Lecture 19: November 3

Instructor: Alistair Sinclair *Scribes: Zitong Yang and Banghua Zhu*

19.1 Recap

Last time we began a discussion of the *hard-core model* (or *independent set model*), one of the classical models in statistical physics and combinatorics. Given a graph $G = (V, E)$, the configurations are the independent sets of vertices $I \subset V$ in G. We give a weight $w(I) = \lambda^{|I|}$ to each independent set, so that the partition function is

$$Z_G(\lambda) := \sum_I w(I) = \sum_k a_k \lambda^k,$$

where $\lambda > 0$ is a parameter (known as the "fugacity") and a_k is the number of independent sets of size k. This is a spin system (or Markov random field) with pairwise interactions between neighbors (edge potentials) that are hard constraints: if vertex v is included in the set (which we write as $\sigma(v) = 1$), then none of its neighbors is allowed to be in the set ($\sigma(u) = 0$ for all $u \sim v$). There is also a vertex potential that assigns a factor of λ for each vertex v with $\sigma(v) = 1$.

We also define the associated Gibbs distribution:

$$\pi_\lambda(I) := \frac{w(I)}{Z_G(\lambda)} = \frac{\lambda^{|I|}}{Z_G(\lambda)}$$

Last time we saw a very simple reduction that showed that approximating $Z_G(\lambda)$ for general graphs at any fixed value $\lambda > 0$ is not possible in poynomial time unless $\mathsf{NP} = \mathsf{RP}$. However, when we restrict attention to graphs of bounded degree, which are typically those we encounter in applications in statistical physics and graphical models, the problem becomes interesting.

The goal for the next couple of lectures is to prove the following major theorem, which we stated last time.

Theorem 19.1. *For graphs $G = (V, E)$ of maximum degree $\Delta \geq 3$, there exists a threshold value $\lambda_c(\Delta) = \frac{(\Delta-1)^{(\Delta-1)}}{(\Delta-2)^{\Delta}}$ s.t.*

1. [Wei06] *if $\lambda < \lambda_c(\Delta)$, then there exists an FPTAS (deterministic!) for $Z_G(\lambda)$;*
2. [Sly10] *if $\lambda > \lambda_c(\Delta)$, then there does not exist an FPRAS for $Z_G(\lambda)$ unless $\mathsf{NP} = \mathsf{RP}$.*

Here $\lambda_c(\Delta) = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$ is the uniqueness threshold for the infinite Δ-regular tree, as discussed in the last lecture. Recall that, for $\lambda < \lambda_c(\Delta)$, there is a unique Gibbs measure on the infinite Δ-regular tree; while for $\lambda > \lambda_c(\Delta)$ there are two such (extremal) Gibbs measures, corresponding to occupation of the odd/even sites respectively.

Recall also that uniqueness of the Gibbs measure is equivalent to *Weak Spatial Mixing*, defined as follows.

Definition 19.2. *Let $S \subset V$ be any subset of vertices of the infinite tree, and τ be any fixed configuration on S. Let $p_r^\tau = \Pr[$root occupied given $\tau]$. Then* weak spatial mixing (WSM) *holds if there exists a constant $c > 0$ s.t.*

$$|p_r^\tau - p_r^{\tau'}| \le \exp(-c \cdot dist(r, s)).$$

for any two configurations τ, τ' on S.

19.2 The uniqueness threshold $\lambda_c(\Delta)$.

Let $d = \Delta - 1$ (so that we're working with the infinite d-ary tree). Then

$$\lambda_c(\Delta) = \frac{d^d}{(d-1)^{d+1}} \sim \frac{e}{\Delta},$$

as $d \to \infty$.

To understand where $\lambda_c(\Delta)$ comes from, we revisit a classical analysis of Spitzer [Spi75] and Kelly [Kel85]. Given an infinite d-ary tree T_v rooted[1] at v, define $p_v := \Pr_{T_v}[\sigma(v) = 0]$ be the probability of v being occupied. We can express p_v in terms of restricted partition functions (with obvious notation) as follows[2]:

$$p_v = \frac{Z_{T_v}[\sigma(v) = 0]}{Z_{T_v}[\sigma(v) = 0] + Z_{T_v}[\sigma(v) = 1]}.$$

Then we can use the tree structure to write these partition functions in terms of those of the subtrees rooted at the children $v_1, \dots, v_d$ of v:

$$p_v = \frac{\Pi_{i=1}^d Z_{T_{v_i}}}{\Pi_{i=1}^d Z_{T_{v_i}} + \lambda \Pi_{i=1}^d Z_{T_{v_i}}[\sigma(v_i) = 0]}.$$

