CS294 Partition Functions	**Fall 2023**

Lecture Note 18

Instructor: Alistair Sinclair

Disclaimer: *These notes have not been subjected to the usual scrutiny accorded to formal publications. They may be distributed outside this class only with the permission of the Instructor.*

In this lecture we continue our discussion of the *hard-core model* of independent sets in bounded degree graphs.

18.1 Setup

For a graph $G = (V, E)$ with maximum degree bounded by Δ let a_k denote the number of independent sets in G of size k and for a fixed real $\lambda > 0$ define the *hard-core model partition function* $Z_G(\lambda)$ as follows:

$$Z_G(\lambda) := \sum_{k \geq 0} a_k \lambda^k.$$

The main algorithmic task of interest to us is computing the value of $Z_G(\lambda)$ given G and λ as input. Recall that our goal is to establish the following dichotomy for this problem in bounded degree graphs.

Theorem 18.1. *For any fixed $\Delta \geq 3$ define $\lambda_c(\Delta) := \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$. Then:*

1. *If $\lambda < \lambda_c(\Delta)$: there exists a fully polynomial time (deterministic) approximation scheme for computing $Z_G(\lambda)$ on all graphs of maximum degree Δ.*
2. *If $\lambda > \lambda_c(\Delta)$ then there is no fully polynomial time randomized approximation scheme for $Z_G(\lambda)$ on all graphs of maximum degree Δ unless $\mathsf{NP} = \mathsf{RP}$.*

A proof of part 1 of Theorem 18.1 was covered in the previous lecture. In this lecture, we work towards a proof of part 2, which is due to Allan Sly [Sly10]. This proof is via a randomized reduction from Max Cut, which is NP-hard.

18.2 Reminder: hard-core model on the infinite tree

For every λ there is a Gibbs measure μ for the hard-core model on the infinite Δ-regular tree with the same marginal $p^* = p^*(\lambda)$ on every vertex, which we refer to as the *occupation probability.* Further, when $\lambda \leq \lambda_c(\Delta)$, μ is the unique Gibbs measure. On the other hand, when $\lambda > \lambda_c(\Delta)$, there are two extremal Gibbs measures, μ_{odd} and μ_{even}, where μ_{odd} is more heavily supported on vertices at odd distance from the root and μ_{even} is more heavily supported on vertices at even distance from the root. Concretely, there is an *odd occupation probability* $p_1 = p_1(\lambda)$ and an *even occupation probability* $p_2 = p_2(\lambda)$, with $0 < p_1 < p^* < p_2 < 1$, such that the marginal of μ_{odd} on even distance vertices is p_1 and on odd distance vertices is p_2, and vice versa for μ_{even}.

This phase transition in the uniqueness of Gibbs measures on the tree is intimately connected with the computational phase transition articulated in Theorem 18.1.

18.3 Key gadget: random bipartite graph

A key gadget used in the reduction is a random bipartite graph $\boldsymbol{G}$ on left vertex set L and right vertex set R, each of size n, of maximum degree Δ sampled in the following way, which we call $\mathbf{M}(\Delta)$:

- Let $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_\Delta$ be Δ independent and uniformly random perfect matchings between L and R.
- Let $\boldsymbol{G}$ be the graph obtained by taking $\boldsymbol{M}_1 + \cdots + \boldsymbol{M}_\Delta$ and deleting duplicate edges.

Before we delve into how this gadget is used in the reduction we sketch a proof of a weaker form of hardness, namely hardness for algorithms based on Glauber dynamics.

Definition 18.2. *An* (α, β)-independent set *is one with* αn *vertices in* L *and* βn *vertices in* R*. (Note that necessarily* $\alpha + \beta \leq 1$*.) We denote the set of all* (α, β)*-independent sets in a bipartite graph* G *by* $I^{\alpha,\beta}(G)$*. With a given* (α, β) *pair, we associate the following (simplified) partition function:*

$$Z_G^{\alpha,\beta}(\lambda) := \sum_{I \in I^{\alpha,\beta}(G)} \lambda^{(\alpha+\beta)n}.$$

A key ingredient is very precise control of the expected value of $Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)$. This approach was first proposed in [DFJ99] for $\lambda = 1$, and then carried out in detail for general λ in [MWW09].

Proposition 18.3. $\mathrm{E}\left[Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)\right] = \exp(\Phi_\lambda(\alpha, \beta) \cdot n \cdot (1 + o(1)))$ *where*

$$\Phi_\lambda(\alpha, \beta) = (\alpha + \beta) \ln \lambda + H(\alpha) + H(\beta) + \Delta \left[(1 - \beta) H\left(\frac{\alpha}{1 - \beta}\right) - H(\alpha)\right]$$

and $H(x) := -x \ln x - (1 - x) \ln(1 - x)$ *denotes binary entropy.*

Proof. We begin by explicitly writing out $\mathrm{E}[Z_{\boldsymbol{G}}^{\alpha,\beta}]$:

$$\begin{aligned} \mathrm{E}\left[Z_{\boldsymbol{G}}^{\alpha,\beta}\right] &= \lambda^{(\alpha+\beta)n} \sum_{\substack{I_L \subseteq L : |I_L| = \alpha n \\ I_R \subseteq R : |I_R| = \beta n}} \Pr[I_L \cup I_R \text{ is an independent set in } \boldsymbol{G}] \\ &= \lambda^{(\alpha+\beta)n} \binom{n}{\alpha n} \binom{n}{\beta n} \Pr[I_L \cup I_R \text{ is an independent set in } \boldsymbol{G}]. \end{aligned} \quad (18.1)$$

