The Lovász Sieve

Assume that we have a family of "bad" events. How can we make sure that there is some non-zero probability that none of the bad events will happen? By the union bound $\Pr[A \cup B] \leq \Pr[A] + \Pr[B]$, this probability is non-zero if the sum of probabilities of all these bad events is smaller than 1. In one sense this is best possible: when bad events are pairwise disjoint, the condition cannot be weakened. If we know that the bad events are independent, we can get a much better bound, by multiplying all the probabilities that each single bad event does not happen. This will work as long as each bad event has probability smaller than 1. But this will immediately fail, if at least two of the bad events are not independent.

In such cases—when there is some relatively small amount of dependence between events—one can use a powerful generalization of the union bound, known as the *Lovász Local Lemma*.

1. The Lovász Local Lemma

An event A is *mutually independent* of a collection of events if conditioning on any sub-collection $B_1, \ldots, B_m$ of these events does not affect the probability of A, that is,

$$\Pr[A \,|\, C_1 \cdots C_m] = \Pr[A]$$

for all $C_i \in \{B_i, \overline{B}_i\}$, $i = 1, \ldots, m$. Note that A might be independent of *each* of the events $B_1, \ldots, B_m$, but not be mutually independent of them. To see this, consider flipping a fair coin twice and the three events: B_1, B_2, A, where B_i is the event that the i-th flip is a head and A is the event that both flips are the same. Then A is independent of B_1 and of B_2 but $\Pr[A \,|\, B_1 B_2] = 1$.

Let $A_1, \ldots, A_n$ be events. A graph $G = (V, E)$ on the set of vertices $V = \{1, \ldots, n\}$ is said to be a *dependency graph* if, for all i, A_i is mutually independent of all the events A_j such that j is *not* adjacent to i in G, i.e., for which $\{i, j\} \notin E$. We emphasize that A_i must not only be independent of each such A_j individually but also must be independent of any boolean combination of the A_j's. Such a graph G may be not uniquely defined, but we will not care about this. We will only be interested in the smallest possible degree of such a graph, which we call the *degree of dependence* of the events $A_1, \ldots, A_n$.

The following fact is known as the *Lovász Local Lemma*.

Lemma 19.1 (Erdős–Lovász 1975). *Let $A_1, \ldots, A_n$ be events with $\Pr[A_i] \leq p$ for all i, and let d be the degree of their dependence. If $\mathrm{e}p(d+1) \leq 1$ then $\Pr[\overline{A}_1 \overline{A}_2 \cdots \overline{A}_n] > 0$.*

As in the original proof of Erdős and Lovász, we will prove the lemma under the slightly stronger condition $4pd \leq 1$, and later show that the lemma remains true under weaker condition $\mathrm{e}p(d+1) \leq 1$, as well.

In the proof we will use two properties of the conditional probability which follow fairly easily from its definition as $\Pr[A \mid B] = \Pr[AB] / \Pr[B]$:

$$\Pr[A \mid BC] = \frac{\Pr[AB \mid C]}{\Pr[B \mid C]} \tag{114}$$

and

$$\Pr[A \mid BC] \cdot \Pr[B \mid C] \cdot \Pr[C] = \Pr[ABC] \,. \tag{115}$$

Proof (Spencer 1995). Fix a dependency graph G of our events of degree d. We prove by induction on m that for any m events (calling them $A_1, \ldots, A_m$ for convenience only)

$$\Pr[A_1 \,|\, \overline{A}_2 \cdots \overline{A}_m] \leq 2p.$$

For $m = 1$ this is obvious. Let $2, \ldots, k$ be the vertices from $\{2, \ldots, m\}$ which are adjacent to 1 in the dependency graph G. Using the identity (**114**), we can write

$$\Pr[A_1 \mid \overline{A}_2 \cdots \overline{A}_m] = \frac{\Pr[A_1 \overline{A}_2 \cdots \overline{A}_k \mid \overline{A}_{k+1} \cdots \overline{A}_m]}{\Pr[\overline{A}_2 \cdots \overline{A}_k \mid \overline{A}_{k+1} \cdots \overline{A}_m]}. \tag{116}$$

