Stanford University **Spring 2018**
Math 233A: Non-constructive methods in combinatorics **Instructor: Jan Vondrák**
Lecture date: January 18, 2018 **Original scribe: László Miklós Lovász**

Lecture 4. Shearer's Lemma

5.1 Introduction and motivation

Suppose we are given a dependency graph G with vertex set $V(G) = [n]$, and probabilities $p_1, p_2, ..., p_n$. The Local Lovász Lemma gives a sufficient condition for ensuring that a set of events $\mathcal{E}_1, \mathcal{E}_2, ..., \mathcal{E}_n$ with dependencies consistent with G and satisfying $\mathbb{P}(\mathcal{E}_i) \leq p_i$ does not cover the entire probability space, that is

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

This theorem is not sharp, however. Suppose, for example, that $G = K_n$, the complete graph on n vertices, and all the p_i are equal to p. Then the condition given by the LLL is that $p \leq \frac{1}{en}$. However, the union bound shows that $p < \frac{1}{n}$ is sufficient. Even for K_2 (two dependent events), the asymmetric LLL requires the condition $\sqrt{p_1} + \sqrt{p_2} \leq 1$ to conclude that $\mathbb{P}[\overline{\mathcal{E}}_1 \cap \overline{\mathcal{E}}_2] > 0$, while the correct condition is obviously $p_1 + p_2 < 1$. Shearer's Lemma gives a sufficient and necessary condition the conclusion that $\mathbb{P}[\bigcap_{i=1}^{n} \overline{\mathcal{E}}_i] > 0$. The downside is that, unlike the Local Lovász Lemma, the conditions of Shearer's Lemma are often very difficult to check.

5.2 Definitions and Statement of Lemma

In order to state Shearer's Lemma, we need to define some multivariate polynomials related to the independent sets of a graph. We will think of the graph G as being fixed, and then we will denote by $\mathbf{p}$ the vector $(p_i : i \in V(G))$.

Definition 5.1 *Given a graph G, we denote by* $\mathrm{Ind}(G)$ *the independent sets of G, that is*

$$\mathrm{Ind}(G) = \{I \subseteq V(G) : I \text{ contains no edges}\}.$$

Definition 5.2 *Given a graph G, we define a polynomial q_I over $\mathbb{R}^{V(G)}$ for any $I \subseteq V(G)$ as*

$$q_I(\mathbf{p}) = \sum_{\substack{J \in \mathrm{Ind}(G) \\ I \subseteq J}} (-1)^{|J \setminus I|} \prod_{j \in J} p_j.$$

Note that $q_I = 0$ unless $I \in \mathrm{Ind}(G)$.

Definition 5.3 *Given a graph G, we define a polynomials $\breve{q}_S$ over $\mathbb{R}^{V(G)}$ for any $S \subseteq V(G)$ as*

$$\breve{q}_S(\mathbf{p}) = \sum_{\substack{J \in \mathrm{Ind}(G) \\ J \subseteq S}} (-1)^{|J|} \prod_{j \in J} p_j.$$

We are now ready to state Shearer's Lemma.

Lemma 5.4 (Shearer's Lemma) *Given a graph G on n vertices and $\mathbf{p} \in (0,1)^n$, the following are equivalent.*

1. $\forall I \in \mathrm{Ind}(G) : q_I(\mathbf{p}) > 0$.
2. $\forall S \subseteq V(G) : \breve{q}_S(\mathbf{p}) > 0$.
3. *For any events $\mathcal{E}_1, \mathcal{E}_2, ..., \mathcal{E}_n$ for which G is a dependency graph, if*

$$\mathbb{P}\left[\mathcal{E}_i\right] \leq p_i$$

for each i, then

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

Moreover, if these conditions are satisfied, then

$$\mathbb{P}\left[\bigcap_{i=1}^{n} \overline{\mathcal{E}}_i\right] \geq q_\emptyset(\mathbf{p}) = \breve{q}_{V(G)}(\mathbf{p}).$$

Example 5.5 *Suppose again that $G = K_n$. Then the independent sets are the empty set and singleton sets. We have that the probability of avoiding $\bigcup_{i=1}^{n} \mathcal{E}_i$ is at least*

$$q_\emptyset(\mathbf{p}) = \sum_{I \in \mathrm{Ind}(K_n)} (-1)^{|I|} p^I = 1 - \sum_{i=1}^{n} p_i,$$

which is exactly the union bound. (For a singleton $\{i\}$, $q_{\{i\}} = p_i > 0$ by assumption.)

