SoS lower bounds on random 3XOR

Ian Mertz

March 24, 2018

1 Upper and lower bounds for SoS

In the last lecture we saw how the MAXCUT problem can be solved with the optimal approximation ratio of 0.878 by degree 5 SoS, in an argument that mirrors the famed Goemans-Williamson algorithm by constraining the degree 2 moments to be both $\{0, 1\}$ valued and to obey some nice properties in relation to one another. This is an example of how moving up in the SoS hierarchy can give us tangible progress towards the ultimate goal, which is the intersection of the feasible region with the Boolean hypercube. From the perspective of the proof system SoS, introducing higher degree terms means that any *pseudodistribution* has to obey both consistency with the input axioms and local consistencies with one another at a higher degree, which can become much harder if the tautology can be refuted very locally.

But what if moving up in the hierarchy doesn't help us? We saw how degree 0 SoS gave almost the worst possible integrality gap for MAXCUT, $\frac{1}{2} + \epsilon$ for any $\epsilon > 0$, which barely beats choosing the cut at random. Also once we hit degree 5 and obtain the optimal approximation ratio under $\mathsf{P} \neq \mathsf{NP}$, and using the fact that degree d SoS proofs can be found in time $n^{O(d)}$, we immediately get that, modulo $\mathsf{P} \neq \mathsf{NP}$, degree $\omega(1)$ is needed to push past this 0.878 barrier.

In this lecture we see an far stronger lower bound, showing that for a problem in P, degree d SoS has an integrality gap of $\frac{1}{2} + \epsilon$ for $d = \Theta(n)$. The problem we use is 3XOR, where the constraints are of the form

$$x_i x_j x_k = a_{ijk} \qquad (x_i, x_j, x_k, a_{ijk} \in \{\pm 1\})$$

This is the canonical way of representing the $\{0, 1\}$ valued constraints $x_i \oplus x_j \oplus x_k = a_{ijk}$, noting that the $\mod$ operation can be tricky to express in a polynomial over the reals. While we've moved into a different domain, namely $\{\pm 1\}$, these variables can be transformed into $\{0, 1\}$ valued variables via the simple transformation $z_i = \frac{1 - x_i}{2}$. Also new is the fact that the input axioms are of degree 3 instead of linear, but since our degree lower bound is $\Theta(n)$ this difference is insignificant.

The 3XOR problem varies wildly in terms of difficulty depending on what problem we want to solve. Håstad's 3XOR lemma states that for $\delta, \epsilon > 0$, it is NP hard to decide whether a given 3XOR instance ϕ on $m = O(n)$ constraints has value at least $1 - \delta$ or at most $\frac{1}{2} + \epsilon$, where the value of ϕ is the maximum fraction of clauses satisfiable simultaneously. However if we shift the goalposts a bit and ask the complexity of distinguishing whether ϕ has value 1, in other words if ϕ is satisfiable, or value at most $\frac{1}{2} + \epsilon$, then the problem can be solved in time n^3 using the classic technique of Gaussian elimination. In particular, this polytime algorithm holds for the *random* 3XOR problem, where in each of the m constraints of ϕ we pick $x_i x_j x_k$ uniformly at random and then either choose the a_{ijk}s to be consistent with some assignment $\alpha \in \{\pm 1\}^n$ (ϕ has value 1) or uniformly at random, in which case we will show that with high probability a random 3XOR instance has value at most $\frac{1}{2} + \epsilon$.

This is somewhat shocking at first considering the strength of SoS; consider, for example, the fact that under Unique Games SoS achieves the best integrality gap for every NP-complete problem. The lower bound is based on a constraint satisfaction problem for Gaussian elimination, and uses randomness to ensure that with high probability, a) we have a minimal number of constraints satisfied (only $\frac{1}{2}+\epsilon$), and b) finding a contradiction among the constraints is highly "non-local", in that it takes fixing a huge number of terms to falsify even one axiom.

