Stanford University	**Spring 2016**
Math 233: Non-constructive methods in combinatorics	**Instructor: Jan Vondrák**
Lecture date: Apr 13, 2016	**Scribe: Erik Bates**

Lecture 7. The Cluster Expansion Lemma

We have seen the Lovász Local Lemma and its stronger variant, Shearer's Lemma, which is unfortunately quite unwieldy in applications. Quite recently, researchers in mathematical physics discovered an intermediate form of the lemma, which seems to give results close to Shearer's Lemma but it is much more easily applicable. First let us review a connection between Shearer's Lemma and statistical physics which inspired this development.

7.1 The hard core model

For a graph G on n vertices, the *hard core repulsive gas model* is a model where particles can appear on the vertices of G, in such a way that neighboring vertices are not simultaneously occupied. (G can be arbitrarily but typically, regular lattice graphs are of interest in physics.) The set of admissible configurations is thus identified with the independent sets of G. There is a *fugacity* parameter λ_v associated with each vertex v, and the probability of a configuration $I \in \text{Ind}(G)$ (that is, particles are placed exactly at the vertices in I) is

$$\mathbb{P}(I) = \frac{1}{Z_G} \prod_{v \in I} \lambda_v,$$

where Z_G is the *partition function*

$$Z_G(\lambda_1, \lambda_2, \ldots, \lambda_n) = \sum_{I \in \text{Ind}(G)} \prod_{v \in I} \lambda_v.$$

Observe that Z_G is the multivariate generating function of independent sets, just like Shearer's polynomials (except for the alternating signs). An object of interest is the Taylor expansion of $\log Z_G$ around 0, also called the *Mayer expansion.* Various sufficient conditions for the absolute convergence of this series have been identified over the years. It eventually became apparent that these conditions are related to the local lemmas we have discussed. In particular, a condition presented in (Dobrushin, 1996) is:

The Mayer expansion is absolutely convergent for all $|\lambda_i| \leq p_i$, if there are $y_1, y_2, \ldots, y_n > 0$ such that

$$p_i \leq \frac{y_i}{\prod_{j \in \Gamma^+(i)} (1 + y_j)} \quad \forall\, i = 1, 2, \ldots, n.$$

If we substitute $y_i = \frac{x_i}{1 - x_i}$, $x_i \in (0, 1)$, we get the LLL condition:

$$\frac{y_i}{\prod_{j \in \Gamma^+(i)} (1 + y_j)} = \frac{x_i}{1 - x_i} \cdot \frac{1}{\prod_{j \in \Gamma^+(i)} \frac{1}{1 - x_j}} = x_i \prod_{j \in \Gamma(i)} (1 - x_j).$$

(Scott and Sokal, 2005) clarified the connection between the Mayer expansion and the LLL: The Mayer expansion is absolutely convergent for all $|\lambda_i| \leq p_i$ if and only if Shearer's conditions are satisfied for $p_1, p_2, \ldots, p_n$ and the graph G. Hence Dobrushin's result follows from the fact that LLL conditions imply Shearer's conditions.

7.2 Cluster expansion conditions

The cluster expansion lemma (Bissacot et al., 2011) is an intermediate form of the LLL, stronger than the LLL but weaker than Shearer's Lemma. Hence, we have the following ordering from weaker to stronger:

1. Symmetric LLL
2. Asymmetric LLL
3. Cluster expansion local lemma (CLL)
4. Shearer's Lemma

Theorem 7.1 (CLL) *Let $\mathcal{E}_1, \mathcal{E}_2, \ldots, \mathcal{E}_n$ be events on some probability space, with a (negative) dependency graph G, and*

$$\mathbb{P}(\mathcal{E}_i) \leq p_i.$$

If there exist $y_1, y_2, \ldots, y_n > 0$ such that

$$p_i \leq \frac{y_i}{\sum\limits_{\substack{I \subseteq \Gamma^+(i) \\ I \in \mathrm{Ind}(G)}} \prod\limits_{j \in I} y_j} \quad \forall\, i = 1, 2, \ldots, n, \tag{7.1}$$

then

$$\mathbb{P}\left(\bigcap_{i=1}^{n} \overline{\mathcal{E}}_i\right) > 0.$$

We can make a quick comparison to see that this result is, in fact, stronger than the LLL:

$$\frac{y_i}{\sum\limits_{\substack{I \subseteq \Gamma^+(i) \\ I \in \mathrm{Ind}(G)}} \prod\limits_{j \in I} y_j} \geq \frac{y_i}{\sum\limits_{S \subseteq \Gamma^+(i)} \prod\limits_{j \in S} y_j} = \frac{y_i}{\prod\limits_{j \in \Gamma^+(i)} (1 + y_j)},$$

where the latter quantity is the bound required in the LLL, as discussed above. Hence, the gain in CLL is that we sum up only over independent subsets of the neighborhood, as opposed to all subsets.