Example 5.6 *If G is the empty graph, then we have*

$$q_\emptyset(\mathbf{p}) = \prod_{i=1}^{n} (1 - p_i) > 0,$$

which is indeed the probability of avoiding n independent events.

5.3 Proof of Shearer's Lemma

As before, we denote

$$\overline{P}_S = \mathbb{P}\left[\bigcap_{i \in S} \overline{\mathcal{E}}_i\right].$$

We will also omit the argument $\mathbf{p}$ in the polynomials. Further, we use the following shorthand notation: $S - a = S \setminus \{a\}$, $S + a = S \cup \{a\}$, and $p^I = \prod_{i \in I} p_i$.

First, we show that Property 2 implies Property 3 (which is the main implication, analogous to the LLL).

Proof: [2 ⇒ 3] We will prove that in fact $\overline{P}_S \geq \breve{q}_S$ for all $S \subseteq [n]$. For $S = \emptyset$, both of these are clearly equal to 1. Inductively, we want to prove that for any $a \in S \subseteq [n]$, we have

$$\frac{\overline{P}_S}{\overline{P}_{S-a}} \geq \frac{\breve{q}_S}{\breve{q}_{S-a}}. \tag{5.1}$$

By a telescoping product, this implies that $\overline{P}_S \geq \breve{q}_S$ and the desired implication.

We will prove (5.1) by induction on the size of S. Pick any $a \in S$. As in the proof of the LLL, we know that

$$\overline{P}_S \geq \overline{P}_{S-a} - p_a \overline{P}_{S\setminus\Gamma^+(a)}. \tag{5.2}$$

Note that each independent set containing a consists of $\{a\}$ and an independent subset of $V(G) \setminus \Gamma^+(a)$. By separating independent sets based on whether they contain a or not, we have

$$\begin{aligned}\breve{q}_S(\mathbf{p}) = \sum_{\substack{J\in\mathrm{Ind}(G)\\ J\subseteq S}} (-1)^{|J|}p^J &= \sum_{\substack{J\in\mathrm{Ind}(G)\\ J\subseteq S-a}} (-1)^{|J|}p^J + \sum_{\substack{J'\in\mathrm{Ind}(G)\\ J'\subseteq S\setminus\Gamma^+(a)}} (-1)^{|J'+a|}p^{J'+a} \\ &= \breve{q}_{S-a}(\mathbf{p}) - p_a \sum_{\substack{J'\in\mathrm{Ind}(G)\\ J'\subseteq S\setminus\Gamma^+(a)}} (-1)^{|J'|}p^{J'} = \breve{q}_{S-a}(\mathbf{p}) - p_a\breve{q}_{S\setminus\Gamma^+(a)}(\mathbf{p}). \end{aligned} \tag{5.3}$$

Note the similarity between (5.2) and (5.3). By the inductive hypothesis, we can assume (5.1) for each element in $(S-a) \setminus (S \setminus \Gamma^+(a))$, and by a telescoping product we obtain

$$\frac{\overline{P}_{S\setminus\Gamma^+(a)}}{\overline{P}_{S-a}} \leq \frac{\breve{q}_{S\setminus\Gamma^+(a)}}{\breve{q}_{S-a}}.$$

Combining this with (5.2) and (5.3), we deduce that

$$\frac{\overline{P}_S}{\overline{P}_{S-a}} \geq 1 - p_a\frac{\overline{P}_{S\setminus\Gamma^+(a)}}{\overline{P}_{S-a}} \geq 1 - p_a\frac{\breve{q}_{S\setminus\Gamma^+(a)}}{\breve{q}_{S-a}} = \frac{\breve{q}_S}{\breve{q}_{S-a}}$$

which completes the proof. □

Proof:[1 ⇔ 2] Next, we show that Property 1 and 2 are equivalent. To show that 2 implies 1, we have

$$q_I(\mathbf{p}) = \sum_{\substack{L\in\mathrm{Ind}(G)\\ L\subseteq V\setminus\Gamma^+(I)}} (-1)^{|L|}p^I p^L = p^I\breve{q}_{V\setminus\Gamma^+(I)}(\mathbf{p}). \tag{5.4}$$