$I_L \cup I_R$ is an independent set in $\boldsymbol{G}$ if and only if all of the Δ matchings $\boldsymbol{M}_1, \ldots, \boldsymbol{M}_\Delta$ used to sample $\boldsymbol{G}$ match vertices in I_L to vertices in $R \setminus I_R$. There are $\binom{(1-\beta)n}{\alpha n}$ ways to match I_L to $R \setminus I_R$ and a total of $\binom{n}{\alpha n}$ ways to match I_L to R. Thus the probability of all of $M_1, \ldots, M_\Delta$ matching I_L to $R \setminus I_R$ is $\left[\frac{\binom{(1-\beta)n}{\alpha n}}{\binom{n}{\alpha n}}\right]^\Delta$. Plugging this value into (18.1) gives us:

$$\mathrm{E}\left[Z_{\boldsymbol{G}}^{\alpha,\beta}\right] = \lambda^{(\alpha+\beta)n} \binom{n}{\alpha n} \binom{n}{\beta n} \cdot \left[\frac{\binom{(1-\beta)n}{\alpha n}}{\binom{n}{\alpha n}}\right]^\Delta. \quad (18.2)$$

The theorem statement can now be concluded **[exercise!]** from (18.2) via an application of Stirling's approximation of binomial coefficients, $\ln \binom{n}{cn} \sim nH(c)$. □

We now look at some analytic properties of the function $\Phi_\lambda(\alpha, \beta)$ in the relevant domain, namely the triangle $T := \{(\alpha, \beta) : \alpha, \beta \geq 0, \alpha + \beta \leq 1\}$.

Fact 18.4. *The function $\Phi_\lambda(\alpha,\beta)$ satisfies the following properties on T:*

(i) The maximum value of $\Phi_\lambda(\alpha,\beta)$ on the line $\alpha=\beta$ is achieved at $\alpha=\beta=p^*$.

(ii) If $\lambda<\lambda_c(\Delta)$, then $\alpha=\beta=p^*$ is also the location of the unique global maximum in T.

(iii) If $\lambda>\lambda_c(\Delta)$ then $\alpha=\beta=p^*$ is a saddle point and the only maxima are achieved at (p_1,p_2) and (p_2,p_1).

(iv) p^*, p_1 and p_2 are continuous functions of λ and $p_1(\lambda)-p^*(\lambda), p_2(\lambda)-p^*(\lambda)\to 0$ as $\lambda\to\lambda_c(\Delta)$.

Here p^*, p_1, p_2 are the occupation probabilities associated with the Gibbs measures on the Δ-regular tree, as stated in Section 18.2.

Proof. We first compute the partial derivatives of $\Phi := \Phi_\lambda$:

$$\begin{aligned}
\frac{\partial\Phi}{\partial\alpha} &= \ln\lambda-\ln\alpha-(\Delta-1)\ln(1-\alpha)+\Delta\ln(1-\alpha-\beta) &(18.3)\\
\frac{\partial\Phi}{\partial\beta} &= \ln\lambda-\ln\beta-(\Delta-1)\ln(1-\beta)+\Delta\ln(1-\alpha-\beta) &(18.4)\\
\frac{\partial^2\Phi}{\partial\alpha^2} &= -\frac{1}{\alpha}+\frac{\Delta-1}{1-\alpha}-\frac{\Delta}{1-\alpha-\beta} &(18.5)\\
\frac{\partial^2\Phi}{\partial\beta^2} &= -\frac{1}{\beta}+\frac{\Delta-1}{1-\beta}-\frac{\Delta}{1-\alpha-\beta} &(18.6)\\
\frac{\partial^2\Phi}{\partial\alpha\partial\beta} &= -\frac{\Delta}{1-\alpha-\beta} &(18.7)
\end{aligned}$$

First it is not hard to check using (18.3) and (18.4) that Φ cannot have a maximum anywhere on the boundary of T. Now since both $\frac{\partial^2\Phi}{\partial\alpha^2}$ and $\frac{\partial^2\Phi}{\partial\beta^2}$ are negative, Φ has a local maximum iff its Hessian has positive determinant. Using (18.5)–(18.7) and simplifying, this condition is

$$\alpha+\beta+\Delta(\Delta-2)\alpha\beta\le 1. \tag{18.8}$$

Also, from (18.3) and (18.4) any maximum (α,β) of Φ satisfies

$$\beta=f(\alpha);\qquad \alpha=f(\beta), \tag{18.9}$$

where $f(x) := (1-x)\left[1-\left(\frac{x}{\lambda(1-x)}\right)^{1/\Delta}\right]$.