An important application of this lower bound is in *hardness reductions*, or shifting the pseudodistribution to give hardness results for other problems. We will show at the end how to immediately achieve an integrality gap of $\frac{7}{8}+\epsilon$ for random instances of 3SAT, which is known to be both easy to achieve (a random assignment suffices) and impossible to beat under $\mathsf{P} \neq \mathsf{NP}$.

But probably the most important part of the result for this course is the proof, which is incredibly elegant and simple but uses a lot of the techniques that show up time and time again in proof complexity lower bounds. A rough outline of the proof is as follows:

1. Figure out every pseudodistribution value that has to be fixed and fix them appropriately, setting the rest to be perfectly random. This may or may not be well-defined.

2. Show that the resulting moment matrix is positive semi-definite by using a partition of the monomials into equivalence classes, assuming the pseudodistribution is well-defined.

3. Show that the pseudodistribution is well defined with high probability, using both the *expansion* of the 3XOR instance, the fact that expansion implies good *boundary expansion*, and that good boundary expansion implies that any contradiction in the way we've defined the pseudodistribution must occur at a very high degree, setting our d to be just below that threshold.

The third step in particular has many ingredients crucial in a number of fundamental lower bounds in proof complexity, such as the pigeonhole principle and the Tseitin tautologies.

Before going into the pseudodistribution we show that a random 3XOR instance is indeed highly unsatisfiable with high probability.

Lemma 1.1. *Let ϕ be an instance of 3XOR on n variables with $m = c_\epsilon n$ constraints (for some constant c_ϵ depending only on ϵ) be chosen as follows: for each constraint we choose $i, j, k \sim [n], a_{ijk} \sim \{\pm 1\}$ iid. Then with probability at least $1 - 2^{-n}$, every assignment $x \in \{\pm 1\}^n$ satisfies at most $(\frac{1}{2}+\epsilon)m$ constraints, where the probability is over the choice of ϕ.*

Proof. For a fixed $\alpha \in \{\pm 1\}^n$ we let Y_j^α be the event that the jth constraint is satisfied by $x = \alpha$ and $Y^\alpha = \sum_j Y_j^\alpha$ be the number of constraints satisfied by $x = \alpha$. By construction of the constraints, for a fixed α each Y_j^α is an independent Bernoulli random variable with expectation $\frac{1}{2}$. Therefore Chernoff implies

$$\Pr_\phi[Y^\alpha > (\frac{1}{2}+\epsilon)m] < 2^{O(-\epsilon^2 m)} \quad \forall \alpha \in \{\pm 1\}^n$$

and so by a union bound on all x

$$\Pr_\phi[\exists \alpha \in \{\pm 1\}^n \mid Y^\alpha > (\frac{1}{2}+\epsilon)m] < 2^{n - O(\epsilon^2 m)}$$

Choosing $m = c_\epsilon n$ for an appropriate $c_\epsilon = O(\frac{1}{\epsilon^2})$ makes this probability less than 2^{-n}. □

2 The 3XOR lower bound for SoS

Fix a random 3XOR instance ϕ on n variables with $m = c_\epsilon n$ clauses. By Lemma 1.1 it has soundess at most $\frac{1}{2} + \epsilon$ with overwhelmingly high probability. The degree lower bound on 3XOR then follows from the following lemma, which will be our main task for the rest of this lecture:

Lemma 2.1. *For any such ϕ, with probability 0.99 there exists a pseudodistribution of degree $\Omega(n)$ such that in expectation all constraints of ϕ are satisfied.*

Proof. Let $x_S = \prod_{i \in S} x_i$, and in particular let $x_{ijk} = x_i x_j x_k$ for convenience. Our degree d pseudo-expectation operator $\tilde{\mathbb{E}}[x_S]$ will be required to satisfy the following constraints:

