CS264: Homework #6

Due by midnight on Wednesday, November 5, 2014

Instructions:

(1) Students taking the course pass-fail should complete the exercises. Students taking the course for a letter grade should complete both the exercises and the problems.

(2) All other instructions are the same as in previous problem sets.

Lecture 11 Exercises

Exercise 41

A *binary linear code* is a subset of $\{0,1\}^n$ that arises as the kernel (over $\mathbb{F}_2$) of a matrix A — i.e., as the solutions to a system of the form $Ax = 0$.

Prove that the set of parity check codes described in class (for a fixed variable set V, ranging over all sets C and bipartite graphs $G = (V, C, E)$) is precisely the set of binary linear codes of length $|V|$.

Exercise 42

Recall the main theorem from lecture, for codes derived from graphs that satisfy the bounded-degree and expansion conditions: there is a constant $\delta > 0$ (independent of n) such that, if $\mathbf{z}$ has Hamming distance less than $\delta n/2$ from the code word $\mathbf{w}$, then the unique optimal solution to the linear program (LP) is $\mathbf{w}$.

Prove that this statement holds in general if and only if it holds in the special where $\mathbf{w} = 0$.

Lecture 12 Exercises

Exercise 43

In this exercise, you can assume the following version of Hall's theorem: if $G = (A, B, E)$ is a bipartite graph with $|N(S)| \geq |S|$ for every $A \subseteq B$, then G has a matching in which all nodes of A are matched.

Prove the result needed for the proof in lecture: if $G = (A, B, E)$ is a bipartite graph such that, for a positive integer c, $|N(S)| \geq c|S|$ for every subset $S \subseteq A$, then there is a subset $F \subseteq E$ of edges such that (i) each vertex of B is incident to at most one edge of F; and (ii) each vertex of A is incident to at least c edges of F.

Problems

Problem 20

(15 points) This problem shows how to use the probabilistic method to prove that good bipartite expanders exist. Let d be a positive integer (a constant). Consider vertex sets A and B, with $|A| = n$ and$|B| = c|A|$ for a constant $c \in (0, 1)$. Obtain a random graph G by choosing, independently for each $a \in A$, d neighbors (with replacement) uniformly at random from B. With probability 1, G is a bipartite graph in which all vertices of A have degree‘d.[1]

[1] In lecture we also insisted that the right-hand side vertices have bounded degree. For simplicity, we drop this constraint for this problem.

Prove that there is a constant $\delta > 0$, which can depend on c and/or d but not on n, such that, with probability approaching 1 as $n \to \infty$,

$$|N(S)| \geq \frac{3}{4}d|S|$$

for every set $S \subseteq A$ with $|S| \leq \delta n$.

[Hint: Use the usual maneuvers from randomized algorithms, like the Chernoff and Union bounds. For example, you could consider the probability that a vertex of $N(S)$ has unique neighbor in S.]

Problem 21

This problem gives a simple and more practical alternative decoding algorithm to the one given in Lectures #11 and #12.[2] Given a corrupted code word $\mathbf{z}_0$, the *SS algorithm* does the following:

(SS) While there is at least one variable $i \in V$ such that more than half of the parity checks in $N(i)$ are unsatisfied, modify $\mathbf{z}$ by flipping the value of an arbitrary such variable.

(a) (5 points) Prove that, no matter what the bipartite graph G (defining the parity check code) and initial corrupted word $\mathbf{z}_0$ are, the SS algorithm is guaranteed to terminate in polynomial time.

(b) (8 points) For this and all remaining parts, assume that the graph G defining the code satisfies the first and third conditions described in lecture (the V-side is d-regular, and the expansion condition).

Prove that, if the current solution $\mathbf{z}$ differs from a code word $\mathbf{w}$ in $m \leq \delta n$ variables, then there are more than $dm/2$ edges between currently corrupted variables (i.e., variables on which $\mathbf{w}$ and $\mathbf{z}$ differ) and currently unsatisfied parity checks. Here δ denotes the same constant as in the expansion condition stated in lecture.

[Hint: In addition to the expansion condition, use the fact that a parity check j containing a corrupted variable can only be satisfied if it contains at least two corrupted variables.]

(c) (3 points) Explain why (b) implies that, if the current Hamming distance from $\mathbf{z}$ to the nearest code word is at most δn, then the SS algorithm will flip the value of some variable.

(d) (2 points) Discuss why (c) does not necessarily imply that the Hamming distance between $\mathbf{z}$ and the nearest code word strictly decreases with the number of iterations of the algorithm.

(e) (5 points) Prove that, if the initial code word $\mathbf{z}_0$ has Hamming distance at most $\delta n/2$ from the nearest code word, then in every iteration of the SS algorithm, the current solution $\mathbf{z}$ has Hamming distance less than δn from the nearest code word.

[Hint: use (b) and a monotonicity argument familiar from part (a).]

(f) (2 points) Conclude that if the SS algorithm is initialized with a corrupted code word $\mathbf{z}_0$ with Hamming distance at most $\delta n/2$ from the nearest code word $\mathbf{w}$, then the SS algorithm is guaranteed to terminate with the solution $\mathbf{w}$.

[2] There are, however, some parameter ranges where LP decoding is known to work and this simpler algorithm is not known to work.