To see this expression, note that the subtrees are independent given the value $\sigma(v)$ at the root, and when $\sigma(v) = 1$ we must have $\sigma(v_i) = 0$ for all i and we also pick up a factor of λ. Dividing all terms by the numerator, this becomes

$$p_v = \frac{1}{1 + \lambda \Pi_{i=1}^d p_{v_i}}. \tag{19.1}$$

Now if we postulate the existence of a translation-invariant Gibbs measure on the infinite tree (i.e., p_v is the same for all vertices v), then this value p_v must be a fixed point of the recurrence

$$f(x) = \frac{1}{1 + \lambda x^d}.$$

Denote the fixed point of $f(x)$ by x^* (i.e., $f(x^*) = x^*$. Note that x^* is unique because $f(x)$ is monotonically decreasing on $[0, 1]$. When λ is small, the (negative) slope of $f(x)$ at x^* is not very steep. This means that, if we iterate the recurrence, we will spiral in to x^*: i.e., x^* is a *stable* fixed point, and the recurrence will converge to it from any initial values at the leaves as the depth of the tree goes to infinity. Conversely, when λ is large the negative slope at x^* becomes so steep that the recurrence spirals away from the fixed point, which is unstable. So in this case the recurrence does not converge. See Figure 19.1 for a sketch of these behaviors.

[1] Technically, the infinite Δ-regular tree doesn't have a root; however, we will gloss over this minor detail as the error introduced by having a root is negligible.

[2] Again, technically these partition functions are infinite, but we can just think of a very large finite tree and take the limit.

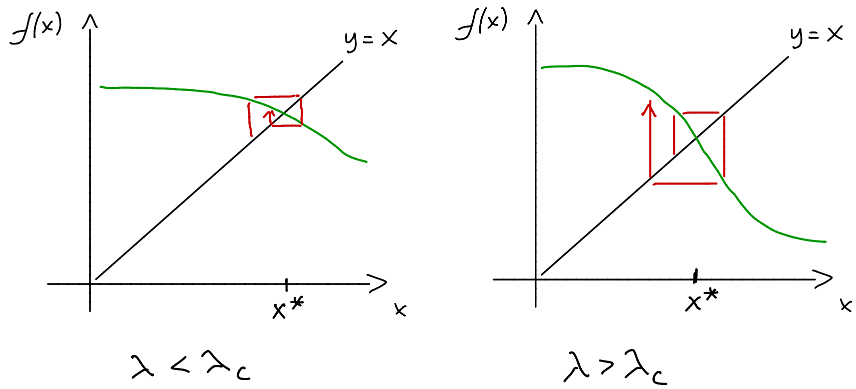

Figure 19.1: Sketch of graph of function f on $[0, 1]$, for small and large values of λ. The red trajectory shows the process of iterating f. For $\lambda < \lambda_c$ the fixed point x^* is stable, while for $\lambda > \lambda_c$ it is not.

It turns out that the feature that characterizes these two different behaviors is the value of the derivative of f' at the fixed point: if $f'(x^*) > -1$ then the fixed point is stable, while if $f'(x^*) < -1$ then it is unstable. And the critical value of λ at which this change in the derivative occurs is precisely $\lambda_c(\Delta)$.

One way to see this is to consider the two-step recurrence $f \circ f$, depicted in the two regimes in Figure 19.2. For small λ this function has just one fixed point (which is stable and coincides with the unique fixed point for f itself), while for large λ it has three fixed points: the (unstable) fixed point of f, together with two stable fixed points, corresponding to the odd and even Gibbs measures respectively. (In these measures, $f \circ f$ converges to two different possible values according to the initial value at the leaves. These measures are *semi-translation invariant*: all vertices on odd levels look the same, as do all vertices on even levels.) Note that these two behaviors are clearly characterized by whether or not $(f \circ f)'(x^*) < 1$.

