Cluster expansion

1. Show that there is a sequence of d-regular graphs G_d so that the smallest complex root (in complex absolute value) of $Z_{G_d}(\lambda)$ is $\Theta(1/d)$.

2. Fix $\Delta > 0$ and $0 < \lambda < \frac{1}{e(\Delta+1)}$. Let G_n be a sequence of n vertex graphs of max degree Δ. Let X_n be the size of a random independent set drawn from the hard-core model on G_n at activity λ.

   (a) Prove that $\mathrm{var}(X_n) = \Omega(n)$. Hint: use the law of total variance and the fact that G has a linear sized set of vertices at pairwise distance at least 3.

   (b) Prove that $\mathrm{var}(X_n) = O(n)$. Hint: use cluster expansion convergence.

   (c) For $k \geq 3$ fixed, prove an asymptotic upper bound on the kth cumulant of X, $\kappa_k(X)$.

   (d) Deduce that X is asymptotically normal; that is, $(X - \mathbb{E}X)/\sqrt{\mathrm{var}(X)} \Rightarrow N(0,1)$.

   (e) Write a formula using the cluster expansion for the cumulant generating function of X, $\log \mathbb{E}e^{tX}$. For what t does this converge?

   (f) Using the previous result prove a large deviation result for X, i.e. the best upper bound you can on the probability

$$\Pr(X \geq (1+\delta)\mathbb{E}X)\,.$$

3. Prove that the clique K_{d+1} has the highest triangle density (number of triangles divided by number of vertices) of any d-regular graph, and that there is gap to any graph that does not contain a K_{d+1} component.

4. Use the previous result and the cluster expansion for the generating function of matchings (monomer-dimer partition function) to prove that for some $\lambda^*(d) > 0$, all $0 < \lambda < \lambda^*$ and all d-regular graphs G not containing a K_{d+1} component we have

$$\frac{1}{|V(G)|}\log Z_G^{\mathrm{match}}(\lambda) > \frac{1}{d+1}\log Z_{K_{d+1}}^{\mathrm{match}}(\lambda)\,.$$

   The monomer-dimer partition function is

$$Z_G^{\mathrm{match}}(\lambda) = \sum_{M \in \mathcal{M}(G)} \lambda^{|M|}$$

   where the sum is over all matchings of G. A matching is an independent set in $L(G)$, the line graph of G.

   (Harder) Can you extend this to prove that for all d-regular G and all $0 < \lambda < \lambda^*$

$$\frac{1}{|V(G)|}\log Z_G^{\mathrm{match}}(\lambda) \geq \frac{1}{d+1}\log Z_{K_{d+1}}^{\mathrm{match}}(\lambda)\,.$$

5. Let G be a Δ-regular bipartite graph on bipartition (L, R) each of size n, with $\Delta = cn^{1/3}$. Suppose each vertex in L has fugacity $\lambda_L = \ell n^{-2/3}$ and each vertex in R has fugacity $\lambda_R = rn^{-2/3}$. Let $\mathbf{I}_L, \mathbf{I}_R$ be the number of occupied vertices in L and R respectively.

(a) Use the cluster expansion to write an asymptotic formula for $\mathbb{E}|\mathbf{I}_L|$ and $\mathbb{E}|\mathbf{I}_R|$.

(b) Use the cluster expansion to write an asymptotic formula for $\mathrm{cov}(\mathbf{I}_L, \mathbf{I}_R)$.

(c) Prove that after suitable centering and scaling the random vector $(\mathbf{I}_L, \mathbf{I}_R)$ converges to a bivariate Gaussian.

6. Let G be a biregular, bipartite graph with bipartition (L, R) and suppose every vertex in L has degree Δ_L and every vertex in R has degree Δ_R. Assume that $\Delta_R > \Delta_L$.

   (a) When Δ_R is much bigger than Δ_L, what do you expect typical uniformly random independent sets from G to look like?

   (b) Write the hard-core partition function of G as the partition function of a polymer model measuring deviations from the generalized ground state of the independents sets with no vertex from R.

   (c) How large must Δ_R be as a function of Δ_L to guarantee convergence of the cluster expansion for the polymer model when $\lambda = 1$?

7. (Total variation distance) For $m \geq n$ consider the following distributions of configurations of m balls in n labeled bins: 1) place each of the m balls independently in uniformly chosen random bins; 2) start with one ball in each bin and place the remaining $m - n$ balls independently in uniformly chosen random bins. Call the two distribution μ_1, μ_2 respectively (both depend on n and m).

   (a) Find good strategy for the following game: I pick μ_1 or μ_2 with probability $1/2$ each and show you one sample from the given distribution; from the sample you have to guess which distribution it came from. For what $m = m(n)$ can you win this game with probability $1 - o(1)$?

   (b) What does the strategy and probability of winning have to do with $\|\mu_1 - \mu_2\|_{TV}$?

   (c) Can you find the optimal threshold in $m = m(n)$ for $\|\mu_1 - \mu_2\|_{TV} \to 0$?