Now we verify that the occupation probabilities p_1, p_2 associated with the hard-core model satisfy the same relation as (18.9) above, i.e., $p_2=f(p_1)$ and vice versa. Let v be an odd vertex in the infinite Δ-regular tree (so its neighbors, $u_1,\dots,u_\Delta$ are even). Then in the even Gibbs measure π_E the occupation probability for v is p_1 and for the u_i it is p_2. Thus we have

$$\begin{aligned}
p_1=\pi_E[\sigma(v)=1] &= \frac{\lambda}{1+\lambda}\pi_E[\sigma(u_i)=0\ \forall i]\\
&= \frac{\lambda}{1+\lambda}\left(\pi_E[\sigma(v)=1]+\pi_E[\sigma(v)=0]\prod_{i=1}^{\Delta}\pi_E[\sigma(u_i)=0\mid\sigma(v)=0]\right)\\
&= \frac{\lambda}{1+\lambda}\left(p_1+(1-p_1)(1-\hat p_2)^\Delta\right), &(18.10)
\end{aligned}$$

where $\hat{p}_2 := \pi_E[\sigma(u) = 0 \mid \sigma(v) = 0]$ for an arbitrary even vertex u and odd neighbor v. Inverting (18.10) gives $\hat{p}_2 = 1 - \left(\frac{p_1}{\lambda(1-p_1)}\right)^{1/\Delta}$, and hence, since $p_2 = (1-p_1)\hat{p}_2$, we get that $p_2 = f(p_1)$, as claimed. An identical argument shows that $p_1 = f(p_2)$.

Finally, our observations in a previous lecture about fixed points for the tree recurrence imply that, when $\lambda < \lambda_c(\Delta)$, $\alpha = \beta = p^*$ is the unique solution to $\alpha = f(\beta), \beta = f(\alpha)$, while for $\lambda > \lambda_c(\Delta)$ there are exactly two additional solutions, namely p_1, p_2. Moreover, it's easy to check that $p^* = \frac{1}{\Delta}$ at $\lambda = \lambda_c(\Delta)$, and since f is increasing p^* must be also, so $p^* > \frac{1}{\Delta}$ when $\lambda > \lambda_c(\Delta)$. Hence by condition (18.8), p^* is not a local maximum, since $2p^* + \Delta(\Delta-2)(p^*)^2 > \frac{2}{\Delta} + \frac{\Delta(\Delta-2)}{\Delta^2} = 1$. Therefore, the maximum of Φ on the triangle T is indeed achieved at the two points $(\alpha, \beta) = (p_1, p_2)$ and $(\alpha, \beta) = (p_2, p_1)$.

The observation about continuity (part (iv)) is left as an **exercise**. $\square$

The behavior of the function Φ_λ is summarized in Figure 18.1.

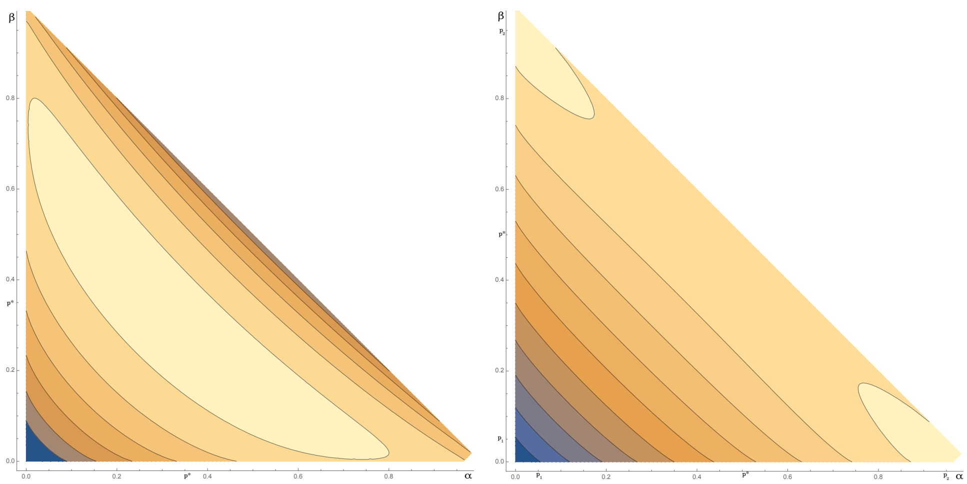

Figure 18.1: Contour plot on the left denotes level sets of Φ_λ when $\lambda < \lambda_c$. Contour plot on the right denotes level sets of Φ_λ when $\lambda > \lambda_c$.

Now we provide some intuition for why we expect Glauber dynamics to mix slowly when $\lambda > \lambda_c$. Suppose (as we shall see shortly) that $Z_{\boldsymbol{G}}(\lambda)$ is well concentrated around its expected value. Now observe that $Z_{\boldsymbol{G}}(\lambda) = \sum_{\alpha,\beta} Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)$. When $\lambda > \lambda_c(\Delta)$, this sum is dominated by $Z_{\boldsymbol{G}}^{p_1 \pm o(1), p_2 \pm o(1)}(\lambda)$ and $Z_{\boldsymbol{G}}^{p_2 \pm o(1), p_1 \pm o(1)}(\lambda)$. Thus, most of the mass of the Gibbs distribution is on $(p_1 \pm o(1), p_2 \pm o(1))$ and $(p_2 \pm o(1), p_1 \pm o(1))$-independent sets. If we start the Glauber dynamics at a (p_1, p_2)-independent set, to reach a $(p_2 \pm o(1), p_1 \pm o(1))$-independent set the walk must pass through a balanced independent set, since Glauber dynamics is only allowed to change the status of one vertex per step. Since balanced independent sets have total weight at most $nZ_{\boldsymbol{G}}^{p^*,p^*}(\lambda)$, which is exponentially smaller, this is a bottleneck for the Glauber dynamics.