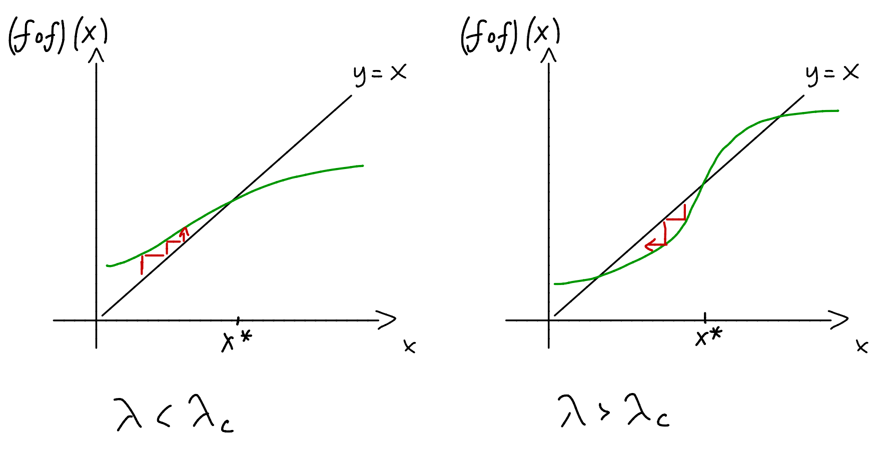

Figure 19.2: Sketch of graph of function $f \circ f$ on $[0, 1]$, for small and large values of λ. When $\lambda < \lambda_c$ there is a single (stable) fixed point, x^*; when $\lambda > \lambda_c$ the fixed point x^* becomes unstable and there are two additional stable fixed points.

To see the above claim about the derivative of f itself, note that

$$(f \circ f)'(x) = f'(f(x))f'(x)$$

At the fixed point x^* we get

$$(f \circ f)'(x^*) = f'(x^*)f'(x^*) = f'(x^*)^2.$$

So $(f \circ f)'(x^*) < 1$ is equivalent to $|f'(x^*)| < 1$, i.e., $f'(x^*) > -1$, as claimed above.

Finally, to see that $\lambda = \lambda_c(\Delta)$ is the threshold for these two regimes, we look at solutions of the following two equations:

$$x^* = f(x^*) \implies x^* = \frac{1}{1 + \lambda x^{*d}} \qquad \text{(the fixed point equation)}$$

and

$$f'(x^*) = -1 \implies \frac{-\lambda d x^{*(d-1)}}{(1 + \lambda x^{*d})^2} = -1 \qquad \text{(derivative condition)}$$

Plugging the first equation into the second simplifies it to $\lambda d x^{*(d+1)} = 1$. Now it is easy to check that the unique solution to these equations is $\lambda = \lambda_c(d) = \frac{d^d}{(d-1)^{d+1}}$. Moreover, $x^* = \frac{d-1}{d}$. Also, it's easy to verify (**exercise!**) that x^* is monotonically decreasing in λ, and hence that $f'(x^*)$ is monotonically decreasing in λ, which implies that $f'(x^*) > -1$ for $\lambda < \lambda_c(d)$ and $f'(x^*) < -1$ for $\lambda > \lambda_c(d)$, as claimed earlier.

19.3 An FPTAS when $\lambda < \lambda_c(\Delta)$

Now we go back to prove part (i) of Theorem 19.1, i.e., if $\lambda < \lambda_c(\Delta)$ then there exists an FPTAS for approximating $Z_G(\lambda)$. This is a direct result of Weitz's algorithm in [Wei06]. The outline of Weitz's algorithm is as follows:

1. Given a general graph G of maximum degree Δ, and an arbitrary vertex v, we construct the Self-Avoiding Walk Tree of G rooted at v, denoted $T_{SAW}(G, v)$. This tree will have the property that $\Pr[\sigma(v) = 1]$ is the same in G as in the tree! The key to the construction of the tree is the introduction of carefully chosen boundary conditions (fixed spins) on certain leaves.

2. Show that for $\lambda < \lambda_c(\Delta)$, the above correlation decay property (spatial mixing) that holds in the infinite Δ-regular tree continues to hold in this new tree $T_{SAW}(G, v)$. (In the presence of boundary conditions, this is not at all obvious.)

3. Use #1 and #2 to get an FPTAS by truncating T_{SAW} at depth $O(\log n)$, so that the computation of the marginal at the root v can be done exactly on the truncated tree, since it has polynomial size and no cycles. Property #2 then guarantees that the error due to truncation will be small.

We start with the construction of $T_{SAW}(G, v)$. This tree enumerates all *self-avoiding walks* in G starting at v; this means that we enumerate all walks, but we terminate each walk when it revisits a previous vertex (closes a cycle). At that point, we introduce a leaf with a certain boundary condition. (Trivial cycles of length 2 don't count, so we never revisit a vertex immediately after leaving it.)