1. $\tilde{\mathbb{E}}[x_S]$ is defined for all $|S| \leq d$, and extends linearly to all $\tilde{\mathbb{E}}[f]$ of degree at most d
2. $\tilde{\mathbb{E}}[x^2 f] = \tilde{\mathbb{E}}[f]$ (recall that $x \in \{\pm 1\}$, not $\{0, 1\}$)
3. $\tilde{\mathbb{E}}[1] = 1$
4. $\tilde{\mathbb{E}}[x_{ijk}] = a_{ijk}$ for all input axioms
5. $\mathcal{M} \succeq 0$ where $\mathcal{M}_{S,T} = \tilde{\mathbb{E}}[x_S]\tilde{\mathbb{E}}[x_T]$

While the only real constraint on the nontrivial monomials seems to be (4), there are many other low-degree implications we need to take into account. For example, $x_{123} = +1$ and $x_{145} = -1$ implies that $x_{2345} = +1 \times -1 = -1$. Note that the x_1 variables cancel out by (2), and in general what we are left with when multiplying x_S and x_T is the *symmetric difference* $x_{S \triangle T}$. This suggests a procedure for choosing our pseudodistribution, and it turns out that these are the only implications we need to make; all other monomials can be safely left right in between $+1$ and -1 (here we will see the convenience of our representation, as $\tilde{\mathbb{E}}[x] = 0$ instead of $\frac{1}{2}$ for completely random x).

Our pseudoexpectation will be defined by the following algorithm:

```
• let $D \leftarrow \emptyset$
• for all axioms $x_{ijk} = a_{ijk}$:
    – set $\tilde{\mathbb{E}}[x_{ijk}] = a_{ijk}$
    – $D \leftarrow D \cup \{\{i, j, k\}\}$
• while there are $S, T \in D$ such that $S \triangle T \notin D$ and $|S \triangle T| \leq 2d$:
    – $\tilde{\mathbb{E}}[x_{S \triangle T}] = \tilde{\mathbb{E}}[x_S]\tilde{\mathbb{E}}[x_T]$
    – $D \leftarrow D \cup \{S \triangle T\}$
• if there exist $S, T \in D$ such that $\tilde{\mathbb{E}}[x_{S \triangle T}] \neq \tilde{\mathbb{E}}[x_S]\tilde{\mathbb{E}}[x_T]$, return failure
• for all $S \notin D$ such that $|S| \leq d$, set $\tilde{\mathbb{E}}[x_S] = 0$
• for all $f$ of degree at most $d - 2$ and all $x$, set $\tilde{\mathbb{E}}[x^2 f] = \tilde{\mathbb{E}}[f]$
• return $\tilde{\mathbb{E}}$
```

We first assume that the bolded line is not invoked and show that we've satisfied all the conditions on $\tilde{\mathbb{E}}$. Clearly (1)-(4) are satisfied by construction, and so we only need to show (5). We define an equivalence relation $\sim$ on sets of size at most d as follows: $S \sim T$ iff $\tilde{\mathbb{E}}[x_{S \triangle T}] \neq 0$, or equivalently $S \triangle T \in D$ (note that we run over all sets of size at most $2d$ in our algorithm, so we don't need to require that $|S \triangle T| \leq d$). This is clearly reflexive and symmetric, and to see transitivity note that if $S \sim T$ and $T \sim U$, then $S \triangle T, T \triangle U \in D$ at some point, and so $(S \triangle T) \triangle (T \triangle U) = S \triangle U$ will be too. Thus $\sim$ partitions the set of all sets of size at most d into equivalence classes I, and we let S_I be a representative from each class.