We now spell out the details of the above argument. First, we need the following concentration result for the random variable $Z_{\boldsymbol{G}}(\lambda)^{\alpha,\beta}$, due to [MWW09]:

Theorem 18.5. *There is a constant η depending on Δ such that, for α, β satisfying $\left|\alpha - \frac{1}{\Delta}\right| \leq \eta$ and*

$\left|\beta - \frac{1}{\Delta}\right| \leq \eta$, *with high probability over* $G \in \mathbf{M}(\Delta)$,

$$Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda) \geq \frac{1}{n}\mathrm{E}\left[Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)\right].$$

The main technical ingredient in proving the above statement is a technically involved second-moment analysis showing that $\frac{\mathrm{E}_{\boldsymbol{G}}[Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)^2]}{\mathrm{E}_{\boldsymbol{G}}[Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)]^2} = \Theta(1)$. We omit this proof.

Note: The above theorem actually only establishes concentration under the condition that α, β are sufficiently close to $\frac{1}{\Delta}$. Note that $\frac{1}{\Delta}$ is the value of p^* exactly at $\lambda = \lambda_c(\Delta)$, and that all of p^*, p_1, p_2 are close to $\frac{1}{\Delta}$ when λ is close to $\lambda_c(\Delta)$ (by property (iv) in Fact 18.4). Thus Theorem 18.5 establishes the desired concentration of $Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)$ for $\lambda_c(\Delta) < \lambda < \lambda_c(\Delta) + \varepsilon$ for some small $\varepsilon > 0$ (that depends on Δ). Technically this only proves hardness of $Z_G(\lambda)$ for such values of λ. However, in a later paper Galanis et al. [GGŠ+14] proved an analogous concentration result that extends the proof to all values $\lambda > \lambda_c(\Delta)$. In what follows we won't dwell on this issue but will assume that Theorem 18.5 holds for all $\lambda > \lambda_c$.

Next, we show how to formalize the intuition that Glauber dynamics mixes slowly.

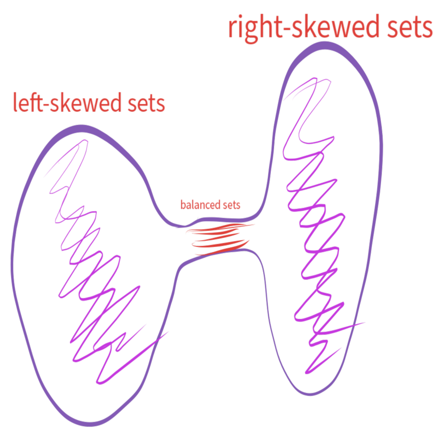

Figure 18.2: An illustration on the bottleneck to rapid mixing

Corollary 18.6 (Mossel-Weitz-Wormald [MWW09])**.** *For any* $\lambda > \lambda_c(\Delta)$*, the Glauber dynamics for sampling independent sets with partition function* $Z_{\boldsymbol{G}}(\lambda)$ *in a random bounded-degree bipartite graph* $G \sim \mathbf{M}(\Delta)$ *has mixing time* $\exp(\Omega(n))$*, with high probability over* $\boldsymbol{G}$*.*

To prove this, we'll need a simple result formalizing the above intuition about bottlenecks in Markov chains. Recall that the *conductance* of a Markov chain P with stationary distribution π is defined as

$$\Phi := \min_{S \subseteq V: 0 \leq \pi(S) \leq 1/2} \frac{C(S,\overline{S})}{\pi(S)}, \tag{18.11}$$

where $C(S,\overline{S}) = \sum_{x \in S, y \in \overline{S}} \pi(x)P(x,y)$. Note that the quotient in (18.11) has a natural interpretation as the probability that the Markov chain, started in the stationary distribution, escapes from S conditional on starting in S. Hence we might expect that an upper bound on Φ leads to a lower bound on the mixing time.

(As we saw in Lecture 11, Φ is also "almost" the dual of the multicommodity flow quantity ρ we used to prove upper bounds on mixing times; in contrast to flows, Φ is much more useful for proving *lower* bounds, because just as any flow gives us an upper bound on ρ (and hence an upper bound on mixing time), any cut $(S, \bar{S})$ gives us an upper bound on Φ (and hence a lower bound on mixing time).)

The following simple fact gives a lower bound on mixing time in terms of Φ. (Note that this is much easier—though also more useful in our context—than the more familiar upper bound, which is known as Cheeger's inequality. For our purposes Cheeger's inequality is subsumed by the flow bound in Lecture 11.)

Proposition 18.7. *For any Markov chain, the mixing time is bounded below by* $\tau_{\text{mix}} \geq \frac{1}{10\Phi}$.

