**Problem Set 4 Solutions**

*Point totals are in the margin; the maximum total number of points was XX.*

1. Concentration of the Longest Common Subsequence

(a) We let $Z_i = (a_i, b_i)$, and let $L = L(Z_1, \ldots, Z_n)$ denote the length of a lcs of $a, b$. Then $X_i = \text{xxpts}$

\[ E(L|Z_1, \ldots, Z_i) \text{ is a martingale (the Doob martingale of } L \text{ w.r.t. } (Z_i)). \]

It is easy to check that $L$ is 2-Lipschitz (if we remove $a_i, b_i$ and their partners from any common subsequence of $a, b$, we get a subsequence at most two shorter; and by reversing the argument we get a similar bound on the increase caused by changing $a_i, b_i$). Since the $Z_i$ are also independent, we can apply Azuma’s inequality with bounded differences of 2 to deduce that

\[ \Pr[|X_n - \mu_n| \geq \lambda] \leq 2 \exp(-\lambda^2/2n). \]

**Note:** We can actually do slightly better by considering instead the filter $Z_{2i-1} = a_i, Z_{2i} = b_i$, which makes $L$ 1-Lipschitz with a difference sequence of length $2n$ and hence replaces the above bound by $2 \exp(-\lambda^2/4n)$.

**Note that independence of the $Z_i$ is crucial here, in addition to the Lipschitz property; see part (ii) below for an illustration of what can go wrong in the absence of independence.**

(b) (i) No difference; the argument above is oblivious to the alphabet size.

(ii) In the absence of independence, we can’t claim any non-trivial concentration. (For example, suppose we have the following values for $a, b$, each with probability $\frac{1}{4}$: $(a = b = 0^n), (a = b = 1^n), (a = 0^n, b = 1^n)$ and $(a = 1^n, b = 0^n)$. Then $E(L) = \frac{n}{2}$, but $|L - E(L)| = \frac{n}{2}$ with probability 1.)

**Note that the arguments of part (a) do still work if $a_i$ and $b_i$ are dependent, provided that $a_i, a_j$ are independent, and $b_i, b_j$ are independent, for $i \neq j$.**

(iii) Here the argument above still holds, but the function $L$ becomes 3-Lipschitz. Thus we get the slightly weaker bound

\[ \Pr[|X_n - \mu_n| \geq \lambda] \leq 2 \exp(-\lambda^2/18n). \]

**Note:** The alternative argument also extends, making $L$ 1-Lipschitz over a sequence of length $3n$ and giving the better bound $2 \exp(-\lambda^2/6n)$.

2. A contact process on a graph

(a) Let $\mathcal{F}_t$ be the $\sigma$-field generated by the outcomes of the first $t$ steps of the process. We show that $\text{xxpts}$

\[ (X_t) \text{ is a martingale w.r.t. the filter } (\mathcal{F}_t). \]

Fix the configuration on the graph after $t$ steps, and assume $X_t \notin \{0, 2m\}$; let $W_t, B_t$ denote the sets of white and black vertices respectively. Then we may write the difference $D_{t+1} = X_{t+1} - X_t$ as

\[ D_{t+1} = \sum_{u \in B_t} d_u Z_u - \sum_{u \in W_t} d_u Z_u, \quad (1) \]

where $d_u$ is the degree of vertex $u$, and $Z_u$ is the indicator r.v. of the event that $u$ changes color. Note that $E(Z_u) = \frac{\text{disc}(u)}{2d_u}$, where $\text{disc}(u)$ is the number of neighbors of $u$ with the opposite color to $u$. Thus

\[ E(D_{t+1}|\mathcal{F}_t) = \sum_{u \in B_t} d_u \times \frac{\text{disc}(u)}{2d_u} - \sum_{u \in W_t} d_u \times \frac{\text{disc}(u)}{2d_u} \]
3. Vertex-disjoint cycles

(b) Let \( T \) be the termination time, which is clearly a stopping time. We apply the Optional Stopping Theorem to the martingale in part (a) with this stopping time. To check the conditions for the OST, note that the differences \( E(X_t - X_{t-1} \mid \mathcal{F}_t) \) are clearly bounded (by \( 2m \)), and also that \( E(T) \) is finite since we can write a finite family of linear equations in the variables \( t_X \), the expected termination time starting from configuration \( X \). The OST now gives \( E(X_T) = E(X_0) = X_0 \). So, letting \( p \) be the probability of termination in the all-white configuration, we have

\[
p \times 2m + (1 - p) \times 0 = X_0,
\]

and hence \( p = \frac{X_0}{2m} \).