Here is the rule for assigning boundary conditions. Suppose the tree encounters a cycle starting and ending at u, i.e., $u \to w_1 \to w_2 \to w_3 \cdots \to w_i \to u$ where the w_i are distinct vertices. Then we make the second occurrence of u a leaf, and we assign the boundary condition at this leaf according to the order in which we traverse the cycle: assuming some ordering on the neighbors of each vertex, then we assign the spin at the leaf to be 0 if $w_1 < w_l$ in this order at u, and 1 otherwise. Since each cycle can be traversed in two directions, each cycle will give rise to two leaves with opposite boundary conditions (corresponding to the two orders of traversal).

We illustrate this construction with a concrete example in Figure 19.3.

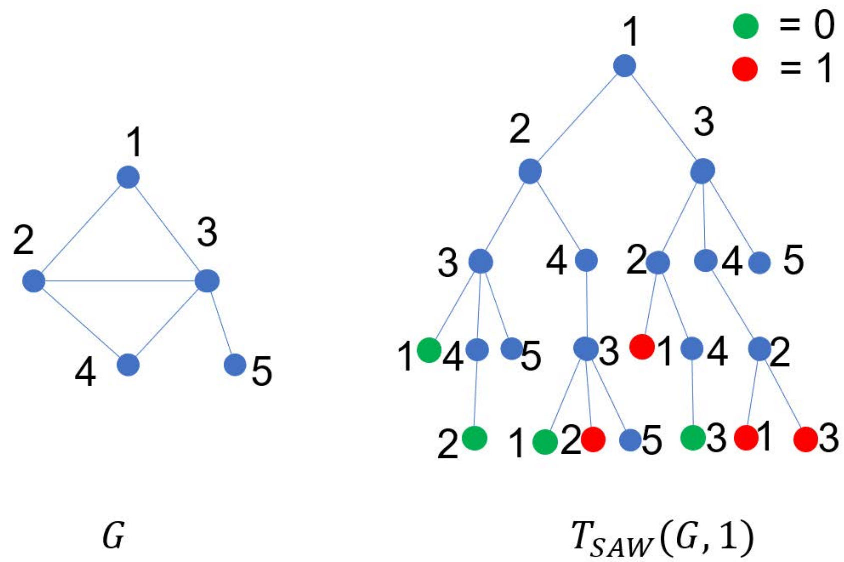

Figure 19.3: Illustration of the construction of the SAW tree $T_{SAW}(G,1)$ (right) from the graph G. The root-leaf paths of $T_{SAW}(G,1)$ correspond precisely to self-avoiding walks in G starting at vertex 1. When a walk terminates (closes a cycle), we assign either a 0 (green) or a 1 (red) boundary condition as specified by the rule in the text. Vertices that are leaves in G (like vertex 5) become leaves in the tree with no boundary condition.

Now we are ready to prove that with this construction, the probability of the root vertex being occupied in the tree is the same as that in the graph, which is the key guarantee in point #1 of the outline above.

Claim 19.3. $\Pr_G[\sigma(v)=0] = \Pr_{T_{SAW}(G,v)}[\sigma(v)=0]$.

Proof. Throughout the proof, we denote the SAW tree $T_{SAW}(G,v)$ by T_v, and the subtree rooted at any interior vertex u by T_u. Also, to simplify the algebra we will work with occupation ratios rather than probabilities. Denote

$$R_v = \frac{\Pr[\sigma(v)=1]}{\Pr[\sigma(v)=0]}.$$

It suffices to prove that $R_v^G = R_v^{T_v}$, since this clearly implies the claim in the theorem.

For the SAW tree, following essentially the same computation that led us to (19.1), we can write the following recurrence (**exercise!**) for the occupation ratio at the root in terms of the occupation ratios of the roots of the subtrees (denoted now by u_i):

$$R_v^{T_v} = \frac{Z_{T_v}[\sigma(v)=1]}{Z_{T_v}[\sigma(v)=0]} = \lambda \prod_{i=1}^{d} \frac{1}{1+R_{u_i}^{T_{u_i}}}. \tag{19.2}$$