This is clearly positive if each p_i and $\breve{q}_S$ is positive. For the reverse, we claim that

$$\breve{q}_S = \sum_{\substack{I\in\mathrm{Ind}(G)\\ I\cap S=\emptyset}} q_I. \tag{5.5}$$

This is clearly positive if q_I is positive for every $I \in \mathrm{Ind}(G)$ (including $I = \emptyset$ which is always present in the sum). To prove (5.5), we use an inclusion-exclusion calculation:

$$\sum_{\substack{I\in\mathrm{Ind}(G)\\ I\cap S=\emptyset}} q_I(\mathbf{p}) = \sum_{\substack{I\in\mathrm{Ind}(G)\\ I\cap S=\emptyset}} \sum_{\substack{J\in\mathrm{Ind}(G)\\ I\subseteq J}} (-1)^{|J\setminus I|} p^J = \sum_{J\in\mathrm{Ind}(G)} p^J \sum_{I\subseteq J\setminus S} (-1)^{|J\setminus I|}$$

$$= \sum_{J\in\mathrm{Ind}(G)} (-1)^{|J|} p^J \sum_{I\subseteq J\setminus S} (-1)^{|I|} = \sum_{\substack{J\in\mathrm{Ind}(G)\\ J\setminus S=\emptyset}} (-1)^{|J|} p^J = \breve{q}_S(\mathbf{p})$$

since the summation $\sum_{I\subseteq J\setminus S}(-1)^{|I|}$ is zero unless $J \setminus S = \emptyset$. We have thus proven that Properties 1 and 2 are equivalent. □

Last, we need to show that Property 3 implies the other two. In other words, we want to show that if there exists an independent set I such that $q_I(\mathbf{p}) \leq 0$, then we can find events $\mathcal{E}_i, i \in V$ consistent with G, with $\mathbb{P}[\mathcal{E}_i] \leq p_i$, and the events cover the entire probability space. We also show that if Shearer's conditions are satisfied, then there is an instance such that $\Pr[\bigcap_{i=1}^n \mathcal{E}_i] = q_\emptyset(\mathbf{p})$; i.e., Shearer's lemma is tight.

Note that if $\mathbf{p} = 0$, then $\breve{q}_S(\mathbf{p}) = 1$ for each S. This means that, since we have a finite set of continuous functions (polynomials), there exists an ϵ such that if each $|p_i| \leq \epsilon$, then each $\breve{q}_S(\mathbf{p}) > 0$. If we further assume that for each i we have $0 < p_i < \epsilon$, then we know from (5.4) that we must in fact have $q_I(\mathbf{p}) > 0$ for each independent set I.

Let the Shearer region be

$$\mathcal{S} = \{\mathbf{p} \in (0,1)^n \mid \forall I \in \mathrm{Ind}(G); q_I(p) > 0\}.$$

Look at $\lambda\mathbf{p}$ for $\lambda \in (0,1)$. We know that there is $\delta > 0$ such that if $0 < \lambda < \delta$, then $\lambda\mathbf{p} \in \mathcal{S}$. Also, $\lambda\mathbf{p} \notin \mathcal{S}$ for some $\lambda > 0$ (for example at the point where $\lambda p_i \geq 1$ for some i). Let $\lambda^* = \sup\{\lambda : \lambda\mathbf{p} \in \mathcal{S}\}$, and $\mathbf{p}^* = \lambda^*\mathbf{p}$. Then, since $\mathcal{S}$ is clearly an open set, $\mathbf{p}^*$ is not in it. However, by continuity of the polynomials, we must have that for all I, $q_I(\mathbf{p}^*) \geq 0$ (and in addition, for all S we have $\breve{q}_S(\mathbf{p}^*) \geq 0$). From (5.3), we can see that for any $a \in S \subseteq V$, $0 \leq \breve{q}_S(\mathbf{p}) \leq \breve{q}_{S-a}(\mathbf{p})$. Since $\mathbf{p}^* \notin \mathcal{S}$, at least one of the polynomials is 0, and hence we must have $\breve{q}_V(\mathbf{p}^*) = q_\emptyset(\mathbf{p}^*) = 0$. Hence, for any point $\mathbf{p} \notin \mathcal{S}$, there is $\lambda \in [0,1]$ such that $q_\emptyset(\lambda p) = 0$.