(c) Since \( (X_t) \) is a martingale on the integer interval \([0, 2m]\), the techniques of Lecture 23 show that the expected termination time is \( E(T) \leq \left(\frac{2m}{\sigma^2}\right)^2 \), where \( \sigma^2 \) is a lower bound on \( E(D_{t+1}^2 \mid \mathcal{F}_t) = \text{Var}(D_{t+1} \mid \mathcal{F}_t) \).

Hence it is sufficient to prove that \( \text{Var}(D_{t+1} \mid \mathcal{F}_t) \) is bounded below by a constant.

Now from (1) and the fact that the \( Z_u \) are independent given \( \mathcal{F}_t \), we have \( \text{Var}(D_{t+1} \mid \mathcal{F}_t) = \sum_u \text{Var}(d_u Z_u) \).

And for each \( u \) we may compute

\[
\text{Var}(d_u Z_u) = \frac{\text{disc}(u)}{2d_u} \times d_u^2 - \left( \frac{\text{disc}(u)}{2d_u} \times d_u \right)^2 = \frac{\text{disc}(u)}{4} (2d_u - \text{disc}(u)) \geq \frac{\text{disc}(u)d_u}{4}.
\]

But clearly at any time before termination we must have \( \text{disc}(u) \geq 1 \) for at least two vertices \( u \), so \( \text{Var}(D_{t+1} \mid \mathcal{F}_t) \geq \frac{1}{4} \sum_u \text{disc}(u)d_u \geq \frac{1}{4} \), as required.

3. Vertex-disjoint cycles

(a) The probability that a particular component is not represented among the out-neighbors of \( v \) is

\[
(1 - \frac{1}{2})^k \leq \exp(-k/c) \leq \exp(-3 \ln k) = k^{-3}.
\]

Taking a union bound over all \( c \) components yields

\[
\Pr[A_v] \leq ck^{-3} \leq \frac{k}{3 \ln k}k^{-3} = \frac{1}{3k^2 \ln k}.
\]

(b) We may apply the Mutual Independence Principle (Proposition 24.4 from Lecture 24), where the underlying independent random variables are the assignments of vertices to components. Thus event \( A_v \) is independent of all events \( A_u \) such that

\[
\left(\{v\} \cup N_{\text{out}}(v)\right) \cap \left(\{u\} \cup N_{\text{out}}(u)\right) = \emptyset,
\]

where \( N_{\text{out}}(v) \) denotes the set of out-neighbors of \( v \). Since the size of each out-neighborhood is \( k \), this set includes all but at most \((k + 1)^2\) of the \( A_u \). Hence \( |D_v| \leq (k + 1)^2 \), as required.

(c) Note that if none of the events \( A_v \) happens then the construction of part (a) yields \( c \) non-empty components, each of which contains a cycle. And these cycles are plainly vertex-disjoint. Now we may apply the LLL (symmetric version, Lemma 24.2 of Lecture 24, alternative condition \( 4pd \leq 1 \) stated immediately after the lemma) with \( p = (3k^2 \ln k)^{-1} \) by (part (a)) and \( d = (k + 1)^2 \) by (part (b)) to conclude that \( \Pr[\bigwedge_{v \in V} A_v] > 0 \) provided \( 4pd \leq 1 \), i.e., provided

\[
\frac{4(k + 1)^2}{3k^2 \ln k} \leq 1.
\]

This is clearly true for all sufficiently large \( k \), and by a simple calculation it is true for all \( k \geq 7 \). So the claim holds for all such \( k \). But the claim also holds trivially when \( k < 7 \) since then \( c = 1 \).
4. A hands-on approach to making the Lovász Local Lemma algorithmic

