on approximate counting, but it isn't difficult to convert this into an approximate sampler; see also Remark 3. [JKP20] also showed that if we replace an arbitrary bipartite α-expander with a uniformly random Δ-regular graph, then $\lambda \geq \Omega\left(\frac{\log^2 \Delta}{\Delta}\right)$ suffices, which almost matches $\lambda_c(\Delta)$ up to logarithmic factors in Δ.

3.1 Bipartite Independent Set as a Mixture of Polymer Models

To prove Theorem 3.1, we leverage polymer models and the cluster expansion. Since we are operating in the regime of large λ, we cannot view our input hardcore model as an abstract polymer model and apply the cluster expansion directly. We instead decompose it into a mixture of two polymer models, corresponding to two natural ground states. The intuition is that, particularly for an expander graph, the large independent sets G are either dominated by vertices in L or by vertices in R, but not both.

To formalize this, let G^2 be the graph obtained from G by connecting two vertices if and only if their graph distance in G is at most 2. Define an abstract polymer model $(\mathcal{C}_L, \sim_L, \boldsymbol{w}(\lambda))$ as follows:

- A polymer in $\mathcal{C}_L$ is any subset of vertices $\gamma \subseteq L$ such that $|\gamma| \leq |L|/2$ and the induced subgraph $G^2[\gamma]$ is connected in G^2. One should think of γ as being a maximal connected component of $I \cap L$ in G^2, where $I \subseteq V$ is some independent set.
- We say two polymers γ, γ' are compatible under $\sim_L$ if $G^2[\gamma \sqcup \gamma']$ is *not* connected in G^2.
- The weight of a polymer γ is given by $w_\gamma(\lambda) = \frac{\lambda^{|\gamma|}}{(1+\lambda)^{|N(\gamma)|}}$. Note that $N(\gamma) \subseteq R$ is precisely the set of vertices which are forced to be out of the independent set if we condition the vertices of γ to be in. In particular,

$$\sum_{\substack{I \subseteq V \text{ indep.} \\ I \cap L = \gamma}} \lambda^{|I|} = (1 + \lambda)^{|R|} \cdot w_\gamma(\lambda), \qquad \forall \text{ polymers } \gamma \in \mathcal{C}_L. \tag{4}$$

One should think of the polymers in $\mathcal{C}_L$ as representing *deviations* from independent sets dominated by vertices in R.

We define a second polymer model $(\mathcal{C}_R, \sim_R, \boldsymbol{w}(\lambda))$ in a completely analogous way, with all occurrences of L replaced by R. Let Ξ_L, Ξ_R denote their corresponding polymer partition functions. The following result gives our desired approximation.

Proposition 3.2. *Let $G = (L \sqcup R, E)$ be a bipartite α-expander. Then for every $\lambda > e^{11/\alpha}$, the number*

$$(1 + \lambda)^{|R|} \cdot \Xi_L(\boldsymbol{w}(\lambda)) + (1 + \lambda)^{|L|} \cdot \Xi_R(\boldsymbol{w}(\lambda)) \tag{5}$$

is a $(1 \pm e^{-n})$-multiplicative approximation to $Z_G(\lambda)$.

To prove this, we need the following technical lemma, which says that the restriction on the polymer size $|\gamma| \le |L|/2$ in $\mathcal{C}_L$ (and respectively for $\mathcal{C}_R$) is essentially without loss of generality. We proof is provided in Appendix A.

Lemma 3.3. *Let I be an independent set in a bipartite α-expander. Then $|I \cap L| \le |L|/2$ or $|I \cap R| \le |R|/2$.*

Proof of Proposition 3.2. The essence of Eq. (5) comes from Eq. (4), which counts the contribution of independent sets which intersect any particular polymer γ. For convenience, let us call a subset of vertices $S \subseteq V$ *sparse* if the maximal connected components of $S \cap L \subseteq L$ have size at most $|L|/2$ (and analogously for the components of $S \cap R$). An immediate consequence of Eq. (4) is

$$Z_G(\lambda) = \sum_{\substack{I \subseteq V \text{ indep.} \\ I\cap L \text{ or } I \cap R \text{ is sparse}}} \lambda^{|I|} + \underbrace{\sum_{\substack{I \subseteq V \text{ indep.} \\ I\cap L \text{ and } I \cap R \text{ not sparse}}} \lambda^{|I|}}_{=0 \text{ by Lemma 3.3}}.$$