Proof. Let S with $\pi(S) \leq \frac{1}{2}$ be a set that achieves the minimum in the definition of conductance. Consider the initial distribution

$$p^{(0)}(x) = \begin{cases} \frac{\pi(x)}{\pi(S)} & x \in S; \\ 0 & \text{otherwise.} \end{cases}$$

Then after one step of the chain we have

$$\begin{aligned} \|p^{(1)} - p^{(0)}\|_{\text{TV}} &= \max_A |p^{(1)}(A) - p^{(0)}(A)| \\ &= p^{(1)}(\bar{S}) - p^{(0)}(\bar{S}) \\ &= \sum_{x \in S} \sum_{y \in \bar{S}} p^{(0)}(x) P(x, y) \\ &= \Phi. \end{aligned}$$

Now it's easy to check **[exercise!]** that $\|p^{(t+1)} - p^{(t)}\|_{\text{TV}} \leq \|p^{(t)} - p^{(t-1)}\|_{\text{TV}}$ for all t, since P is a contraction on probability distributions. Hence we deduce from above that $\|p^{(t)} - p^{(0)}\|_{\text{TV}} \leq t\Phi$. Thus we have

$$\|p^{(t)} - \pi\|_{\text{TV}} \geq \|p^{(0)} - \pi\|_{\text{TV}} - \|p^{(t)} - p^{(0)}\|_{\text{TV}} \geq \frac{1}{2} - t\Phi, \tag{18.12}$$

since $\|p^{(0)} - \pi\|_{\text{TV}} = \pi(\bar{S}) \geq \frac{1}{2}$. Finally, the last expression in (18.12) is at least $\frac{1}{e}$ provided $t \leq (\frac{1}{2} - \frac{1}{e})\Phi^{-1}$, which implies that $\tau_{\text{mix}} \geq (\frac{1}{2} - \frac{1}{e})\Phi^{-1} \geq \frac{1}{10}\Phi^{-1}$. ☐

Proof of Corollary 18.6. We give a lower bound on the conductance of the Glauber dynamics. Define three sets of independent sets in G:

$$S_L = \{I\colon\ |I \cap L| > |I \cap R|\},\ S_R = \{I\colon\ |I \cap R| > |I \cap L|\},\ S_0 = \{I\colon\ |I \cap L| = |I \cap R|\}.$$

We will show that, when $\lambda > \lambda_c(\Delta)$,

$$\max\left\{\frac{\pi(S_0)}{\pi(S_L)}, \frac{\pi(S_0)}{\pi(S_R)}\right\} = \exp(-\Omega(n)). \tag{18.13}$$

This will complete the proof, as follows. Note first that we may assume w.l.o.g. that $\pi(S_R) \leq \frac{1}{2}$ (else replace S_R by S_L below). Since in the Glauber dynamics every independent set $I \in S_R$ has neighbors either within S_R or in S_0, we can bound the conductance of the Markov chain as follows:

$$\Phi \leq \frac{C(S_R, \overline{S_R})}{\pi(S_R)} \leq \frac{\pi(S_0)}{\pi(S_R)} = \exp(-\Omega(n)).$$

Proposition 18.7 then implies that the mixing time is $\exp(\Omega(n))$.

To prove (18.13), fix any $\lambda > \lambda_c(\Delta)$. This parameter fixes the occupation probabilities $p_1 < p^* < p_2$. Then for almost every random $G \sim \mathbf{M}(\Delta)$ we have

$$\frac{1}{n}\mathrm{E}[Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)] \overset{\text{Theorem 18.5}}{\leq} Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda) \overset{\text{Markov's ineq.}}{\leq} n\mathrm{E}[Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)].$$

By Fact 18.4, we know that there exist constants M, m such that

$$\begin{aligned} (\alpha,\beta) = (p_1,p_2) \text{ or } (p_2,p_1) \quad &\Rightarrow \quad \mathrm{E}[Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)] = \exp(M \cdot n(1+o(1))); \\ \alpha = \beta \quad &\Rightarrow \quad \mathrm{E}[Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)] \le \exp(m \cdot n(1+o(1))), \end{aligned}$$

and that $M - m \ge \epsilon(\lambda)$. Here $\epsilon(\lambda)$ is a positive constant that depends on λ. Therefore we have, for almost every $G \sim \mathbf{M}(\Delta)$,

$$\frac{\pi(S_0)}{\pi(S_R)} \le \frac{\sum_{\alpha=\beta} Z_{\boldsymbol{G}}^{\alpha,\beta}(\lambda)}{Z_{\boldsymbol{G}}^{p_1,p_2}(\lambda)} \le \frac{n^2 \exp(m \cdot n(1+o(1)))}{\frac{1}{n}\exp(M \cdot n(1+o(1)))} = n^3 \exp(-\epsilon(\lambda)n(1+o(1))) = \exp(-\Omega(n)),$$

with an identical argument for S_L. □

Note: This argument can be extended to any dynamics that modifies $o(n)$ spins in one step. The only change is that we now define S_0, S_L, S_R to be the sets of independent sets that are close to being balanced, and those that are far from being balanced (dominated by L and R respectively). The rest of the proof follows the same logic. The details are left as an **exercise**.