We claim that

$$\tilde{\mathbb{E}}[x_S]\tilde{\mathbb{E}}[x_T] = \sum_I \tilde{\mathbb{E}}[x_{S \triangle S_I}]\tilde{\mathbb{E}}[x_{T \triangle S_I}]$$

First consider the case when $S \sim T$. Then $S, T \in I$ for some I, which implies that $S \sim S_I$ and $T \sim S_I$ and thus

$$\tilde{\mathbb{E}}[x_{S \triangle S_I}]\tilde{\mathbb{E}}[x_{T \triangle S_I}] = \tilde{\mathbb{E}}[x_{S \triangle S_I \triangle T \triangle S_I}] = \tilde{\mathbb{E}}[x_{S \triangle T}] = \tilde{\mathbb{E}}[x_S]\tilde{\mathbb{E}}[x_T]$$

by definition of $\sim$, whereas for all other classes I',

$$\tilde{\mathbb{E}}[x_{S \triangle S_{I'}}]\tilde{\mathbb{E}}[x_{T \triangle S_{I'}}] = 0 \times 0 = 0$$

Thus the sum over all I gives us

$$\sum_I \tilde{\mathbb{E}}[x_{S \triangle S_I}]\tilde{\mathbb{E}}[x_{T \triangle S_I}] = \tilde{\mathbb{E}}[x_S]\tilde{\mathbb{E}}[x_T] + \sum_{I'} 0 = \tilde{\mathbb{E}}[x_S]\tilde{\mathbb{E}}[x_T]$$

Now consider when $S \not\sim T$. Then S and T are in different equivalence classes, and so for every I, at least one of $\tilde{\mathbb{E}}[x_{S \triangle S_I}]$ and $\tilde{\mathbb{E}}[x_{T \triangle S_I}]$ is 0. Thus

$$\sum_I \tilde{\mathbb{E}}[x_{S \triangle S_I}]\tilde{\mathbb{E}}[x_{T \triangle S_I}] = \sum_I 0 = 0$$

Because $|S|, |T| \leq d$, if $\tilde{\mathbb{E}}[x_S]\tilde{\mathbb{E}}[x_T] \neq 0$ then $|S \triangle T| \leq 2d$ and so $\tilde{\mathbb{E}}[x_{S \triangle T}] \neq 0$, which implies that $S \sim T$ and contradicts our assumption. Thus $\tilde{\mathbb{E}}[x_S]\tilde{\mathbb{E}}[x_T] = \sum_I \tilde{\mathbb{E}}[x_{S \triangle S_I}]\tilde{\mathbb{E}}[x_{T \triangle S_I}]$ as claimed, and so we can write $\mathcal{M}_{S,T} = \tilde{\mathbb{E}}[x_S]\tilde{\mathbb{E}}[x_T] = \sum_I \tilde{\mathbb{E}}[x_{S \triangle S_I}]\tilde{\mathbb{E}}[x_{T \triangle S_I}]$, or in other words

$$\mathcal{M} = \sum_I x_I x_I^T \qquad x_I = (x_{S \triangle S_I})_{|S| \leq d}$$

which implies that $\mathcal{M} \succeq 0$ by the characterization of PSD matrices as sums of outer products.

Now we turn our attention to the bolded condition, and show that we never contradict ourselves by defining $\tilde{\mathbb{E}}[x_S] = +1$ and $\tilde{\mathbb{E}}[x_S] = -1$ at the same time. For this we will need to shift our view of random 3XOR to a bipartite graph picture. Let $B = C \cup V$ be a bipartite graph where $|C| = m$ and $|V| = n$, and for every vertex $C_j \in C$ we give it a uniformly random neighborhood of size 3 as well as a random label $a_j \sim \{\pm 1\}$. This is equivalent to the original random 3XOR generating procedure, interpreting the neighborhood of C_j as the three variables in the jth constraint. We let the neighborhood of C_j be referred to as $\Gamma(C_j)$, and more generally $\Gamma(T) = \cup_{C_j \in T} \Gamma(C_j)$

We say that B is a (t, β)-*expander* if for all left hand side sets $T \subset C$, $|T| \leq t$, we have that $|\Gamma(T)| \geq \beta |T|$. Roughly speaking the more our graph expands, the larger number of variables that show up in each monomial that we derive from a small subset of the constraints, and the larger our monomials the harder it is to find a contradiction because each individual variable is "less constrained". To begin analyzing how a contradiction could occur in our algorithm we show that our random 3XOR instance has good expansion with high probability.

