CS 2429
Proof Complexity and Mathematical Programming, Homework
Due: April 10, 2018

1. Prove that the following are equivalent. Let A be a symmetric, n-by-n matrix.
   - A is PSD.
   - The polynomial $p_A(x) = \sum_{i,j} A_{i,j} x_i x_j$ is a sum of squares.
   - There is a set of correlated random variables $(X_1, \ldots, X_m)$ such that for every i, j $E_{i,j}[X_i X_j] = A_{i,j}$ and for every i, X_i is a (distributed like) a normal random variable with mean 0 and variance $A_{i,i}$.

2. Let D be a degree d SOS pseudo-distribution over $\{0,1\}^n$. Let $S \subseteq [n]$. Prove that the distribution corresponding to restricting $x \in \{0,1\}^n$ to S is also a degree d pseudo-distribution.

3. The next two questions are about the ability (or disability) of Sherali Adams to detect cliques in random graphs. Let $G_{n,1/2}$ be the Erdos-Renyi distribution on random undirected graphs with n vertices: for each edge $e_{i,j}$ choose $e_{i,j} = 1$ with probability $1/2$ and otherwise choose $e_{i,j} = 0$.
   (a) Prove that almost surely (with probability that goes to 1 as n goes to infinity), G drawn from $\mathcal{G}_{n,1/2}$ has no clique of size greater than $2 \log n$.
   (b) Prove that for any $\epsilon > 0$, almost surely G drawn from $\mathcal{G}_{n,1/2}$ has a clique of size $2(1-\epsilon) \log n$.

4. Now consider the following constraints, $Clique_G$ for a fixed graph G.

   $Clique_G$, states that G contains a clique of size at least $100 \log n$. The variables are s_i, $i \in [n]$ and the constraints are:
   - $\sum_i s_i \geq 100 \log n$,
   - For all (i,j) such that $e_{i,j} = 1$, $s_i + s_j \leq e_{i,j} + 1$

   For a random G from $\mathcal{G}_{n,1/2}$, $Clique_G$ is unsatisfiable (by the previous homework question).

   (a) Prove that with high probability, over G drawn from $\mathcal{G}_{n,1/2}$, $Clique_G$ has a degree $O(\log n)$ SA refutation.
   (b) Prove that for n sufficiently large, with probability at last $3/4$ over G drawn from $\mathcal{G}_{n,1/2}$, any Sherali-Adams refutation of $Clique_G$ requires degree $\Omega(\log n)$.
   (c) What if we change the LP to say that G has a clique of size $t(n) = n^\epsilon$. What is the maximum $\epsilon > 0$ such that detecting cliques of size n^ϵ is hard for degree $o(t(n))$ SA? Prove your answer.