18.4 Reduction from Max Cut: preliminaries

We are now going to show the much stronger negative result that when $\lambda > \lambda_c(\Delta)$, there does not exist a FPRAS for approximating the independent-set partition function $Z_{\boldsymbol{G}}(\lambda)$ unless NP = RP. This proof is from [Sly10]. It proceeds by reducing the Max Cut problem, which is NP-hard, to approximation of $Z_{\boldsymbol{G}}(\lambda)$. More specifically, we are going to show that, given an arbitrary input graph H for MaxCut, and any Δ and $\lambda > \lambda_c(\Delta)$, we can construct a graph H^G (where G is a (suitably modified) random bipartite graph of degree Δ embedded into H as a gadget), so that approximating $Z_{H^G}(\lambda)$ allows us to find a maximum cut in H with high probability.

18.4.1 Random bipartite graph

We start by defining a distribution over random bounded-degree bipartite graphs G.

Definition 18.8. *A bipartite graph G from $\mathbf{M}(\Delta, \theta, \psi)$ is sampled as follows:*

1. *G has $n + n^{\theta+\psi}$ vertices on each side of the bipartition. The sides are denoted by "phases" $\oplus$ and $\ominus$ respectively.*
2. *Partition the vertices within the same phase into sets W^+, U^+ (respectively, W^-, U^-) of sizes n and $n^{\theta+\psi}$. (It will turn out that $\theta + \psi < 1$ so $|U^\pm| = o(|W^\pm|)$.)*
3. *First add $\Delta - 1$ independent random perfect matchings between the two phases, and then add an additional perfect matching between W^+ and W^-. Discard multiple edges. (Note that vertices in U^+, U^- have maximum degree $\Delta - 1$.)*
4. *Construct n^θ disjoint Δ-regular trees, each with n^ψ leaves in U^+ (resp. U^-). We call the roots of these trees terminals, and denote them V^+, V^- respectively. Note that $|V^+| = |V^-| = n^\theta$, and that the degrees of the terminals are all $\Delta - 1$.*

This construction is illustrated in Figure 18.3.

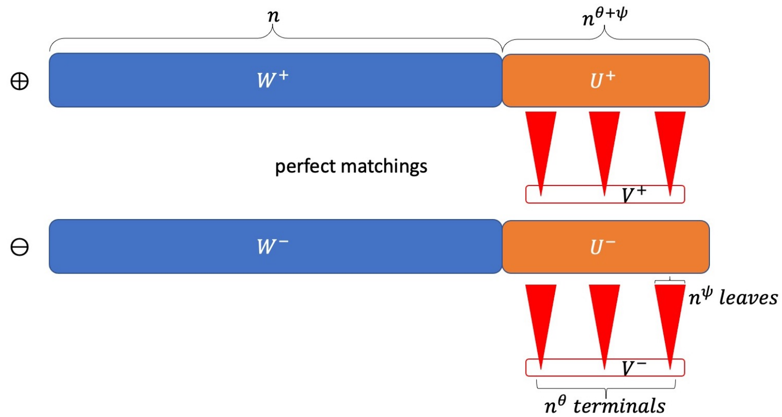

Figure 18.3: A random bipartite graph from $\mathbf{M}(n, \theta, \psi)$

We next define the *phase* of a configuration σ on G, and also two related product measures. (Note that a configuration σ is an assignment of spins 1 (occupied) or 0 (unoccupied) to vertices of G; each configuration corresponds to a unique independent set $I(\sigma) = \{v \in V : \sigma(v) = 1\}$.)

Definition 18.9. *The* phase *of a configuration σ on G is defined as*

$$S(\sigma) = \begin{cases} \oplus & \text{if } |I(\sigma) \cap W^+| > |I(\sigma) \cap W^-| \\ \ominus & o.w. \end{cases}$$

For a configuration σ, denote by σ_V the restriction of σ to the terminal vertices $V = V^+ \cup V^-$.

Definition 18.10. *The product measures $Q_V^{\pm}$ on configurations σ_V are defined by:*

$$Q_V^+(\sigma_V) = (p_2)^{|I(\sigma_V) \cap V^+|}(1-p_2)^{|V^+|-|I(\sigma_V)\cap V^+|}(p_1)^{|I(\sigma_V)\cap V^-|}(1-p_1)^{|V^-|-|I(\sigma_V)\cap V^-|},$$

$$Q_V^-(\sigma_V) = (p_1)^{|I(\sigma_V) \cap V^+|}(1-p_1)^{|V^+|-|I(\sigma_V)\cap V^+|}(p_2)^{|I(\sigma_V)\cap V^-|}(1-p_2)^{|V^-|-|I(\sigma_V)\cap V^-|}.$$

Thus Q_V^+ assigns spins to vertices in V according to independent Bernoulli trials, with occupation probabilities p_2 in V^+ and $p_1 < p_2$ in V^-. Q_V^- does the same with the roles of V^+, V^- reversed.

Using these definitions we can now state the following main technical lemma that will be used to validate the reduction.

