HOMEWORK 6
Due: Thursday, May 1

Definitions: Given boolean functions $f_1, \dots, f_d$, a "k-query test" is a probability distribution $\mathbf{Pr}$ on "checks" of the form

$$\phi(f_{i_1}(x^{(1)}), \dots, f_{i_k}(x^{(k)})).$$

Here the random variables are ϕ, a predicate $\{-1,1\}^k \to \{\text{pass, fail}\}$, the indices $i_1, \dots, i_k$, and the strings $x^{(1)}, \dots, x^{(k)}$. The "predicate support" of the test, denoted Φ, is the collection of all k-bit predicates used with nonzero probability.

1. DPCTs. A "Dictator Projection-Consistency Test" (DPCT) consists of two parts:

(i) A k-query test over just two functions, $f_1 : \{-1,1\}^K \to \{-1,1\}$ and $f_2 : \{-1,1\}^{dK} \to \{-1,1\}$. We write the coordinates for f_2 as $[dK] = J_1 \cup \dots \cup J_K$, where each $|J_i| = d$.

(ii) Two (independent) *randomized* decoding procedures $\mathrm{Dec}_1 : f_1 \mapsto i \in [K]$ and $\mathrm{Dec}_2 : f_2 \mapsto j \in [dK]$. The second one should be "permutation-invariant", meaning that for all permutations σ on $[dK]$ and all functions f_2, the distribution $\mathrm{Dec}_2(f_2 \circ \sigma)$ should be the same as the distribution $\sigma(\mathrm{Dec}_2(f_2))$.

The DPCT has "completeness" c if the following holds: If $f_1 = \chi_i$ and $f_2 = \chi_j$ where $j \in J_i$, then $\mathbf{Pr}[\text{pass}] \geq c$. The DPCT has "$(\epsilon, \delta)$-soundness" s if, whenever $\mathbf{Pr}[\text{pass}] \geq s + \epsilon$, the probability that $\mathrm{Dec}_2(f_2) \in J_{\mathrm{Dec}_1(f_1)}$ is at least δ, where the probability is over the (independent) decoding procedures.

Assume there exists a family of DPCTs, one for each value of K and d, and that for all constant $\epsilon > 0$ there exists constant $\delta > 0$ such that each DPCT has completeness c and (ϵ, δ)-soundness s. Show that for all constant $\eta > 0$, the c vs. $s + \eta$ Max-Φ problem is NP-hard, where Φ is the predicate support of the tests.

2. DATs.

a. Suppose that we have a k-query test operating on just a single function $f : \{-1,1\}^K \to \{-1,1\}$. Show that $\mathbf{Pr}[\text{pass}]$ can be expressed as a sum of products of k of f's Fourier coefficients.

b. Let $\mathcal{F} = \{f_1, \dots, f_d\}$ be a collection of functions $\{-1,1\}^K \to \{-1,1\}$. Suppose we take the one-function test from part (a) and "apply it to $\mathcal{F}$", meaning that we convert each check

$$\phi(f(x^{(1)}), \dots, f(x^{(k)}))$$

to the check

$$\phi(f_{i_1}(x^{(1)}), \dots, f_{i_k}(x^{(k)})),$$

where the i_j's are independently and uniformly chosen from $[d]$. Show that $\mathbf{Pr}[\text{pass}]$ can be written as the *same* sum of products of k of F's Fourier coefficients, where $F : \{-1,1\}^K \to [-1,1]$ denotes the function $F(x) = \mathrm{avg}_{i \in [d]} f_i(x)$.

c. A "Dictator Average Test" (DAT) is like a DPCT, except that its k-query test only operates on single functions $f : \{-1,1\}^K \to \{-1,1\}$, and it has only a single, permutation-invariant decoding procedure $\mathrm{Dec} : f \mapsto i \in [K]$. It has completeness at least c if $\mathbf{Pr}[\text{pass}] \geq c$ whenever f is a dictator, $f = \chi_i$ for some i. It has (ϵ, δ)-soundness s if, whenever it is applied to a *collection* $\mathcal{F} = \{f_1, \dots, f_d\}$ and $\mathbf{Pr}[\text{pass}] \geq \epsilon$, the probability that $\mathrm{Dec}(f_{i_1}) = \mathrm{Dec}(f_{i_2})$ is at least δ, where $i_1, i_2 \in [d]$ are chosen independently and uniformly (and the decodings are independently done).

Let k be a universal constant. Assume there exists a family of DATs, one for each value of K, and that for all constant $\epsilon > 0$ there exists constant $\delta > 0$ such that each DAT has completeness c and (ϵ, δ)-soundness s. Show that for all constant $\eta > 0$, the $c - \eta$ vs. $s + \eta$ Max-Φ problem is NP-hard assuming the Unique Games Conjecture, where Φ is the predicate support of the tests.