We bound the numerator

$$\begin{aligned} \Pr[A_1 \overline{A}_2 \cdots \overline{A}_k \mid \overline{A}_{k+1} \cdots \overline{A}_m] &\leq \Pr[A_1 \mid \overline{A}_{k+1} \cdots \overline{A}_m] \\ &= \Pr[A_1] \leq p \end{aligned}$$

since A_1 is mutually independent of $A_{k+1}, \ldots, A_m$. The denominator, on the other hand, can be bounded by the induction hypothesis

$$\begin{aligned} \Pr[\overline{A}_2 \cdots \overline{A}_k \mid \overline{A}_{k+1} \cdots \overline{A}_m] &= 1 - \Pr[A_2 \cup \cdots \cup A_k \mid \overline{A}_{k+1} \cdots \overline{A}_m] \\ &\geq 1 - \sum_{i=2}^{k} \Pr[A_i \mid \overline{A}_{k+1} \cdots \overline{A}_m] \\ &\geq 1 - 2p(k-1) \geq 1/2, \end{aligned}$$

because $k - 1 \leq d$ and $2pd \leq 1/2$. Thus

$$\Pr[A_1 \mid \overline{A}_2 \cdots \overline{A}_m] \leq p/(1/2) = 2p,$$

completing the induction. Finally, by (**??**), by (115)

$$\Pr[\overline{A}_1 \cdots \overline{A}_n] = \prod_{i=1}^{n} \Pr[\overline{A}_i \mid \overline{A}_1 \cdots \overline{A}_{i-1}] \geq (1-2p)^n > 0.$$

□

When the events A_i are not symmetric (i.e., when their probabilities might be very different) a more general form of the Lovász sieve is appropriate. This generalization is due to Spencer (1977).

LEMMA 19.2. *Let $G = (V, E)$ be a dependency graph of events $A_1, \ldots, A_n$. Suppose there exist real numbers $x_1, \ldots, x_n$, $0 \leq x_i < 1$, so that, for all i,*

$$\Pr[A_i] \leq x_i \cdot \prod_{\{i,j\} \in E} (1 - x_j).$$

Then

$$\Pr[\overline{A}_1 \overline{A}_2 \cdots \overline{A}_n] \geq \prod_{i=1}^{n} (1 - x_i).$$

In particular, with positive probability no event A_i holds.

PROOF. The induction hypothesis of the earlier proof is replaced by

$$\Pr[A_1 \mid \overline{A}_2 \cdots \overline{A}_m] \leq x_1, \tag{114}$$

and, using the same identity (**??**), the denominator of (**??**) is set equal to

$$\prod_{j=2}^{k} \Pr[\overline{A}_j \mid \overline{A}_{j+1} \cdots \overline{A}_m], \tag{116}$$

which by the induction hypothesis, is at least

$$\prod_{j=2}^{k} (1 - x_j) = \prod_{\{1,j\} \in E} (1 - x_j)\,.$$

□

The Lovász sieve works well when we have "much independence" between the events. In a similar vein, there is also an estimate, due to Razborov (1988), which works well if the events are "almost k-wise independent."

Let $A_1, \ldots, A_n$ be events, each of which appears with the same probability $\Pr[A_i] = p$. If *all* these events are mutually independent, then

$$\Pr\left[\bigcup_{i=1}^{n} A_i\right] = 1 - \Pr\left[\bigcap_{i=1}^{n} \overline{A}_i\right] = 1 - (1-p)^n \geq 1 - \mathrm{e}^{-pn}\,.$$

The mutual independence is a very strong requirement. It turns out that a reasonable estimate can be obtained also in the case when $\Pr\left[\bigcap_{i\in I} A_i\right]$ is only "near" to $p^{|I|}$ for the sets I of size up to some number k; in this case the events $A_1, \ldots, A_n$ are also called *almost k-wise independent*.