(a) We apply exactly the same reasoning as in Lecture 24 to the reduced formula \( \phi' \). Thus we consider a random assignment to the variables of \( \phi' \), and define a “bad” event \( A_i \) for each clause of \( \phi' \). Since each clause of \( \phi' \) contains at least \( \frac{k}{2} \) variables, \( \Pr[A_i] \leq \frac{1}{2^{k/2}} = p \). Also, the dependency graph of \( \phi' \) still has degree at most \( d = 2^{k/20} \), so by the LLL we see that \( \phi' \) is satisfiable provided

\[
4pd = 4 \cdot 2^{-k/2} \cdot 2^{k/20} < 1.
\]

This certainly holds for all sufficiently large \( k \).

(b) Stage 1 produces a reduced formula \( \phi' \) that is guaranteed to be satisfiable, so we just need to find an assignment for it in order to complete the partial assignment produced in Stage 1. By the Claim, with probability \( 1 - o(1) \) the dependency graph \( G' \) of \( \phi' \) has connected components of size at most \( C \log n \). Thus we can partition the clauses of \( \phi' \) into groups of size \( \leq C \log n \) such that no two groups share a variable. Each such group involves at most \( kC \log n \) variables, so we can afford to test all \( 2^{kC \log n} = n^{O(k)} \) assignments for it exhaustively. By part (a), we are guaranteed to find a satisfying assignment for each group, and since the variables in the groups are disjoint we can combine these with the partial assignment from Stage 1 to get a satisfying assignment for \( \phi \). The overall running time is polynomial in \( n \) (assuming \( k \) is a constant).

(c) Suppose \( G' \) contains a connected component of size \( m \). We prove the contrapositive by showing how to construct a set \( S \) of size \( \geq m/d^3 \) with the stated properties. Pick any vertex \( v \) of the component and put it in \( S \); then remove \( v \) together with the ball of radius 3 around it from \( G' \). Pick a remaining vertex of the component adjacent to a removed vertex, put it in \( S \) and and remove it together with its radius-3 ball. Continue in this way until the component is exhausted. Since the degree of \( G' \) is bounded by \( d \), we remove at most \( 1 + d + d(d - 1) + d(d - 1)^2 \leq d^3 \) vertices at each step, so \( |S| \geq m/d^3 \). And by construction, \( S \) is connected in \( G_4 \) and all its elements are at pairwise distance at least four.

(d) The survival of a clause of \( \phi \) depends only on the disposition of its own variables and those in clauses with which it shares a variable. For two clauses at distance at least four in \( G \), these sets of variables are disjoint and hence the survival events are independent. Moreover, for any given clause to survive, either it or one of its neighbors (in \( G \) must have received unsatisfying assignments to at least \( k/2 \) variables, so by a union bound the probability of survival is at most \( 2^{-k/2(d + 1)} \). Putting these two observations together yields the claimed bound.

(e) Let \( S \) be a connected component in \( G_4 \) of size \( r \). We can specify \( S \) uniquely by a spanning tree \( T_S \). \( T_S \) can in turn be uniquely specified by a closed walk \( (v_0, v_1, \ldots, v_{2r-2}) \), where \( v_{2r-2} = v_0 \) and the walk proceeds around the “outside” of the tree (as we saw in Lecture 26 for a different purpose). Since there are \( n \) choices for \( v_0 \) and at most \( d^4 \) (the degree of \( G_4 \)) choices for each successive \( v_i \), the number of such walks in \( G_4 \) is at most \( n(d^4)^{2r-2} \leq n d^{8r} \), as required.

(f) Taking a union bound over all possible “bad” sets \( S \) of size \( r \) as specified in part (c), the probability that any such \( S \) exists is at most

\[
nd^{8r} \left(2^{-k/2(d + 1)}\right)^r \leq n \left(2^{1-k/2d^3}\right)^r.
\]

Setting \( d = 2^{k/20} \) and \( r = (C \log n)/d^3 \), this becomes \( n(2^{-k/20+1})^{C \log n/d^3} \), which tends to 0 as \( n \to \infty \) for any fixed \( k \), taking \( C \) large enough.