Turning now to G, to get an expression for R_v^G we first split the vertex v into d copies, $v_1, v_2, \cdots, v_d$, where each v_i is connected only to u_i. This gives us a new graph G' in which all the v_i are leaves. Now we can write the occupation ratio in G as:

$$R_v^G = \frac{\Pr_G[\sigma(v)=1]}{\Pr_G[\sigma(v)=0]} = \lambda^{-(d-1)} \frac{\Pr_{G'}[\sigma(v_1)=\cdots=\sigma(v_d)=1]}{\Pr_{G'}[\sigma(v_1)=\cdots=\sigma(v_d)=0]}.$$

The factor $\lambda^{-(d-1)}$ comes from the fact that v being occupied in G gives only a factor of λ, while all the v_i being occupied in G' gives a factor of λ^d. We can further rewrite the above formula via a cascading product of conditional probabilities:

$$\begin{aligned} R_v^G = \lambda^{-(d-1)} & \frac{\Pr_{G'}[\sigma(u_1)=1|\sigma(v_2)=\cdots=\sigma(v_d)=0]}{\Pr_{G'}[\sigma(u_1)=0|\sigma(v_2)=\cdots=\sigma(v_d)=0]} \times \frac{\Pr_{G'}[\sigma(u_2)=1|\sigma(v_1)=1,\sigma(v_3)=\cdots=\sigma(v_d)=0]}{\Pr_{G'}[\sigma(u_2)=0|\sigma(v_1)=1,\sigma(v_3)=\cdots=\sigma(v_d)=0]} \\ & \times\cdots\times \frac{\Pr_{G'}[\sigma(u_d)=1|\sigma(v_1)=\cdots=\sigma(v_{d-1})=1]}{\Pr_{G'}[\sigma(u_d)=0|\sigma(v_1)=\cdots=\sigma(v_{d-1})=1]}. \end{aligned} \tag{19.3}$$

(To see this, use the fact that $\frac{\Pr[AB]}{\Pr[CD]} = \frac{\Pr[A|D]}{\Pr[C|D]} \cdot \frac{\Pr[B|A]}{\Pr[D|A]}$ for any four events A, B, C, D.) Now (19.3) can be written as

$$R_v^G = \lambda^{-(d-1)} \prod_{i=1}^{d} R_{v_i}^{G',\tau_i}, \tag{19.4}$$

where R_v^{G',τ_i} denotes the occupation ratio of v_i in G' with the boundary condition τ_i given by

$$\tau_i(v_j) = \begin{cases} 1 & j < i; \\ 0 & j > i. \end{cases} \tag{19.5}$$

Now since v_i is a leaf in G', we can write its occupation ratio just in terms of that of its single neighbor u_i, i.e.,

$$R_{v_i}^{G',\tau_i} = \frac{\lambda Z_{G'-v_i}^{\tau_i}[\sigma(u_i)=0]}{Z_{G'-v_i}^{\tau_i}[\sigma(u_i)=0] + Z_{G'-v_i}^{\tau_i}[\sigma(u_i)=1]} = \frac{\lambda}{1+R_{u_i}^{G'-v_i,\tau_i}}.$$

Plugging in the above formula to (19.4), we have

$$R_v^G = \lambda^{-(d-1)} \prod_{i=1}^{d} \frac{\lambda}{1+R_{u_i}^{G'-v_i,\tau_i}} = \lambda \prod_{i=1}^{d} \frac{1}{1+R_{u_i}^{G'-v_i,\tau_i}}. \tag{19.6}$$

But now comparing the expressions in (19.6) and (19.2), we see that they are identical in form! Moreover, it is not hard to check (**exercise!**) that the SAW tree of the graph $G'-v_i$ rooted at u_i, with boundary conditions τ_i at the leaves v_j, is precisely the same as the subtree T_{u_i} of T_v. (In particular, the boundary condition τ_i defined in (19.5) corresponds precisely to the rule for closing cycles at v specified in the construction of the SAW tree.) Thus inductively we can argue that $R_{u_i}^{G'-v_i,\tau_i} = R_{u_i}^{T_{u_i}}$, implying that indeed $R_v^G = R_v^{T_v}$, as required. □

Exercise: You are strongly encouraged to hand-turn the above proof on the toy example in Figure 19.3 to get a feel for how it works.

References

[Kel85] F.P. Kelly. Stochastic models of computer communication systems. *Journal of the Royal Statistical Society B*, 47:379–395, 1985.

[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.

[Spi75] F. Spitzer. Markov random fields on an infinite tree. *Annals of Probability*, 3:387–398, 1975.

[Wei06] D. Weitz. Counting independent sets up to the tree threshold. *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC)*, pages 140–149, 2006.