LEMMA 19.3 (Razborov 1988). *Let $n > 2k$ be any natural numbers, let $0 < p, \delta < 1$, and let $A_1, \ldots, A_n$ be events such that, for every subset $I \subseteq \{1, \ldots, n\}$ of size at most k,*

$$\left|\Pr\left[\bigcap_{i\in I} A_i\right] - p^{|I|}\right| \leq \delta.$$

Then

$$\Pr\left[\bigcup_{i=1}^{n} A_i\right] \geq 1 - \mathrm{e}^{-pn} - \binom{n}{k+1}(\delta k + p^k).$$

Note that if the events are k-wise independent, then $\delta = 0$ and the obtained estimate worse by an additive term $\binom{n}{k+1}p^k$ than that for mutual independence.

PROOF. Let us first consider the case where k is even. Let $B_1, \ldots, B_n$ be independent events, each having the success probability p. Applying the Bonferroni inequalities to $\Pr\left[\bigcup_{i=1}^{n} A_i\right]$ and $\Pr\left[\bigcup_{i=1}^{n} B_i\right]$, we obtain that

$$\Pr\left[\bigcup_{i=1}^{n} A_i\right] \geq \sum_{\nu=1}^{k} (-1)^{\nu+1} \sum_{|I|=\nu} \Pr\left[\bigcap_{i\in I} A_i\right] \tag{117}$$

and

$$\Pr\left[\bigcup_{i=1}^{n} B_i\right] \leq \sum_{\nu=1}^{k} (-1)^{\nu+1} \sum_{|I|=\nu} p^{|I|} + \sum_{|I|=k+1} p^{k+1}. \tag{118}$$

The assumption of the lemma that $A_1, \ldots, A_n$ are almost k-wise independent implies that the right-hand side in (117) is at least

$$\sum_{\nu=1}^{k} (-1)^{\nu+1} \sum_{|I|=\nu} p^{|I|} - \delta k \binom{n}{k}. \tag{119}$$

On the other hand, the independence of $B_1, \ldots, B_n$ implies that

$$\Pr\left[\bigcup_{i=1}^{n} B_i\right] = 1 - (1-p)^n \geq 1 - \mathrm{e}^{-pn}. \tag{120}$$

Combining (117), (118), (119) and (120) yields

$$\begin{aligned}\Pr\left[\bigcup_{i=1}^{n} A_i\right] &\geq 1 - \mathrm{e}^{-pn} - \delta k \binom{n}{k} - p^{k+1}\binom{n}{k+1} \\ &\geq 1 - \mathrm{e}^{-pn} - \binom{n}{k+1}(\delta k + p^{k+1}).\end{aligned}$$

In the case where k is odd, we use the above argument with $k-1$ substituted for k. □

2. Disjoint cycles

By a *digraph* we will mean a directed graph without parallel edges. Such a graph is *k-regular* if every vertex has exactly k outgoing edges.

THEOREM 19.4. *Every k-regular digraph has a collection of $\lfloor k/(3\ln k)\rfloor$ vertex-disjoint cycles.*

PROOF. Let $G=(V,E)$ be a k-regular digraph. Set $r := \lfloor k/(3\ln k)\rfloor$, and color the vertices uniformly at random using colors $\{1,\dots,r\}$. That is, each vertex v gets a particular color independently and with the same probability $1/r$. Let A_v be the event that v does not have any out-neighbor of the same color as v. (An out-neighbor of v is the second endpoint of an edge leaving v.) We need only to show that $\Pr[\cap_{v\in V}\overline{A_v}]>0$.

Since each vertex has k out-neighbors, we have that

$$\Pr[A_v] = \left(1-\frac{1}{r}\right)^k < \mathrm{e}^{-k/r} \le \mathrm{e}^{-3\ln k} = k^{-3}\,.$$

For a vertex v, let $N(v)$ be the set consisting of v and all its k out-neighbors. Then A_v is mutually independent of the events in $\{A_u : N(u)\cap N(v)=\emptyset\}$. Since this set contains at most $(k+1)^2$ events, the degree of dependence of the events A_v is $d\le (k+1)^2$. Hence, to apply the Lovász Local Lemma we only need that $4k^{-3}(k+1)^2\le 1$, which is true for $k\ge 6$. For $k<6$ the theorem is trivially true since then $r=1$. □

Alon, McDiarmid and Molloy (1996) proved that, in fact, $\Omega(k^2)$ vertex-disjoint cycles exist and conjectured that at least $\binom{k+1}{2}$ cycles should exist.