Similarly, we may expand Eq. (5) as

$$\begin{aligned}
&(1+\lambda)^{|R|} \cdot \Xi_L(\boldsymbol{w}(\lambda)) + (1+\lambda)^{|L|} \cdot \Xi_R(\boldsymbol{w}(\lambda)) \\
&= \sum_{\substack{I \subseteq V \text{ indep.} \\ I\cap L \text{ is sparse}}} \lambda^{|I|} + \sum_{\substack{I \subseteq V \text{ indep.} \\ I\cap R \text{ is sparse}}} \lambda^{|I|} && \text{(Using Eq. (4))} \\
&= Z_G(\lambda) + \sum_{\substack{I \subseteq V \text{ indep.} \\ I \text{ is sparse}}} \lambda^{|I|}. && \text{(Double counting and previous display)}
\end{aligned}$$

This already shows that $Z_G(\lambda)$ is a lower bound for Eq. (5). It remains to show that

$$\sum_{\substack{I \subseteq V \text{ indep.} \\ I \text{ is sparse}}} \lambda^{|I|} \le e^{-n} \cdot Z_G(\lambda).$$

Since $Z_G(\lambda) \ge \lambda^{|L|} + \lambda^{|R|} \ge \lambda^{n/2}$, what we will show is that any sparse independent set in a bipartite α-expander must have size much smaller than $n/2$. This is intuitive because if $I \cap L$ is large, then by expansion, so is $N(I \cap L) \subseteq R$. But since $N(I \cap L)$ must be disjoint from $I \cap R$, it must force $I \cap R$ to be small. We now formalize this intuition.

Let $I \subseteq V$ be any sparse independent set, and let $A = I \cap L, B = I \cap R$. Since the maximal connected components of A in G^2 have size $\le |L|/2$, we can apply expansion to each individual polymer in A, and combine them to obtain

$$|N(A)| \ge (1+\alpha) \cdot |A|,$$

even if A itself has size greater than $|L|/2$. Similarly, $|N(B)| \ge (1+\alpha) \cdot |B|$. So far, we've only used sparsity. But because I is independent, B must be disjoint from $N(A)$ and vice versa. It follows that

$$\begin{aligned}
n &= |L| + |R| \\
&\ge |A| + |N(A)| + |B| + |N(B)| && (I \text{ is independent}) \\
&\ge (2+\alpha) \cdot |I|, && (I \text{ is sparse and } I = A \sqcup B)
\end{aligned}$$

i.e. all sparse independent sets $I \subseteq V$ satisfy $|I| \le \frac{n}{2+\alpha}$. Using a trivial bound on the total number of sparse independent sets, it follows that

$$\frac{\sum_{\substack{I \subseteq V \text{ indep.} \\ I \text{ is sparse}}} \lambda^{|I|}}{Z_G(\lambda)} \le \frac{2^n \cdot \lambda^{\frac{n}{2+\alpha}}}{\lambda^{n/2}} \le e^{-n} \qquad \text{(Using } \lambda > e^{11/\alpha}\text{)}$$

as desired. □

3.2 Convergence of the Cluster Expansion for $\mathcal{C}_L, \mathcal{C}_R$

Now that we have the polymer models we want to work with, we now need to verify that their cluster expansions converge as we want them to. This is so that we can approximate each of $\Xi_L(\boldsymbol{w}(\lambda)), \Xi_R(\boldsymbol{w}(\lambda))$, which together approximate $Z_G(\lambda)$ via Proposition 3.2.

Proposition 3.4. *There exists a universal constant $C > 0$ such that the following holds: For every $\alpha > 0$, $\Delta \in \mathbb{N}$, $\lambda > C\Delta^{4/\alpha}$, and bipartite α-expander $G = (L \sqcup R, E)$ of maximum degree Δ, the two polymer models $\mathcal{C}_L, \mathcal{C}_R$ defined in Section 3.1 satisfy the Kotecký–Preiss Condition Eq. (3) with $a(\gamma) = b(\gamma) = |\gamma|$ for all polymers γ.*

To show this, we need the following lemma, which controls the number of neighboring incompatible polymers.