Conversely, assume that $\mathbf{p} \in \mathcal{S}$ and consider $\mathbf{p}' \leq \mathbf{p}$ which differs in one coordinate i. We have

$$\breve{q}_S(\mathbf{p}') = \breve{q}_S(\mathbf{p}) - (p'_i - p_i) \cdot \breve{q}_{S\setminus\Gamma^+(i)}(\mathbf{p}) \geq \breve{q}_S(\mathbf{p})$$

by the positivity of the polynomials at $\mathbf{p}$. Hence, all the polynomials $\breve{q}_S$ are also positive at $\mathbf{p}'$. We can continue like this with other coordinates and prove that for any $\mathbf{p}' \leq \mathbf{p}$, $\breve{q}_S(\mathbf{p}') \geq \breve{q}_S(\mathbf{p})$. In particular this implies that $\breve{q}_S(\lambda\mathbf{p}) > 0$ for all $\lambda \in [0,1]$.

To summarize, for any point $\mathbf{p}$, either $\mathbf{p} \in \mathcal{S}$ in which case $q_\emptyset(\lambda\mathbf{p}) > 0$ for all $\lambda \in [0,1]$, or $\mathbf{p} \notin \mathcal{S}$ in which case $q_\emptyset(\lambda\mathbf{p}) = 0$ for some $\lambda \in [0,1]$.

Remark 5.7 *Considering the discussion above, we can now add a Property 4 equivalent to the conditions of Shearer's lemma:*

4. $\forall\, \lambda \in [0,1];\ q_\emptyset(\lambda\mathbf{p}) > 0.$

Clearly Property 3 is downward closed, so if $\mathbf{p} \notin \mathcal{S}$ and we construct a set of events for $\mathbf{p}^* = \lambda^* \mathbf{p}$ that disproves Property 3, then we have also constructed a set of events for $\mathbf{p}$ that disproves Property 3. We will thus replace $\mathbf{p}$ with $\mathbf{p}^*$ in what follows, and we will assume that each $q_I(\mathbf{p}) \geq 0$ and each $\breve{q}_S(\mathbf{p}) \geq 0$.

Shearer's tight instance. By the above, we assume that $q_S(\mathbf{p}) \geq 0$ and $\breve{q}_S(\mathbf{p}) \geq 0$ for all $S \subseteq V$. If we plug $S = \emptyset$ into (5.5), we obtain the equation

$$1 = \breve{q}_{\emptyset}(\mathbf{p}) = \sum_{I \in \mathrm{Ind}(G)} q_I(\mathbf{p}) = \sum_{S \subseteq V} q_S(\mathbf{p}).$$

Since each $q_S(\mathbf{p})$ is nonnegative, we can define a collection of events $\mathcal{E}_1, \ldots, \mathcal{E}_n$ such that:

$$\mathbb{P}\left[\bigcap_{i \in S} \mathcal{E}_i \cap \bigcap_{j \in V \setminus S} \overline{\mathcal{E}_j}\right] = q_S(\mathbf{p})$$

for every $S \subseteq V$. Note that this specifies fully a probability space with events $\mathcal{E}_1, \ldots, \mathcal{E}_n$. We have to prove that this is a probability distribution that is consistent with G, and each event $\mathcal{E}_i$ has probability p_i.

First, we claim the following

Claim 5.8 *For any independent set I, we have*

$$\mathbb{P}[I] := \mathbb{P}\left[\bigcap_{i \in I} \mathcal{E}_i\right] = \prod_{i \in I} p_i.$$

Proof: Let I be an independent set. Then

$$\mathbb{P}[I] = \sum_{\substack{J \in \mathrm{Ind}(G) \\ I \subseteq J}} q_J = \sum_{\substack{J \in \mathrm{Ind}(G) \\ I \subseteq J}} \sum_{\substack{\widetilde{J} \in \mathrm{Ind}(G) \\ J \subseteq \widetilde{J}}} (-1)^{|\widetilde{J} \setminus J|} p^{\widetilde{J}} = \sum_{\substack{\widetilde{J} \in \mathrm{Ind}(G) \\ I \subseteq \widetilde{J}}} p^{\widetilde{J}} \sum_{J : I \subseteq J \subseteq \widetilde{J}} (-1)^{|\widetilde{J} \setminus J|} = p^I$$

by the same inclusion-exclusion argument as before. □

In particular, we see that $\mathbb{P}[\mathcal{E}_i] = p_i$. We also clearly have the following, since $q_S(\mathbf{p}) = 0$ for $S \notin \mathrm{Ind}(G)$.