The cluster expansion lemma was originally proved in (Bissacot et al., 2011) by analytic arguments involving the convergence of the Mayer expansion. Here we present a combinatorial proof from (Harvey and Vondrak , 2015). We will show that Shearer's conditions are implied by the hypotheses. This way we obtain a relationship between Shearer's coefficients and the y_i coefficients in CLL. We note that if one wishes to avoid Shearer's Lemma, one can replace $\breve{q}_S$ by $\overline{P}_S = \Pr[\bigcap_{i \in S} \overline{\mathcal{E}}_i]$

and obtain a self-contained proof essentially without any change (only in (7.3) we get an inequality instead of an equality).

Proof of Theorem 7.1. For $S \subseteq V = V(G) = [n]$, define

$$Y_S := \sum_{\substack{I \subseteq S \\ I \in \mathrm{Ind}(G)}} y^I = \sum_{\substack{I \subseteq S \\ I \in \mathrm{Ind}(G)}} \prod_{i \in I} y_i.$$

With this notation, the CLL assumption (7.1) reads

$$p_i \leq \frac{y_i}{Y_{\Gamma^+(i)}}. \tag{7.2}$$

We make two observations: First, recall from Lectures 5 and 6 the identity

$$\breve{q}_S = \breve{q}_{S-a} - p_a \cdot \breve{q}_{S \setminus \Gamma^+(a)}. \tag{7.3}$$

An analogous identity for Y_S also holds:

$$Y_S = Y_{S-a} + y_a \cdot Y_{S \setminus \Gamma^+(a)}. \tag{7.4}$$

Indeed,

$$Y_S = \sum_{\substack{I \subseteq S \\ I \in \mathrm{Ind}(G)}} y^I = \sum_{\substack{I \subseteq (S-a) \\ I \in \mathrm{Ind}(G)}} y^I + \sum_{\substack{I \subseteq S \\ I \in \mathrm{Ind}(G) \\ a \in I}} y^I = Y_{S-a} + y_a \sum_{\substack{I \subseteq S \setminus \Gamma^+(a) \\ I \in \mathrm{Ind}(G)}} y_I = Y_{S-a} + y_a \cdot Y_{S \setminus \Gamma^+(a)}.$$

Second, observe that Y_S is sub-multiplicative in the following sense: If $S \cap T = \varnothing$, then

$$Y_{S \cup T} = \sum_{\substack{I \subseteq S \cup T \\ I \in \mathrm{Ind}(G)}} y^I \leq \sum_{\substack{I \subseteq S \\ I \in \mathrm{Ind}(G)}} \sum_{\substack{I \subseteq T \\ I \in \mathrm{Ind}(G)}} y^{I \cup J} = \sum_{\substack{I \subseteq S \\ I \in \mathrm{Ind}(G)}} y^I \sum_{\substack{I \subseteq T \\ I \in \mathrm{Ind}(G)}} y^J. = Y_S Y_T$$

The inequality holds because if $I \in \mathrm{Ind}(G)$, then $I \cap S$ and $I \cap T$ are also independent.

We will show by induction on $|S|$ that

$$\frac{\breve{q}_S}{\breve{q}_{S-a}} \geq \frac{Y_{V \setminus S}}{Y_{(V \setminus S)+a}} \quad \forall\, S \subseteq V, a \in S. \tag{7.5}$$

Note that Y is indexed by the sets complementary to the indices of $\breve{q}$. Thus the sets indexing Y shrink in the induction, while the sets indexing $\breve{q}$ grow. The reason for this can be traced back to (7.4) which contains a sign opposite to (7.3).