3. Colorings

A striking feature of the Lovász sieve is the lack of conditions on the total number n of events – only the degree of their dependence is important. This is particularly useful when dealing with large families whose members share not too many points in common. Let us demonstrate this with several typical examples.

First, let us consider 2-colorings of hypergraphs. Recall that a family of sets $\mathcal{F}$ is 2-*colorable* if it is possible to color the points of the underlying set in red and blue, so that no member of $\mathcal{F}$ is monochromatic. A family is *k-uniform* if all its members have size k.

In Chap. 3 (see Theorem 3.4) we proved that if the family $\mathcal{F}$ is relatively small then it is 2-colorable: Every k-uniform family of fewer than 2^{k-1} sets is 2-colorable.

Let us recall the argument. Suppose $\mathcal{F}$ is a k-uniform family with at most $2^{k-1}-1$ sets. Consider a random coloring, each element independently colored red or blue with probability 1/2. Any one member of $\mathcal{F}$ will then be monochromatic with probability $2\cdot 2^{-k}=2^{1-k}$, and so the probability that *some* member will be monochromatic, does not exceed $|\mathcal{F}|\cdot 2^{1-k}$, which is strictly smaller than 1. Therefore, at least one coloring must leave no member of $\mathcal{F}$ monochromatic.

Now suppose that $\mathcal{F}$ has more than 2^k members. Then the above random coloring will be doomed since the chances of it to be a proper 2-coloring will tend to zero. Fortunately, we do not require a high probability of success, just a positive probability of success. For example, if $\mathcal{F}$ is a family of m mutually disjoint k-element subsets of some set, then the events A_i="the i-th member of $\mathcal{F}$ is monochromatic" are mutually independent, and so the probability that none of them holds is exactly $\left(1-2^{-(k-1)}\right)^m$, which is positive no matter how large m is. Therefore, $\mathcal{F}$ is 2-colorable.

Of course for general families $\mathcal{F}$, the events $A_1,\dots,A_m$ are not independent as some pairs of members may intersect. In such situations the Lovász sieve shows its surprising power.

THEOREM 19.5 (Erdős–Lovász 1975). *If every member of a k-uniform family intersects at most 2^{k-3} other members, then the family is 2-colorable.*

PROOF. Suppose $\mathcal{F} = \{S_1, \dots, S_m\}$ is a family of k-element subsets of some set X. Consider a random coloring of X, each point independently colored red or blue with probability 1/2. Let A_i denote the event that S_i is monochromatic. Then $\Pr[A_i] = p$ where $p = 2(1/2)^{|S_i|} = 2^{1-k}$. Our goal is to show that $\Pr[\overline{A}_1 \cdots \overline{A}_m] > 0$. Define a dependency graph by joining A_i and A_j if and only if $S_i \cap S_j \neq \emptyset$. By the assumption, this graph has degree at most $d = 2^{k-3}$. Since $4dp = d2^{3-k} \leq 1$, Lemma **??** yields the result. □

In the general (not necessarily uniform) case we have the following.

THEOREM 19.6 (Beck 1980). *Let $\mathcal{F}$ be a family of sets, each of which has at least k ($k \geq 2$) points. Also suppose that for each point v,*

$$\sum_{S \in \mathcal{F}: v \in S} (1 - 1/k)^{-|S|} 2^{-|S|+1} \leq \frac{1}{k}.$$

Then $\mathcal{F}$ is 2-colorable.