Lemma 2.2. *For $m = c_\epsilon n$ and any constant $\delta > 0$ there exists a constant $\eta > 0$ depending on δ and c_ϵ such that with probability 0.99, B is $(\eta n, 2 - \delta)$-expanding.*

Proof. Let Y_S be the event that the set $S \subseteq [m]$ of size $s \leq \eta n$ has expansion less than $2 - \delta$. There are at most $\binom{n}{(2-\delta)s}$ possible neighborhoods, and each vertex has $\binom{(2-\delta)s}{3}$ possible individual neighborhoods from this total neighborhood, each occurring with probability $\frac{1}{n^3}$. By extension there are $\binom{\binom{(2-\delta)s}{3}}{s}$ possible settings of all the edges on S, each of which occurs with probability $(\frac{1}{n^3})^s$. Thus we get that

$$\begin{aligned} \Pr[Y_S > 1] &\leq \binom{n}{(2-\delta)s}\binom{\binom{(2-\delta)s}{3}}{s}(\frac{1}{n^3})^s \\ &\leq (\frac{ne}{(2-\delta)s})^{(2-\delta)s}((\frac{(2-\delta)se}{3})^3\frac{1}{s})^s\frac{1}{n^{3s}} \\ &\leq C(\frac{n^{2-\delta}}{s^{2-\delta}} \cdot s^2 \cdot \frac{1}{n^3})^s \\ &\leq C(\frac{s^\delta}{n^{1+\delta}})^s \\ &\leq C(\frac{s}{n})^{\delta s}n^{-s} \end{aligned}$$

and taking the sum over all $\binom{m}{s}$ possible S gives us

$$\begin{aligned} \Pr[\exists S, |S| \leq s \mid Y_S > 1] &\leq m^s \cdot C(\frac{s}{n})^{\delta s}n^{-s} \\ &\leq C(\frac{s}{n})^{\delta s}(\frac{c_\epsilon n}{n})^s \\ &= (c_{\epsilon,\delta}\frac{s}{n})^{\delta s} \end{aligned}$$

which is at most 0.01 for $s \leq \eta n$ as long as $\eta \leq \frac{1}{2c_{\epsilon,\delta}}$. □

Another concept we use is *boundary expansion*. Similarly to expansion we say that B is a (t, γ)-*boundary expander* if for all left hand side sets $T \subset C$, $|T| \leq t$, we have that $|B\Gamma(T)| \geq \beta|T|$, where $B\Gamma(T)$ is the set of all vertices in $\Gamma(T)$ with exactly one neighbor in T. A simple observation connects expansion and boundary expansion.

Lemma 2.3. *If B is a (t, β)-expander and C is 3-regular, then B is a $(t, 2\beta - 3)$-bounndary expander.*

Proof. Fix any set $S \subseteq C$ of size $s \leq t$. Letting the number of edges on S be $E(S)$, we have that

$$\begin{aligned} 3|S| &= E(S) \\ &\geq |B\Gamma(S)| + 2|\Gamma(S)/B\Gamma(S)| \\ &= 2|\Gamma(S)| - |B\Gamma(S)| \\ &\geq 2\beta|S| - |B\Gamma(S)| \end{aligned}$$

The lemma follows by adding $|B\Gamma(S)| - 3|S|$ to both sides. □

Thus if we choose δ in our expansion lemma to be strictly less than 0.5, say $\delta = 0.3$, then our graph will have constant boundary expansion for sets up to the same size. When $\delta = 0.3$ the boundary expansion will be $2 \cdot 1.7 - 3 = 0.4$, and so from here on out we assume B is a $(\eta n, 0.4)$-boundary expander. Here we pause and flesh out our intuition about how expansion will help us by looking at boundary expansion. If we have a large set $T \subseteq C$ such that $T \leq \eta n$, it must have have a boundary of size $0.4|T| = \Omega(n)$. Now if we consider the monomial that is the symmetric difference of all constraints in T, that symmetric difference *must* include the boundary, as each variable in the boundary appears in exactly one constraint in T. Therefore in order to "collect up" enough axioms in C in order to derive a contradiction, we will need to define some intermediate pseudoexpectations of very high degree monomials, and if we set our degree just below this threshold we can avoid being able to get enough axioms to get a contradiction

