Lower Bounds 2: Planted Clique

Lecture: Sam Hopkins
Scribe: Angelos Assos

November 2022

1 Introduction

As discussed in the previous lectures, it is important to know the capabilities, as well as the limitations of our powerful tool, the Sum of Squares Algorithm. Especially in problems that one believes that Sum Of Squares is the best Poly-time algorithm that can capture the complexity of this problem, lower bounds can hint us as to what can be done by Poly-Time algorithms in these particular problems.

2 Planted Clique Problem

The problem we are going to view through the lens of Sum of Squares is the Planted Clique problem. Let us start by defining it:

Definition 2.1 (Planted Clique Problem)**.** Consider the following two graph distributions:

1. $\mathcal{G}(n, \frac{1}{2})$ - the Random distribution,
2. $\mathcal{G}(n, \frac{1}{2})$ together with a 'planted' $k-$clique - the planted distribution,

Given a graph G has been sampled from one of the above two graph distributions, determine from which one it was sampled.

Notice that, for the Planted Clique problem, we get a sample from one of two distributions and the goal is to determine where the sample came from. Essentially, proving no-go results for the planted clique, will imply average case hardness for the problem; that is, given a probability distribution over possible inputs there is no polynomial time algorithm that solves the problem, given a sample from that distribution.
One good indicator that can indicate to us where the graph was sampled from is the maximum clique of the graph. We expect that if a graph G sampled from the Random Distribution, or otherwise if it is sampled from $\mathcal{G}(n, \frac{1}{2})$, then we expect the size of the maximum clique to be $2\log n$. Therefore, when we have that $k > 2.001\log n$, we should be able to Brute Force and determine whether

there exists a clique of size at least $2.001 \log n$ - then we are confident that the graph was sampled from the planted distribution, in $n^{O(\log n)}$ time.
Let us now shift our focus to 2 sub-problems:

Search: Given a Graph from the planted distribution, find a k-clique
Refutation: Given an arbitrary G, output `CERTIFY` if you can guarantee there does not exist a k-clique, otherwise output $'??'$. It should hold that:

$$Pr_{G \sim \mathcal{G}(n, \frac{1}{2})}(\texttt{output CERTIFY}) \geq 1 - o(1)$$

Note that having both the Search and the Refutation can guarantee us the distinguishability that the Planted Clique Problem requires.
So after all, when should Planted Clique be easy? The typical degrees in a Random graph G should be $\frac{n}{2} \pm O(\sqrt{n})$. When adding a planted clique of size k, we increase that to $\frac{n}{2} \pm O(\sqrt{n}) + k$, thus when $k \gg \sqrt{n \log n}$, one can just look at the maximum degree of the graph, which is going to be a good indicator for the max clique, and infer from which distribution the graph was sampled.
Earlier, we mentioned that we can Brute Force checking for a $2.001 \log n$ size clique. This can be done in $n^{O(\log n)}$, and SOS can also match that: for a clique of size $k = \frac{n}{2^d}$ we can have an $n^{O(d)}$ degree Sum of Squares that verifies if there exists a k clique.
In the next sections we are going to see that using SoS, does not help us infer anything about the maximum clique of a random graph. More specifically, in the case of the random graph it only tells us that the max clique has to be less than $O(\sqrt{n})$ - as depicted by theorem 1.

3 SOS algorithm for Planted Clique refutation

Given G and $x_1, x_2, ..., x_n \in \{0, 1\}$ we write the following constraints for a clique of size k:

$$C_k = \begin{cases} x_i^2 = x_i \\ x_i x_j = 0, i \not\sim j \texttt{ in } G \\ \sum_i x_i = k \end{cases}$$

We are going to check if there is a degree d refutation of C_k, in time $n^{O(d)}$. How can we refute this? Note that it was an exercise in Problem set 1 to prove that for $k \gg \sqrt{n}$, we have a refutation. Our hope, is to prove there is **no** SoS refutation of degree $d = O(1)$

Theorem 1. With high probability, if $G \sim \mathcal{G}(n, \frac{1}{2})$, then we have a degree d pseudo expectation $\tilde{\mathbb{E}}$ such that:

$$\tilde{\mathbb{E}} \models C_{n^{0.5-\epsilon}}$$

where $d = \Omega(\frac{\epsilon^2 \log n}{\log \log n})$

For the rest of the lecture we are going to go through the full proof of why there is no SoS refutation for $d = 2$, sketch the proof of $d = 4$ case, and talk briefly about what happens when $d \gg 4$

3.1 Degree $d = 2$ refutation - full proof

Let's reiterate what we want: We want with high probability to exist a pseudo-expectation $\tilde{\mathbb{E}}$ for which:

$$\tilde{\mathbb{E}} \models x_i^2 = x_i, x_i x_j = 0 \ \ \forall i \not\sim j \ \texttt{ and } \tilde{\mathbb{E}} \sum x_i = \Omega(\sqrt{n})$$

The proof is going to be similar to the proof for the Max-Cut lower bound from Lecture 7. We will need the to construct the pseudo-expectation, which is essentially just a map from a graph G to $\tilde{\mathbb{E}} x_i$ and $\tilde{\mathbb{E}} x_i x_j$.
Let's start by seeing what is forced upon us. We need to have

$$\tilde{\mathbb{E}} x_i x_j = 0, i \not\sim j$$

Let us set $\tilde{\mathbb{E}} x_i = \frac{k}{n}, \forall i \in [n]$ and $\tilde{\mathbb{E}} x_i x_j = \lambda$, which holds for edges in the graph, i.e. for $i \sim j$. We also have due to Cauchy Schwartz:

$$\tilde{\mathbb{E}}(\sum_i x_i)^2 \geq (\tilde{\mathbb{E}} \sum x_i)^2 \implies \sum_{i,j} \tilde{\mathbb{E}} x_i x_j \geq k^2 \implies \lambda \geq \frac{k^2}{|E|} \approx \frac{k^2}{n^2}$$