Claim 5.9 *For any $S \notin \mathrm{Ind}(G)$, we have*

$$\mathbb{P}[S] := \mathbb{P}\left[\bigcap_{i \in S} \mathcal{E}_i\right] = 0.$$

In particular, this shows

$$\{i, j\} \in E(G) \quad \Rightarrow \quad \mathcal{E}_i \cap \mathcal{E}_j = \emptyset.$$

Thus, the tight instance can be summarized by saying that neighboring events are disjoint and non-neighboring events are independent. Next, we show that G is indeed a dependency graph for these events.

Claim 5.10 *For any $i \in V$, $\mathcal{E}_i$ is mutually independent of $\{\mathcal{E}_j : j \in V \setminus \Gamma^+(i)\}$.*

Proof: Take any $J \subseteq V \setminus \Gamma^+(i)$. If J is an independent set, then $J \cup \{i\}$ is also independent. In this case, Claim 5.8 implies

$$\mathbb{P}\left[\mathcal{E}_i \cap \bigcap_{j \in J} \mathcal{E}_j\right] = \mathbb{P}\left[\bigcap_{j \in J \cup \{i\}} \mathcal{E}_j\right] = p_i \prod_{j \in J} p_j = \mathbb{P}(\mathcal{E}_i) \cdot \mathbb{P}\left[\bigcap_{j \in J} \mathcal{E}_j\right].$$

If J is not independent, then of course $J \cup \{i\}$ is not independent. In this case, Claim 5.9 gives

$$\mathbb{P}\left(\mathcal{E}_i \cap \bigcap_{j \in J} \mathcal{E}_j\right) = 0 = \mathbb{P}(\mathcal{E}_i) \cdot 0 = \mathbb{P}(\mathcal{E}_i) \cdot \mathbb{P}\left(\bigcap_{j \in J} \mathcal{E}_j\right).$$

The proof is completed by applying inclusion-exclusion: If $J, K \subseteq [n] \setminus \Gamma^+(i)$ are disjoint,

$$\begin{aligned}\mathbb{P}\left[\mathcal{E}_i \cap \bigcap_{j \in J} \mathcal{E}_j \cap \bigcap_{k \in K} \overline{\mathcal{E}}_k\right] &= \sum_{L \subseteq K} (-1)^L \, \mathbb{P}\left[\mathcal{E}_i \cap \bigcap_{j \in J \cup L} \mathcal{E}_j\right] \\ &= \mathbb{P}[\mathcal{E}_i] \sum_{L \subseteq K} (-1)^L \, \mathbb{P}\left[\bigcap_{j \in J \cup L} \mathcal{E}_j\right] \\ &= \mathbb{P}[\mathcal{E}_i] \cdot \mathbb{P}\left[\bigcap_{j \in J} \mathcal{E}_j \cap \bigcap_{k \in K} \overline{\mathcal{E}}_k\right].\end{aligned}$$

We conclude that $\mathcal{E}_i$ is independent of $\{\mathcal{E}_j : j \in V \setminus \Gamma^+(i)\}$. □

Finally, from Claim 5.8, we get by inclusion-exclusion:

$$\mathbb{P}\left[\bigcap_{j \in S} \overline{\mathcal{E}}_j\right] = \sum_{J \subseteq S} \underbrace{(-1)^{|J|} \cdot \mathbb{P}\left[\bigcap_{j \in J} \mathcal{E}_j\right]}_{\neq 0 \text{ only when } J \in \mathrm{Ind}(G)} = \sum_{\substack{J \subseteq S \\ J \in \mathrm{Ind}(G)}} (-1)^{|J|} p^J = \breve{q}_S(\mathbf{p}).$$

Thus the tight instance shows that the lower bounds $\mathbb{P}[\bigcap_{i \in S} \overline{\mathcal{E}}_i] \geq \breve{q}_S(\mathbf{p})$ are tight. In particular, the tight instance satisfies $\mathbb{P}[\bigcap_{i=1}^n \overline{\mathcal{E}}_i] = \breve{q}_V(\mathbf{p}) = q_\emptyset(\mathbf{p})$ which is 0 if Shearer's conditions are violated.