For the base case of $S = V$, we use (7.4) and (7.2) to write

$$Y_V = Y_{V-a} + y_a \cdot Y_{V \setminus \Gamma^+(a)} \geq Y_{V-a} + p_a \cdot Y_{\Gamma^+(a)} Y_{V \setminus \Gamma^+(a)} \geq Y_{V-a} + p_a \cdot Y_V.$$

The last inequality is a result of the sub-multiplicative property shown above. Hence

$$\frac{Y_{V-a}}{Y_V} \leq 1 - p_a = \frac{\breve{q}_a}{\breve{q}_\varnothing}.$$

Now assuming the inductive hypothesis, we check the general case:

$$\begin{aligned} Y_{(V\setminus S)+a} &= Y_{V\setminus S} + y_a \cdot Y_{(V\setminus S)\setminus\Gamma^+(a)} \\ &\geq Y_{V\setminus S} + p_a \cdot Y_{\Gamma^+(a)} Y_{(V\setminus S)\setminus\Gamma^+(a)} \\ &\geq Y_{V\setminus S} + p_a \cdot Y_{(V\setminus S)\cup\Gamma^+(a)}. \end{aligned}$$

Hence

$$\frac{Y_{V\setminus S}}{Y_{(V\setminus S)+a}} \leq 1 - p_a \cdot \frac{Y_{(V\setminus S)\cup\Gamma^+(a)}}{Y_{(V\setminus S)+a}} \leq 1 - p_a \cdot \frac{\breve{q}_{S\setminus\Gamma^+(a)}}{\breve{q}_{S-a}} = \frac{\breve{q}_S}{\breve{q}_{S-a}},$$

where the second inequality is the result of iterated applications of (7.5) (for lower cases of $|S|$): If $\Gamma^+(a) = \{a, a_1, a_2, \ldots, a_k\}$, then

$$\begin{aligned} \frac{Y_{(V\setminus S)\cup\Gamma^+(a)}}{Y_{(V\setminus S)+a}} &= \frac{Y_{(V\setminus S)\cup\Gamma^+(a)}}{Y_{(V\setminus S)\cup\Gamma^+(a)-a_k}} \cdot \frac{Y_{(V\setminus S)\cup\Gamma^+(a)-a_k}}{Y_{(V\setminus S)\cup\Gamma^+(a)-a_k-a_{k-1}}} \cdots \frac{Y_{(V\setminus S)+a+a_1}}{Y_{(V\setminus S)+a}} \\ &\geq \frac{\breve{q}_{S\setminus\Gamma^+(a)}}{\breve{q}_{(S\setminus\Gamma^+(a))+a_k}} \cdot \frac{\breve{q}_{(S\setminus\Gamma^+(a))+a_k}}{\breve{q}_{(S\setminus\Gamma^+(a))+a_k+a_{k-1}}} \cdots \frac{\breve{q}_{S-a-a_1}}{\breve{q}_{S-a}} \\ &= \frac{\breve{q}_{S\setminus\Gamma^+(a)}}{\breve{q}_{S-a}}. \end{aligned}$$

Now that (7.5) has been verified, we have (by another telescoping product)

$$\frac{\breve{q}_S}{\breve{q}_\varnothing} \geq \frac{Y_\varnothing}{Y_S} \geq \frac{1}{\sum\limits_{T\subseteq S} y^T} = \frac{1}{\prod\limits_{i\in S}(1+y_i)} > 0.$$

Shearer's conditions are thus satisfied, and so we may conclude $\mathbb{P}\left(\bigcap_{i=1}^n \overline{\mathcal{E}}_i\right) > 0$. □

7.3 Applications of cluster expansion

In practice, cluster expansion is applicable similarly to the LLL, but it gives better constants.

7.3.1 Latin transversals

Recall the problem of Latin transversals from Lecture 4. For a matrix $A \in \mathbb{Z}^{n\times n}$, we sampled a permutation $\pi \in \mathbb{S}_n$ uniformly at random. The events to be avoided were

$$\mathcal{E}_{i_1,j_1,i_2,j_2} = \{\pi \in \mathbb{S}_n : \pi(i_1) = j_1, \pi(i_2) = j_2\}, \quad i_1 \neq i_2, j_1 \neq j_2, A_{i_1 j_1} = A_{i_2 j_2}$$

so that

$$p := \mathbb{P}(\mathcal{E}_{i_1,j_1,i_2,j_2}) = \frac{1}{n(n-1)}.$$

A negative dependency graph G was defined by

$$\left\{\mathcal{E}_{i_1,j_1,i_2,j_2}, \mathcal{E}_{i_1',j_1',i_2',j_2'}\right\} \in E(G) \quad \Leftrightarrow \quad \{i_1,i_2\}\cap\{i_1',i_2'\} \neq \varnothing \quad \text{or} \quad \{j_1,j_2\}\cap\{j_1',j_2'\} \neq \varnothing.$$