Lemma 3.5 ([Bor+13])**.** *Let $G = (V, E)$ be a graph of maximum degree Δ. Then for every vertex $v \in V$, the number of $S \subseteq V$ such that $v \in S$, $G[S]$ is connected, and $|S| = t$ is at most $(e\Delta)^t$.*

We skip the proof of this lemma. The main idea behind it is to count the number of t-vertex trees in G containing v, which is certainly an upper bound on the number of connected induced subgraphs we're looking at. Once we have restricted attention to trees, it is not hard to see that the worst case is when G is the infinite Δ-regular tree. For this, there is an explicit formula for the number of t-vertex subtrees containing some arbitrarily fixed "root" vertex. Furthermore, it is upper bounded by $(e\Delta)^t$.

Proof of Proposition 3.4. We do the analysis for $\mathcal{C}_L$; the one for $\mathcal{C}_R$ follows by symmetry.

$$\begin{aligned}
\sum_{\gamma' \not\sim_L \gamma} w_{\gamma'}(\lambda) \cdot e^{2|\gamma'|} &\leq \sum_{\gamma' \not\sim_L \gamma} (1+\lambda)^{-\alpha \cdot |\gamma'|} \cdot e^{2|\gamma'|} && (\alpha\text{-expansion}) \\
&\leq \sum_{v \in N[N[\gamma]]} \sum_{\gamma' : \gamma' \ni v} (1+\lambda)^{-\alpha \cdot |\gamma'|} \cdot e^{2|\gamma'|} && (\text{Definition of } \not\sim_L) \\
&\leq \sum_{v \in N[N[\gamma]]} \sum_{t=1}^{\infty} \left(e^3 \Delta^2 (1+\lambda)^{-\alpha}\right)^t \\
& && (\text{Lemma 3.5}, G^2 \text{ has maximum degree} \leq \Delta^2) \\
&\leq \frac{|N[N[\gamma]]|}{\Delta^2} && (\text{Using } \lambda \geq \Omega(\Delta^{4/\alpha})) \\
&\leq |\gamma| && (G \text{ has maximum degree } \Delta)
\end{aligned}$$

□

3.3 Proof of Theorem 3.1

Suppose we wish to estimate $Z_G(\lambda)$ up to $(1 \pm \epsilon)$-multiplicative accuracy for some $\epsilon > 0$. If $\epsilon \leq O(e^{-n})$, then our $\mathsf{poly}(1/\epsilon)$ running time budget from the definition of an FPTAS permits brute force enumeration. Hence, we may assume $\epsilon > \Omega(e^{-n})$, which is the nontrivial error regime. In light of Proposition 3.2, it then suffices to compute a $(1 \pm O(\epsilon))$-multiplicative approximation to each of $\Xi_L(\boldsymbol{w}(\lambda))$ and $\Xi_R(\boldsymbol{w}(\lambda))$. We do this for the former via truncating the cluster expansion; the latter follows symmetrically.

For a parameter $m \in \mathbb{N}$, define

$$T_m(\boldsymbol{w}) \stackrel{\text{def}}{=} \sum_{\substack{\Gamma = (\gamma_1, \ldots, \gamma_k) \text{ cluster} \\ b(\Gamma) \leq m}} \frac{1}{k!} \cdot \varphi(H_\Gamma) \cdot \prod_{\gamma \in \Gamma} w_\gamma$$

where we set $b(\Gamma) \stackrel{\text{def}}{=} \sum_{\gamma\in\Gamma} b(\gamma)$ for convenience. Then