We now formalize this with the concept of deriving a pseudoexpectation value for x_S. Since the only pseudoexpectations defined in our algorithm are either input axioms or symmetric differences of previous sets, we can associate each monomial x_S for which $\tilde{\mathbb{E}}[x_S] \neq 0$ with a set of the input axioms $T \subseteq C$, such that $\triangle_{C_j \in T}\Gamma(C_j) = x_S$ and $\tilde{\mathbb{E}}[x_S] = \prod_{C_j \in T} a_j$. In order to define x_S in this way however, we need to ensure that every intermediate monomial has low degree. We define a sequence for x_S to be $\mathcal{S} = (T_1 \dots T_k)$ for $T_i \subseteq [m]$ such that

- each T_i either has size at most 1 (corresponding to a single input axiom $\tilde{\mathbb{E}}[x_{ijk}] = a_{ijk}$ or the axiom $\tilde{\mathbb{E}}[x_\emptyset] = 1$) or is the symmetric difference of two previous sets $T_j \triangle T_\ell$, $j, \ell < i$
- $\triangle_{C_j \in T_k}\Gamma(C_j) = x_S$
- $|\triangle_{C_j \in T_\ell} \Gamma(C_j)| \leq d$ for all $T_\ell \in \mathcal{S}$
- we call $\prod_{C_j \in T_k} a_j$ the value of $\mathcal{S}$, and note that $\tilde{\mathbb{E}}[x_S]$ receives this value by our algorithm

We define a contradiction at x_S to be two distinct sequences for x_S with different values, or in other words $\mathcal{S}_1 = (T_1 \dots T_k), \mathcal{S}_2 = (T'_1 \dots T'_{k'})$ such that $\triangle_{C_j \in T_k}\Gamma(C_j) = \triangle_{C_j \in T'_{k'}}\Gamma(C_j) = x_S$ but $\prod_{C_j \in T_k} a_j = -1$ and $\prod_{C_j \in T'_{k'}} a_j = +1$. Observe that if such sequences exist then there exists a contradiction at $x_\emptyset$, where $\mathcal{S}_1 = (T_1 \dots T_k, T'_1 \dots T'_{k'}, T_k \triangle T'_{k'})$ and $\mathcal{S}_2 = \{\emptyset\}$, as the value of $\mathcal{S}_1$ is $\prod_{C_j \in T_k \triangle T'_{k'}} a_j = \prod_{C_j \in T_k} a_j \prod_{C_j \in T'_{k'}} a_j = -1$ and the value of $\mathcal{S}_2$ is $\prod_{C_j \in \emptyset} a_j = +1$.

We focus our attention on $\mathcal{S}_1$, which we relabel as $T_1 \dots T_k$ without loss of generality. We first note that $|T_1| = 1$, because the first element in the sequence must be an input axiom. Next, we use boundary expansion to show that $|T_k| > \eta n$. Indeed since we are deriving a value for $x_\emptyset$, we have that $\triangle_{C_j \in T_k}\Gamma(C_j) = \emptyset$, and because of the boundary expansion of C either $|B\Gamma(T_k)| \geq 0.4|T_k|$ or $|T_k| > \eta n$. But because we're deriving the value -1 it cannot be the case that T_k is empty, and so if it has any boundary expansion then $|\triangle_{C_j \in T_k} \Gamma(C_j)| \geq |B\Gamma(T_k)| > 0$, which is a contradiction because $|\triangle_{C_j \in T_k} \Gamma(C_j)| = 0$. Therefore $|T_k| > \eta n$ as desired.