Lemma 18.11 ([Sly10])**.** *For any Δ and $\lambda > \lambda_c(\Delta)$, there exist constants (θ, ψ) (depending on Δ and λ) such that the random bipartite graph $G \sim \mathbf{M}(\Delta, \theta, \psi)$ has $(2 + o(1))n$ vertices and satisfies the following properties w.h.p.:*

(i) In the Gibbs measure on configurations with partition function $Z_{\boldsymbol{G}}(\lambda)$, both phases occur with reasonable probability, i.e., $\Pr_\sigma[S(\sigma) = \oplus] \geq \frac{1}{n}$, $\Pr_\sigma[S(\sigma) = \ominus] \geq \frac{1}{n}$;

(ii) Conditioned on the phase S, the marginal of the Gibbs distribution on spins in V is very close to one of the production distributions $Q_V^{\pm}$, i.e.,

$$\max_{\sigma_V} \left| \frac{\pi[\sigma_V \mid S = \oplus]}{Q_V^+[\sigma_V]} - 1 \right| \leq n^{-2\theta}; \quad \max_{\sigma_V} \left| \frac{\pi[\sigma_V \mid S = \ominus]}{Q_V^-[\sigma_V]} - 1 \right| \leq n^{-2\theta}.$$

Note that property (ii) in this lemma is very strong: it says that, conditional on the phase, the terminal vertices of the gadget (which comprise a polynomial fraction of all the vertices in G) have essentially independent spins with a constant bias depending on the phase. This strong independence property (reflected in the L_∞ norm that appears in property (ii)) is guaranteed by the trees appended to the basic bipartite graph gadget: the vertices in $U^\pm$ also enjoy a weaker conditional independence property, which is sharpened by the trees due to the fact that, in the hard-core model on the Δ-regular tree, correlations from leaves to root decay very rapidly conditional on the odd/even measure. The details involved in proving Lemma 18.11 are quite substantial and we omit the proof here.

In the next section, we'll see how to use the above gadget G, and the properties in Lemma 18.11, to perform the reduction from MaxCut.

18.5 Sly's reduction

We will give a randomized reduction from computing the maximum cut of an arbitrary graph H to approximating the partition function $Z_{H^G}(\lambda)$ of a graph H^G of maximum degree Δ derived by combining H with a degree-Δ gadget G as defined in the previous section, at any choice of $\lambda > \lambda_c(\Delta)$. (The parameters θ, ψ of the graph G will depend on Δ and λ, as specified in Lemma 18.11.) Since MaxCut is NP-hard, this shows that an FPRAS for approximating the partition function of graphs of maximum degree Δ at any $\lambda > \lambda_c(\Delta)$ would imply that NP = RP.

Fix Δ and $\lambda > \lambda_c(\Delta)$, and let (θ, ψ) be the associated parameters for the random graphs G guaranteed by Lemma 18.11. Suppose H is an arbitrary graph (input to MaxCut) with $n^{\theta/4}$ vertices. We now define the graph H^G as follows:

- For each vertex $x \in V(H)$, create a disjoint copy of G and name it G_x. Let V_x^+ and V_x^- denote the respective terminal vertices of G_x. Denote the union of all the copies of G by $\widehat{H}^G$.
- For each edge $\{x, y\} \in E(H)$, add a matching of size $n^{3\theta/4}$ between V_x^+ and V_y^+, and similarly add such a matching between V_x^- and V_y^-. Note that since $|V_x^+| = |V_x^-| = n^\theta$, and the degree of H is at most $n^{\theta/4}$, we can make all these matchings disjoint, so no terminal has its degree increased by more than 1.

This yields a graph H^G of maximum degree Δ.

We now verify that sampling from the hard-core distribution in H^G at parameter λ allows us to find a maximum cut in H. Given a hard-core configuration σ on H^G, let $S = (S_x)_{x \in V(H)}$ denote the vector of phases in each of the subgraphs G_x. Note that each S naturally gives rise to a cut in H, defined by $\mathrm{Cut}(S) = \{\{x, y\} \in E(H) : S_x \neq S_y\}$. Also, define $Z_{H_G}[S]$ to be the partition function Z_{H^G} restricted to configurations with phase vector S.

The following consequence of Lemma 18.11 shows how the effect of the additional matching edges is to place larger weight on larger cuts.

Lemma 18.12. *Suppose G satisfies the two properties of Lemma 18.11. Then*

$$\frac{Z_{\widehat{H}^G}[S]}{Z_{\widehat{H}^G}} \geq n^{-n^{\theta/4}} \tag{18.14}$$

and

$$\frac{Z_{H^G}[S]}{Z_{\widehat{H}^G}[S]} = (\alpha(H) + o(1)) \left[\frac{(1 - p_1 p_2)^2}{(1 - p_1^2)(1 - p_2^2)} \right]^{n^{3\theta/4} |\mathrm{Cut}(S)|}, \tag{18.15}$$

where $\alpha(H) := \left[\left(1-p_1^2\right)\left(1-p_2^2\right)\right]^{n^{3\theta/4}|E(H)|}$.

Note that $(1-p_1p_2)^2 - (1-p_1^2)(1-p_2^2) = (p_1-p_2)^2 > 0$, so the weight factor in (18.15) is exponentially increasing with $|\operatorname{Cut}(S)|$.