If every integer appears as an entry of A at most k times, then it was argued that each event has degree in G of at most $4(n-1)k-1$. In particular, for a 4-tuple $a=(i_1,j_1,i_2,j_2)$, $\Gamma(a)$ is a union of four cliques, each of size at most nk. This means the independent subsets of $\Gamma(a)$ consist of either 0 or 1 vertex from each clique. In order to apply Theorem 7.1, then, it suffices that

$$\frac{1}{n(n-1)} \leq \frac{y}{(1+nky)^4}.$$

To determine the optimal value of y, we differentiate and solve

$$\begin{aligned}(1+nky)^4 - 4ynk(1+nky)^3 &= 0 \\ 1+nky &= 4nky \\ y &= \frac{1}{3nk}.\end{aligned}$$

That is, we can have

$$p \leq \frac{1/(3nk)}{(1+\frac{1}{3})^4} = \frac{3^3}{4^4}\cdot\frac{1}{nk} = \frac{27}{256}\cdot\frac{1}{nk},$$

meaning we need

$$k \leq \frac{27}{256}(n-1).$$

This is an improvement of the requirement derived when using the LLL:

$$k \leq \frac{n}{4e}.$$

7.3.2 Multipartite Turán problem

Here we can obtain a result almost as good as the one found using Shearer's Lemma. Recall from Lecture 6 that the dependency graph was the line graph $D=L(K_r)$. For an edge $\{i,j\}$ in K_r (i.e. a vertex in D), $\Gamma(\{i,j\})$ is the union of two independent cliques, those edges of K_r incident to i, and those edges incident to j. Hence

$$\frac{y}{\sum\limits_{\substack{I\subseteq\Gamma^+(\{i,j\})\\ I\in\mathrm{Ind}(G)}} y^I} \geq \frac{y}{y+(1+(r-2)y)^2} = \frac{1}{1+\frac{1}{y}(1+(r-2)y)^2},$$

where we have considered the independent subsets formed by taking just $\{i,j\}$ or by taking either 0 or 1 element from each of the two disjoint cliques of size $r-2$ (there remain $r-2$ edges incident to i, and the same is true for j). Once again, we optimize y by solving

$$\begin{aligned}-\frac{1}{y^2}(1+(r-2)y)^2 + \frac{2(r-2)}{y}(1+(r-2)y) &= 0 \\ 1+(r-2)y &= 2y(r-2) \\ y &= \frac{1}{r-2}.\end{aligned}$$

We can thus take

$$p \leq \frac{1}{1 + \frac{1}{y}(1 + (r-2)y)^2} = \frac{1}{1 + 4(r-2)} = \frac{1}{4r-7}.$$

Recall that the bound we obtained from the Heilmann-Lieb theorem was $p \leq \frac{1}{4r-8}$, and using more refined information about the roots of Hermite polynomials, it was enough to assume that $p \leq \frac{1}{4r - \Theta(r^{1/3})}$. So the bound from CLL is very close to what we can get from Shearer's Lemma.

References

Rodrigo Bissacot, Roberto Fernández, Aldo Procacci, and Benedetto Scoppola. An improvement of the lovász local lemma via cluster expansion. *Combinatorics, Probability and Computing*, 20 (05):709–719, 2011.

RL Dobrushin. Estimates of semi-invariants for the ising model at low temperatures. *Translations of the American Mathematical Society-Series 2*, 177:59–82, 1996.

Alexander D Scott and Alan D Sokal. The repulsive lattice gas, the independent-set polynomial, and the lovász local lemma. *Journal of Statistical Physics*, 118(5-6):1151–1261, 2005.

Nicholas Harvey and Jan Vondrak. An algorithmic proof of the Lovász Local Lemma via resampling oracles. http://arxiv.org/abs/1504.02044, 2015.