S^3CS 2014 - Sum of Squares - Homework 2.

Exercise 1. Let $v \in \mathbb{R}^n$ be unit ($\|v\|_2 = 1$) vector with at most ϵn nonzero coordinates. Prove that $\|v\|_4^4 \geq 1/(\epsilon n)$.

Exercise 2. (Cauchy-Schwarz, Holder, and friends) Prove that every degree 4 pseudo-distribution $\{(u, w)\}$ over R^{2n}

1. $\tilde{\mathbb{E}}\langle u, w\rangle \leq \tilde{\mathbb{E}}\frac{1}{2}\|u\|^2 + \frac{1}{2}\|w\|^2$
2. $\tilde{\mathbb{E}}\langle u, w\rangle \leq \sqrt{\tilde{\mathbb{E}}\|u\|^2}\sqrt{\tilde{\mathbb{E}}\|w\|^2}$
3. $\tilde{\mathbb{E}}\sum_i u(i)^3 w(i) \leq \left(\tilde{\mathbb{E}}\|u\|_4^4\right)^{3/4} \left(\|w\|_4^4\right)^{1/4}$
4. Conclude the proof of Lemma 2 for pseudo expectations.

Exercise 3. Suppose that $w_1, \ldots, w_n \in \mathbb{R}^k$ are random standard Gaussian vectors:

1. Prove that for every vector v, $\mathbb{E}\langle w, v\rangle^4 = 3\|v\|_2^4$.
2. Let $P(v)$ be the polynomial $\frac{1}{n}\sum_{i=1}^n \langle w_i, v\rangle^4$. Prove that with high probability, every coefficient of P is within an additive constant of $\pm O(polylog(n)/\sqrt{n})$ from its expectation (i.e., the corresponding coefficient of the polynomial $3\|v\|_2^4$)
3. Let $Q(v) = \sum_{a,b,c,d} Q_{a,b,c,d} x^a x^b x^c x^d$ be a homogenous degree 4 polynomial in $P_4[\mathbb{R}^k]$. Prove that $\tilde{\mathbb{E}}Q(v) \leq \sqrt{\sum Q_{a,b,c,d}^2}\sqrt{\tilde{\mathbb{E}}\|v\|_2^4}$.
4. Prove that there is some constant C such that if $n > (\log^C k)k^4$ then $\tilde{\mathbb{E}}P(v) \leq 10\|v\|_2^4$ for every degree 4 pseudo expectation.
5. Conclude the proof of Lemma 1 for pseudo expectations (for the case that n is this large). You can prove the weaker version that (2) holds in pseudo expectation as opposed that the difference betwen the RHS and LHS is a sum of squares. Bonus: can you prove it as long as $n \gg k^2$ (possibly losing some polylog factors)

(assuming we get to do the lower bound)

Exercise 4. Prove that for every $\epsilon > 0$, there exists c such that for every set of cn vectors $a_1, \ldots, a_m \in \mathsf{GF}(2)^n$, if we pick $b \in \mathsf{GF}(2)^m$ at random then with probability at least 0.99, for every $x \in \mathsf{GF}(2)^n$, there number of $i \in [m]$ such that $\langle a_i, x\rangle = b_i$ is at most $(1/2 + \epsilon)m$.

Exercise 5. Prove that for every constant $c \geq 1$, there exists a constant $\delta > 0$ and a bipartite graph $G = (L \cup R, E)$ such that every vertex in L is connected to three vertices in R, and for every set S of size at most ϵn, the size of $\Gamma(S) = \{j \in R : (i,j) \in E\}$ is at least $1.6|S|$. See footnote for hint[1]

Exercise 6. Using the lower bound for random 3LIN as a black box, prove the following statement: There exists a number $\epsilon > 0$ 3SAT formula φ with $m = O(n)$ clauses and n variables, such that every assignment to φ satisfies at most $(7/8 + 0.01)m$ of its clauses, but there exists a degree ϵn pseudo-distribution $\{x\}$ that satisfies all the constraints of φ. See footnote for hint[2]

[1] **Hint:** For k that is more than some small constant, if S is a set of size k and T is a set of size at most $1.6k$, the probability that $\Gamma(S) \subseteq T$ is small enough that we can do a union bound over all such sets S and T. For smaller sets, we can bound the expectation of the number of such sets S that violate this condition (which using Markov with decent probability we will not have much more than that) and this number will be small enough that we can remove all left vertices involved in them.

[2] **Hint:** Use the simple (but useful!) fact that zero is an even number, namely that if $a \oplus b \oplus c = 1$ then $a \vee b \vee c = 1$.