$$
\begin{aligned}
|T_m(\boldsymbol{w}) - \log \Xi_G(\boldsymbol{w})| &\leq \sum_{\substack{\Gamma=(\gamma_1,\ldots,\gamma_k) \text{ cluster} \\ b(\Gamma)>m}} \frac{1}{k!} \left|\varphi(H_\Gamma) \prod_{\gamma\in\Gamma} w_\gamma\right| && \text{(Theorem 2.1)} \\
&\leq e^{-m} \cdot \sum_{v\in V} \sum_{\substack{\Gamma=(\gamma_1,\ldots,\gamma_k) \text{ cluster} \\ b(\Gamma)>m, \gamma\ni v \text{ for some } \gamma\in\Gamma}} \frac{1}{k!} \left|\varphi(H_\Gamma) \prod_{\gamma\in\Gamma} w_\gamma e^{b(\gamma)}\right| \\
&\leq e^{-m} \cdot \sum_{v\in V} a(\{v\}) && \text{(Proposition 3.4 and Theorem 2.2)} \\
&\leq n \cdot e^{-m}. && \text{(Using } a(\gamma) = |\gamma|\text{)}
\end{aligned}
$$

Taking $m \leq O(\log(n/\epsilon))$, we obtain our desired approximation. All that remains is to show how to enumerate the clusters $\Gamma = (\gamma_1, \ldots, \gamma_k)$ such that $b(\Gamma) \leq O(\log(n/\epsilon))$ and compute their associated Ursell functions $\varphi(H_\Gamma)$. Since we chose $b(\gamma) = |\gamma|$, brute force enumeration of such Γ amounts to enumeration of all connected induced subgraphs of G with at most $O(\log(n/\epsilon))$ many vertices. This can by done in $(n/\epsilon)^{O(\log \Delta)}$-time by Lemma 3.5. For computation of the Ursell functions, since each H_Γ only has $k \leq O(\log(n/\epsilon))$ many vertices, we can afford running time which is singly-exponential in k. Unfortunately, naïve brute force requires $2^{O(k^2)}$-time as there can be up to k^2 many edges in H_Γ. For this, we instead employ an algorithm of [Bjö+08] as a blackbox, since the Ursell function $\varphi(H_\Gamma)$ is an evaluation of the *Tutte polynomial* of H_Γ

$$
T_H(x,y) \stackrel{\text{def}}{=} \sum_{A\subseteq E(H)} (x-1)^{k(A)-k(E)} (y-1)^{k(A)+|A|-|V|}. \tag{6}
$$

at $(x, y) = (1, 0)$.

We emphasize that it was crucial to truncate based on the function $b(\cdot)$, as opposed to truncating based only on the number of polymers in Γ, even though the latter is what we did in previous applications of Barvinok's method. This is because by our choice of $b(\gamma) = |\gamma|$, the restriction $b(\Gamma) \leq m$ limits not only the number of polymers in Γ, but also bounds the size of each polymer $\gamma \in \Gamma$ by m. Hence, we only have to consider small polymers, of which there are only polynomially many. This is how we can accommodate the fact that we're looking at independent sets in an exponentially large graph, the incompatibility graph of polymers.

4 Conclusion

The polymer method and cluster expansion are versatile tools in the study of Gibbs distributions. One of the central themes of this framework is to design polymers which accurately capture *deviations* from some underlying "perfect" object(s) (e.g. a system where there are no interactions at all). For instance, in the "low temperature"/"strong interaction" regime, these could be the various ground states of the Gibbs distribution. The fact that the interactions are strong imply that in a typical sample from the Gibbs distribution, the deviations from one of the ground states will be small. This implies that the polymer models capturing these deviations are themselves effectively in "high temperature", and so we can apply expansion methods [JKP20], Markov chains [Che+21], etc. on them. We conclude with some open problems.

Question 1. *In low-temperature settings like Section 3, existing algorithms based on truncating the cluster expansion are rather slow. There are faster MCMC-based algorithms [Che+21], but they are run on collections of polymers, and hence are more complicated to implement. They must also in their own way deal with the fact that there can be exponentially many polymers. Can we design simpler and faster algorithms? For instance, even though Glauber dynamics mixes slowly in the worst case, can we show that with an appropriate easy-to-sample-from initializing distribution, Glauber dynamics mixes in polynomial-time?*

Question 2 (#BIS)**.** *Does there exist an* FPRAS *for approximately counting (unweighted) independent sets in arbitrary* bipartite *graphs? Even the existence of* subexponential-time *approximation algorithms is open.*

#BIS is an important problem in the complexity theory of approximate counting and sampling since many natural problems reduce to it, e.g. counting downsets in a poset [Dye+04], counting stable matchings in the stable marriage problem [CGM12], computing the partition function of the ferromagnetic Potts model/random cluster models [GJ08; GJ12a; GJ12b; GJ13; GJ14], and computing the partition function of the ferromagnetic Ising model with inconsistent external fields [GJ07] (see also [LLZ14]). #BIS is also *complete* for a certain logically-defined complexity class known as $\#\mathsf{RH}\Pi_1$ under approximation-preserving reductions [Dye+04]. Its complexity is wide open.