3. Gadgets.

a. Show that $1 - \epsilon$ vs. $7/8 + \epsilon$ for Max-E3Sat is NP-hard. (Hint: reduce from Hastad's theorem. Remark: indeed, 1 vs. $7/8 + \epsilon$ is known to be NP-hard.)

b. Let "Max-Maj$_3$" be the CSP where the constraints are of the form "Majority$(\ell_{i_1}, \ell_{i_2}, \ell_{i_3}) = 1$", where the ℓ's denote literals x_i or $\overline{x}_i$. Show that approximating Max-Maj$_3$ to factor $2/3 + \epsilon$ is NP-hard. (Hint: same. Remark: there is a factor-2/3 approximation algorithm for Max-Maj$_3$ using SDP, but its analysis is quite tricky.)

c. Suppose that there is a factor-α approximation algorithm for Max-And$_k$, the CSP where each constraint is an AND of up to k literals. Show that there is a factor-α approximation algorithm for "Max-kCSP", the CSP where *any* constraint on up to k variables is allowed.

4. Min-E2Lin-Deletion to Multicut. Suppose there is a factor-α approximation algorithm for Multicut. Show there is a factor-α approximation algorithm for "Min-E2Lin-Deletion" — i.e., a $1 - \epsilon$ vs. $1 - \alpha\epsilon$ algorithm for Max-E2Lin.

5. Semidefinite Programs for Sparsest Cut. Consider the sparsest cut problem given by a n-vertex graph with the edge-capacity vector C and demand vector D, where we want to find a cut $(S, \bar{S})$ minimizing $\frac{C \cdot \delta_S}{D \cdot \delta_S}$. Also, recall that ℓ_2^2 is the set of all distances d such that $\sqrt{d} \in \ell_2$.

a. In Lecture 20, we considered the relaxation

$$\min_{d \in \ell_2^2} \frac{C \cdot d}{D \cdot d}$$

and claimed that this relaxation has a very large integrality gap of $\Omega(n)$. Give an example showing this.

b. We then considered the relaxation

$$\min_{d \in \ell_2^2 \cap \text{metrics}} \frac{C \cdot d}{D \cdot d}.$$

Write a semidefinite program that finds such a metric d.

6. Getting Semidefinite Programs from Linear Programs. Suppose the $\{0, 1\}$ variables for some problem in question are $\{x_1, x_2, \ldots, x_n\}$, and the ILP formulation of the problem seeks to maximize $c^T x$ subject to the integrality constraints $x \in \{0, 1\}^n$ and problem constraints $g_\ell(x) \geq 0$ for all $1 \leq \ell \leq m$. E.g., for vertex cover, we have $m = |E|$ problem constraints

$$g_{ab}(x) = x_a + x_b - 1 \quad \forall (a, b) \in E.$$

Let K be the LP relaxation of this ILP obtained by dropping the integrality constraints. (For vertex cover, this is the standard LP relaxation K_{VC}, which we know has an integrality gap of 2.)

Suppose we have a vector $y \in \mathbb{R}^{2^n}$, with a coordinate for each subset of variables in $[n]$. The *moment matrix* $M_t(y)$ is an $\binom{[n]}{\leq t} \times \binom{[n]}{\leq t}$ matrix indexed by pairs of sets of at most t variables, whose entries are

$$(M_t(y))_{I,J} = y_{I \cup J} \qquad \forall I, J \in \binom{[n]}{\leq t}$$

Note that each linear constraint g_ℓ can also be viewed as a vector in $\mathbb{R}^{2^n}$: the coordinates in this vector for sets of size ≤ 1 correspond to the coefficients of the corresponding variable in the linear constraint (including the constant), and all other coordinates in the vector are 0. E.g., the vector corresponding to g_{ab} has 1 in positions for $\{a\}$ and $\{b\}$ and -1 in the position for $\emptyset$. Finally, given two vectors g and y in $\mathbb{R}^{2^n}$, $(g * y)$ is another vector in $\mathbb{R}^{2^n}$, such that for $H \subseteq [n]$,

$$(g * y)(H) = \sum_{J \subseteq [n]} g(J) y(J \cup H).$$

Recall that $A \succcurlyeq 0$ denotes the constraint that A is a positive-semidefinite matrix. Given a ILP formulation as above, and some integer $t \geq 0$, define the set $P_t(K)$ to consist of the points $y \in \mathbb{R}^{2^n}$ such that

$$M_{t+1}(y) \succcurlyeq 0 \qquad \textbf{and} \qquad M_t(g_\ell * y) \succcurlyeq 0 \quad \forall \ell = 1 \ldots m.$$

Define $Q_t(K)$ to be the "projection" on $\mathbb{R}^n$ of $P_t(K) \cap \{y \mid y_\emptyset = 1\}$; i.e., for each vector $y \in P_t(K) \cap \{y \mid y_\emptyset = 1\}$, the set $Q_t(K)$ contains a n-dimensional vector y' such that $y'_i = y_{\{i\}}$ for all $i \in [n]$.

a. Consider the polytope K_{VC} given by the LP relaxation for vertex cover above. Show that $Q_0(K_{VC}) \subseteq K_{VC}$.

b. Write the constraints corresponding to $P_1(K_{VC})$.