Proof. For (18.14), note that since the graph $\widehat{H}^G$ consists of a collection of disjoint copies of G, the distribution on a configuration on $\widehat{H}^G$ is a product measure over configurations of $(G_x)_{x\in V(H)}$. This implies that the phases are independent. Hence, the claim is immediate from property (i) of Lemma 18.11.

For (18.15), the ratio on the left side is precisely the probability that a configuration σ with phase vector S sampled from the hardcore distribution on $\widehat{H}^G$ is also an independent set of H^G. By property (ii) of Lemma 18.11, conditional on the phases, the spins on the terminals of each G_x are (almost) independent with probabilities p_1, p_2 respectively. Now for any given edge $\{x,y\} \in E(H)$, let $E_{x,y}$ be the event that none of the matching edges between G_x, G_y has both endpoints in an independent set in $\widehat{H}^G$. Then we have:

$$\frac{Z_{H^G}[S]}{Z_{\widehat{H}^G}[S]} = (1+o(1)) \prod_{\{x,y\}\in E(H)} \Pr\left[E_{x,y}|S\right]. \tag{18.16}$$

Here, we consider two cases. By our construction of H^G and property (ii) of Lemma 18.11:

- if $S_x = S_y$, then conditional on S, the probability that each matching edge between V_x^+, V_y^+ (respectively, between V_x^-, V_y^-) does not have both its endpoints occupied is approximately $(1-p_1^2)$ (respectively, $\sim (1-p_2^2)$).
- if $S_x \neq S_y$, then this same conditional probability for both V_x^+, V_y^+ and V_x^-, V_y^- is approximately $(1-p_1p_2)^2$.

It follows that in the product term of (18.16), we get a factor of

$$\left[\left(1-p_1^2\right)\left(1-p_2^2\right)\right]^{n^{3\theta/4}}$$

for each non-cut edge of H, and a factor of

$$\left[(1-p_1p_2)^2\right]^{n^{3\theta/4}}$$

for each cut edge of H. Combining these two estimates, and pulling out a factor of $\alpha(H) = \alpha(H) := \left[\left(1-p_1^2\right)\left(1-p_2^2\right)\right]^{n^{3\theta/4}|E(H)|}$, yields (18.15). The $o(1)$ error term comes from absorbing the $n^{-2\theta}$ errors in property (ii) of Lemma 18.11. ☐

Armed with Lemma 18.12, we can now prove that the reduction works. Suppose that we have an FPRAS for $Z_{G^H}(\lambda)$ at the (arbitrary) value $\lambda > \lambda_c(\Delta)$ used to construct G. Then by standard methods we can sample independent sets in H^G in polynomial time with very small error (which we shall ignore). We identify the sampled independent sets with their phase vectors, which in turn correspond to cuts in H.

Now consider two cuts, C, C' in H, with $|C| > |C'|$, corresponding to phase vectors S, S'. The ratio of sampling probabilities for these cuts is

$$\begin{aligned}
\frac{Z_{H^G}[S]}{Z_{H^G}[S']} &= (1+o(1)) \times \frac{Z_{\widehat{H}^G}[S]}{Z_{\widehat{H}^G}[S']} \times \left[\frac{(1-p_1p_2)^2}{(1-p_1^2)(1-p_2^2)}\right]^{n^{3\theta/4}(|\operatorname{Cut}(S)|-|\operatorname{Cut}(S')|)} \\
&\geq (1+o(1))n^{-n^{\theta/4}} \left[\frac{(1-p_1p_2)^2}{(1-p_1^2)(1-p_2^2)}\right]^{n^{3\theta/4}} \\
&\geq 4^{n^{\theta/4}},
\end{aligned}$$

for sufficiently large n. The first line here comes from (18.15), the second from (18.14), and the third uses the fact observed earlier that $\frac{(1-p_1p_2)^2}{(1-p_1^2)(1-p_2^2)} = (1+\varepsilon)$ for some $\varepsilon > 0$. But the total number of cuts in H is only $2^{|V(H)|} = 2^{n^{\theta/4}}$, so the probability that the algorithm outputs a maximum cut is at least

$$\frac{4^{n^{\theta/4}}}{4^{n^{\theta/4}} + 2^{n^{\theta/4}}} = 1 - 2^{-|V(H)|}.$$

Thus with very high probability we solve MaxCut (exactly) in H in polynomial time. This concludes the proof of Sly's theorem.

References

[DFJ99] M. Dyer, A. Frieze, and M. Jerrum. On counting independent sets in sparse graphs. *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 210–217, 1999.

[GGŠ+14] A. Galanis, Q. Ge, D. Štefankovič, E. Vigoda, and Y. Yin. Improved inapproximability results for counting independent sets in the hard-core model. *Random Structures & Algorithms*, 45:78–110, 2014.

[MWW09] E. Mossel, D. Weitz, and N. Wormald. On the hardness of sampling independent sets beyond the tree threshold. *Probability Theory and Related Fields*, 143:401–439, 2009.

[Sly10] A. Sly. Computational transition at the uniqueness threshold. *Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 287–296, 2010.