References

[Bjö+08] Andreas Björklund, Thore Husfeldt, Petteri Kaski, and Mikko Koivisto. "Computing the Tutte Polynomial in Vertex-Exponential Time". In: *2008 49th Annual IEEE Symposium on Foundations of Computer Science*. 2008, pp. 677–686. DOI: 10.1109/FOCS.2008.40 (cit. on p. 9).

[Bor+13] Christian Borgs, Jennifer Chayes, Jeff Kahn, and László Lovász. "Left and right convergence of graphs with bounded degree". In: *Random Structures & Algorithms* 42.1 (2013), pp. 1–28 (cit. on p. 8).

[CGM12] Prasad Chebolu, Leslie Ann Goldberg, and Russell Martin. "The Complexity of Approximately Counting Stable Matchings". In: *Theoretical Computer Science* Volume 437 (2012), pp. 35–68 (cit. on p. 10).

[Che+21] Zongchen Chen, Andreas Galanis, Leslie Ann Goldberg, Will Perkins, James Stewart, and Eric Vigoda. "Fast algorithms at low temperatures via Markov chains". In: *Random Structures & Algorithms* 58 (2 2021), pp. 294–321 (cit. on pp. 5, 9).

[DGJ09] Martin Dyer, Leslie Ann Goldberg, and Mark Jerrum. "Matrix Norms and Rapid Mixing for Spin Systems". In: *The Annals of Applied Probability* 19.1 (Feb. 2009), pp. 71–107 (cit. on p. 5).

[Dob70] Roland Lvovich Dobrushin. "Prescribing a System of Random Variables by Conditional Distributions". In: *Theory of Probability and its Applications* 15.3 (1970), pp. 458–486 (cit. on p. 5).

[Dob96] Roland Lvovich Dobrushin. "Estimates of Semiinvariants for the Ising Model at Low Temperatures". In: *American Mathematical Society Translations* 177.32 (1996). Topics in statistical and theoretical physics. F. A. Berezin memorial volume, pp. 59–81 (cit. on pp. 1, 4, 5).

[DS85] R. L. Dobrushin and S. B. Shlosman. "Completely Analytical Gibbs Fields". In: *Statistical Physics and Dynamical Systems: Rigorous Results*. Boston, MA: Birkhäuser Boston, 1985, pp. 371–403. ISBN: 978-1-4899-6653-7 (cit. on p. 5).

[Dye+04] Martin Dyer, Leslie Ann Goldberg, Catherine Greenhill, and Mark Jerrum. "The Relative Complexity of Approximate Counting Problems". In: *Algorithmica* 38 (2004), pp. 471–500 (cit. on p. 10).

[FK72] Cornelis Marius Fortuin and Pieter Willem Kasteleyn. "On the random-cluster model: I. Introduction and relation to other models". In: *Physica* 57 (4 1972), pp. 536–564 (cit. on p. 3).

[FV17] Sacha Friedli and Yvan Velenik. "Statistical Mechanics of Lattice Systems. A Concrete Mathematical Introduction". In: (2017) (cit. on pp. 1, 4, 5).

[GJ07] Leslie Ann Goldberg and Mark Jerrum. "The Complexity of Ferromagnetic Ising with Local Fields". In: *Combinatorics, Probability and Computing* 16 (1 2007), pp. 43–61 (cit. on p. 10).

[GJ08] Leslie Ann Goldberg and Mark Jerrum. "Inapproximability of the Tutte polynomial". In: *Inform. and Comput* 206.7 (2008), pp. 908–929 (cit. on p. 10).