PROOF. Let $\mathcal{F} = \{S_1, \dots, S_m\}$ and (again) color the points with red and blue at random, independently of each other and with probability 1/2. Let A_i denote the event that S_i is monochromatic; hence $\Pr[A_i] = 2^{-|S_i|+1}$. Consider the same dependency graph $G = (V, E)$ as above: $\{i, j\} \in E$ if and only if $S_i \cap S_j \neq \emptyset$. We shall prove that the condition of Lemma **??** is satisfied with 19.2

$$x_i := (1 - 1/k)^{-|S_i|}\, 2^{-|S_i|+1}.$$

Indeed, by the definition of the graph G, for every $i = 1, \dots, m$ we have

$$\begin{aligned} x_i \prod_{\{i,j\} \in E} (1 - x_j) &\geq x_i \prod_{v \in S_i} \prod_{j: v \in S_j} (1 - x_j) \\ &\geq x_i \prod_{v \in S_i} \left[1 - \sum_{j: v \in S_j} x_j\right] \geq x_i (1 - 1/k)^{|S_i|}, \end{aligned}$$

since, by the condition of the theorem, $\sum_{j: v \in S_j} x_j \leq 1/k$. Thus,

$$x_i \prod_{\{i,j\} \in E} (1 - x_j) \geq x_i \,(1 - 1/k)^{|S_i|} = 2^{-|S_i|+1} = \Pr[A_i].$$

By the application of Lemma **??** 19.2 we obtain $\Pr[\overline{A}_1 \overline{A}_2 \cdots \overline{A}_n] > 0$, i.e., there is a 2-coloring in which no set of $\mathcal{F}$ is monochromatic. □

Later, Beck (1991) was even able to design an efficient randomized algorithm finding a desired coloring. This was the first time when an algorithmic version of the Lovász Local Lemma was found.

Let us now consider yet another coloring problem. Let $\mathcal{F}$ be a family of k-element sets and suppose that no point appears in more than l of its members. By induction on k, it can be shown (see Exercise **??**) that then it is possible to color the points in $r = l(k-1) + 1$ colors so that no member of $\mathcal{F}$ contains two points of the same color. On the other hand, if we have only $r < k$ colors, then every member of $\mathcal{F}$ will always have at least k/r points of the same color. Is it possible, also in this case (when $r < k$) to find a coloring such that no member has much more than k/r points of one color? The following result says that, if $k = l$ and if we have about $k/\log k$ colors, then such a coloring exists.

THEOREM 19.7 (Füredi–Kahn 1986). *Let k be sufficiently large. Let $\mathcal{F}$ be a k-uniform family of sets and suppose that no point belongs to more than k sets of $\mathcal{F}$. Then it is possible to color the points in $r = \lfloor k/\log k \rfloor$ colors so that every member of $\mathcal{F}$ has at most $v = \lceil 2\mathrm{e} \log k \rceil$ points of the same color.*

In fact, Füredi and Kahn proved a stronger result, where $v = \lfloor 4.5 \log k \rfloor$ and the members of $\mathcal{F}$ have size at most k. The argument then is the same but requires more precise computations.

19.1

PROOF. Color the points of X by r colors, each point getting a particular color randomly and independently with probability $1/r$. Let $A(S,i)$ denote the event that more than v points of S get color i. We are going to apply Lemma **??** to these events. Events $A(S,i)$ and $A(S',i')$ can be dependent only if $S \cap S' \neq \emptyset$. So, we consider the following dependency graph G for these events: the vertex set consists of the pairs (S,i) where $S \in \mathcal{F}$ and $1 \leq i \leq r$, and two vertices (S,i) and (S',i') are joined by an edge if and only if $S \cap S' \neq \emptyset$.

Let d be the maximum degree of G. By the condition on our family $\mathcal{F}$, every member can intersect at most $k(k-1)$ other members, implying that $d \leq (1+k(k-1))r \leq k^3$. By Lemma **??**, 19.1 it remains to show that each of the events $A(S,i)$ can happen with probability at most $1/(4k^3)$.

Since $|S| = k$, the probability that only the points of a subset $I \subseteq S$ get color i, is $(1/r)^{|I|}(1-1/r)^{k-|I|}$. Summing over all subsets I of S, then the event $A(S,i)$ happens with probability at most

$$\sum_{t>v} \binom{k}{t} \left(\frac{1}{r}\right)^t \left(1-\frac{1}{r}\right)^{k-t} \leq \binom{k}{v} \left(\frac{1}{r}\right)^v < \left(\frac{\mathrm{e}k}{vr}\right)^v \leq 2^{-v} < k^{-4}.$$

19.1

By Lemma **??**, with positive probability, none of the events $A(S,i)$ will happen, and the desired coloring exists. □