Now we note that because T_ℓ is either an axiom or a symmetric difference of two previous sets, $|T_\ell| \leq 2\max_{\ell' < \ell} |T_{\ell'}|$. Therefore there must exist some set T in our sequence such that $|T| \in [\frac{\eta n}{2}, \eta n]$. Because $|T| \leq \eta n$, by expansion $|\triangle_{C_j \in T} \Gamma(C_j)| \geq |B\Gamma(T)| \geq 0.4|T| \geq 0.2\eta n$. Set $d = 0.01\eta n$. Then this sequence is invalid as the monomial derived at T has degree strictly larger than d. Therefore such a sequence cannot exist, and so our algorithm never fails. Thus with probability 0.99 a random 3XOR instance with value $\frac{1}{2} + \epsilon$ has a degree $\Theta(n)$ SoS pseudodistribution with value 1. □

3 Hardness reductions

As mentioned earlier in the lecture, the pseudodistribution for 3XOR can be used to show hardness of other problems. In particular, using a natural restriction of random 3SAT instances yielding 3XOR instances we get the following easy corollary.

Corollary 3.1. *Let ϕ be an instance of 3SAT on n variables with $m = c_{epsilon}n$ constraints (for c_ϵ a constant only dependent on ϵ) be chosen as follows: for each constraint we choose $i, j, k \sim [n]$, $e_i, e_j, e_k \sim \{0, 1\}$ iid and take our clause to be $(x_i^{e_i} \vee x_j^{e_j} \vee x_k^{e_k})$ (where $x^0 = x$, $x^1 = \overline{x}$). Then with probability at least* 0.99*,*

- *every assignment $x \in \{\pm 1\}^n$ satisfies at most $(\frac{7}{8} + \epsilon)m$ clauses*
- *there exists a pseudodistribution of degree $\Omega(n)$ such that in expectation all clauses of ϕ are satisfied*

where the probability is over the choice of ϕ.

Proof. Let ϕ be our random instance. For the soundness we leave it as an exercise to the reader to use the same argument as in the 3XOR case. For completeness we define $\phi_\oplus$ to be a 3XOR instance as follows: for clause $C : (x_i^{e_i} \vee x_j^{e_j} \vee x_k^{e_k})$, we have a constraint $C' : x_i x_j x_k = a_{ijk} = (-1)^{e_i + e_j + e_k}$. This is a random instance of 3XOR, as i, j, k were chosen iid and $(-1)^{e_i + e_j + e_k}$ is uniformly distributed over ± 1. Thus with probability 0.99 there exists a pseudodistribution satisifying all constraints in $\phi_\oplus$. The result follows by noting that if we transform this pseudodistribution back to $\{0, 1\}$ valued variables via $x \to \frac{1-x}{2}$, any assignment in the support of the pseudodistribution satisfies C', and any assignment satisfying C' also satisfies C. □

Random 3SAT isn't unique in terms of reducibility to 3XOR; in fact similar arguments can be shown to give good integrality gaps for a number of constant width constraint satisfaction problems. While these gaps are all constant at best, as every k-CSP has gap at most 2^{-k}, it turns out that similar hardness reduction strategies can be used to show *much* stronger lower bounds against SOS. In particular, Håstad's argument showing that Independent Set cannot be approximated to within $O(n^{1-o(1)})$ can be turned into a proof that degree $\Theta(n)$ SoS has the same integrality gap, based on reducing to a k-CSP with the right gap between the integrality gap and the number of ways to satisfy each constraint. This reduction also showcases a somewhat counterintuitive way of arguing lower bounds against SoS: once we have a relatively *weak* lower bound on some CSP, we can use the *power* of SoS to reduce this CSP to a much harder problem while preserving the completeness of the pseudodistribution, yielding a *stronger* lower bound. Hence when it comes to analyzing SoS on new hard problems like planted clique, it may not be clear a priori whether the power of SoS will be a blessing as with MAXCUT or a curse as with Independent Set.