[GJ12a] Leslie Ann Goldberg and Mark Jerrum. "Approximating the partition function of the ferromagnetic Potts model". In: *J. ACM* 59.5 (2012), pp. 1–31 (cit. on p. 10).

[GJ12b] Leslie Ann Goldberg and Mark Jerrum. "Inapproximability of the Tutte polynomial of a planar graph". In: *Computational Complexity* 21 (4 2012), pp. 605–642 (cit. on p. 10).

[GJ13] Leslie Ann Goldberg and Mark Jerrum. "Approximating the Tutte polynomial of a binary matroid and other related combinatorial polynomials". In: *Journal of Computer and System Sciences* 79 (1 2013), pp. 68–78 (cit. on p. 10).

[GJ14] Leslie Ann Goldberg and Mark Jerrum. "The complexity of computing the sign of the Tutte polynomial". In: *SIAM J. Comput.* 43 (2014), pp. 1921–1952 (cit. on p. 10).

[Hay06] Thomas P. Hayes. "A simple condition implying rapid mixing of single-site dynamics on spin systems". In: *FOCS* (2006), pp. 39–46 (cit. on p. 5).

[HJP23] Tyler Helmuth, Matthew Jenssen, and Will Perkins. "Finite-size scaling, phase coexistence, and algorithms for the random cluster model on random graphs". In: *Annales de l'institut Henri Poincare (B) Probability and Statistics* 59.2 (2023), pp. 817–848 (cit. on p. 3).

[HPR20] Tyler Helmuth, Will Perkins, and Guus Regts. "Algorithmic Pirogov-Sinai theory". In: *Probability Theory and Related Fields* 176 (2020), pp. 851–895 (cit. on pp. 1, 2).

[JKP20] Matthew Jenssen, Peter Keevash, and Will Perkins. "Algorithms for #BIS-hard problems on expander graphs". In: *SIAM Journal on Computing* 49 (4 2020) (cit. on pp. 1–3, 6, 9).

[KP86] R. Kotecký and D. Preiss. "Cluster expansion for abstract polymer models". In: *Communications in Mathematical Physics* 103.3 (1986), pp. 491–498 (cit. on p. 5).

[LLZ14] Jingcheng Liu, Pinyan Lu, and Chihao Zhang. "The Complexity of Ferromagnetic Two-spin Systems with External Fields". In: *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2014)*. Ed. by Klaus Jansen, José D. P. Rolim, Nikhil R. Devanur, and Cristopher Moore. Vol. 28. Leibniz International Proceedings in Informatics (LIPIcs). Dagstuhl, Germany: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik, 2014, pp. 843–856. ISBN: 978-3-939897-74-3. DOI: 10.4230/LIPIcs.APPROX-RANDOM.2014.843 (cit. on p. 10).

[MWW07] Elchanan Mossel, Dror Weitz, and Nicholas C. Wormald. "On the hardness of sampling independent sets beyond the tree threshold". In: *Probability Theory and Related Fields* 143 (2007), pp. 401–439 (cit. on p. 6).

[PR17] Viresh Patel and Guus Regts. "Deterministic Polynomial-Time Approximation Algorithms for Partition Functions and Graph Polynomials". In: *SIAM Journal on Computing* 46.6 (2017), pp. 1893–1919 (cit. on p. 4).

[SS05] Alexander D. Scott and Alan D. Sokal. "The Repulsive Lattice Gas, the Independent-Set Polynomial, and the Lovász Local Lemma". In: *Journal of Statistical Physics* 118.5 (2005), pp. 1151–1261 (cit. on pp. 1, 4, 5).

A Unfinished Proofs

Proof of Lemma 3.3. Without loss of generality, we may assume $|L| \leq |R|$. Let $A = I \cap L, B = I \cap R$, and suppose $|B| > |R|/2$. By choosing $S \subseteq B$ such that $|S| = |R|/2$, we see that $|N(B)| \geq |N(S)| \geq (1+\alpha) \cdot |S| = (1+\alpha) \cdot |R|/2$. Since I is an independent, A and $N(B)$ must be disjoint, whence

$$|A| \leq |L| - |N(B)| < |L| - (1+\alpha) \cdot |R|/2 < |L|/2.$$

□