In order to find such a pseudo expectation, we need the matrix $M = (1, x)(1, x)^T$ to be PSD:

$$\tilde{\mathbb{E}} \geq 0 \iff \tilde{\mathbb{E}}(1, x)(1, x)^T \geq 0$$

Now, let us now try to construct the matrix of the pseudo-distribution:

$$M_2 = \left(\begin{array}{c|c} 1 & \vec{x} \\ \hline \vec{x} & A \end{array}\right)$$

Note that the for the matrix A we have:

$$A = (\frac{k}{n} - \lambda) I + \lambda A_G$$

where A_G is the adjacency matrix with 1's in the diagonals. Thus what's left to prove is that the following matrix is PSD:

$$\left(\begin{array}{c|c} 1 & \frac{k}{n}\vec{1} \\ \hline \frac{k}{n}\vec{1} & (\frac{k}{n} - \lambda) I + \lambda A_G \end{array}\right)$$

We focus on the lower right block of the matrix:

$$(\frac{k}{n} - \lambda) I + \lambda A_G = (\frac{k}{n} - \lambda) I + \lambda (\frac{J}{2} + \bar{A}_G)$$

where J is the all 1's matrix, and $\bar{A}_G$ satisfies $\bar{A}_G = A_G - \frac{1}{2}J$, i.e. it is A_G centered around 0. The above transformation lets us play with $\bar{A}_G$, whose values are $\frac{1}{2}$ and $-\frac{1}{2}$ instead of A_G whose values are 1 and 0. We can now use that with high probability $||\bar{A}_G|| \le O(\sqrt{n})$:

$$||\bar{A}_G|| \le O(\sqrt{n}) \implies \lambda \bar{A}_G + O(\lambda\sqrt{n})I \succeq 0$$

Using that fact we get:

$$(\frac{k}{n} - \lambda)I + \lambda(\frac{J}{2} + \bar{A}_G) \succeq (\frac{k}{n} - \lambda)I + \lambda\frac{J}{2} - O(\lambda\sqrt{n})I = (\frac{k}{n} - O(\lambda\sqrt{n}))I + \frac{\lambda}{2}J$$

Finally, for the whole matrix to be PSD, according to the Schur complement we will need:

$$(\frac{k}{n} - O(\lambda\sqrt{n}))I + (\frac{\lambda}{2} - (\frac{k}{n})^2)J \succeq 0$$

We need

$$\lambda \ge (\frac{k}{n})^2, \frac{k}{n} \ge \lambda\sqrt{n}$$

Which can be rewritten as:

$$\frac{k}{n\sqrt{n}} \ge \lambda \ge \frac{k^2}{n^2}$$

where $k \ll \sqrt{n}$. Clearly, it is possible for us to choose λ that satisfies the above, thus we have created a PSD matrix for the pseudo distribution and thus we are done.

3.2 The Trace Method

The natural thing to do is to try to extend the results for bigger d. Let us take a peek on what happens when we have $d = 4$. The matrix of the pseudo-expectation, $M_4 = \big((1,x)^{\otimes 2}\big)\big((1,x)^{\otimes 2}\big)^T$ will look like this:

$$M_4 = \begin{pmatrix} A_0 \in \mathbb{R}^{1\times 1} & A_1 \in \mathbb{R}^{1\times n} & A_2 \in \mathbb{R}^{1\times n^2} \\ A_2^T \in \mathbb{R}^{n\times 1} & A_3 \in \mathbb{R}^{n\times n} & A_4 \in \mathbb{R}^{n\times n^2} \\ A_3^T \in \mathbb{R}^{n^2\times 1} & A_4^T \in \mathbb{R}^{n^2\times n} & A_5 \in \mathbb{R}^{n^2\times n^2} \end{pmatrix}$$

We will be looking for a pseudo expectation that makes the above matrix PSD, subject to some constraints. One example constraint is that we need $\tilde{\mathbb{E}} x_i x_j x_k x_l = 0$ if $ijkl$ does not form a clique. One can notice that even for $d = 4$, it becomes increasingly harder to tune the parameters in order to get what we want, therefore, we might need to different methods to come up with ways to tackle the more general cases.

Let's pause the thinking for the $d = 4$ for a while. We would like to equip ourselves with machinery that is going to help us prove the cases for bigger d.

When we have a scalar random variable x, we can use the moments of x to get a simple tail bound for x as follows:

$$Pr(x > t) = Pr(x^k > t^k) \leq \frac{\mathbb{E}x^k}{t^k}$$

We would like to develop such a tool for matrices too, in order to bound the magnitude of the matrix. One can bound the magnitude of the matrix in the following way:

$$\mathbb{E}[||M||] \leq (\mathbb{E}\lambda_{max}^{2k})^{\frac{1}{2k}} \leq \mathbb{E}[Tr(M)^{2k}]^{\frac{1}{2k}}$$

Note that we might be losing a factor of $n^{\frac{1}{2k}}$, since we just use the largest eigen-value and we know:

$$\lambda_{min} \leq \frac{Tr(M)}{n} \leq \lambda_{max}$$

However, this is not such a big loss for us since for small k, even for $k = \log n$ this loss becomes just a constant.
Let us go through some examples. Let us try the above with $k = 1$, for adjacency matrix of a random graph M, for which we have independent entries with

$$M_{ij} = \begin{cases} 1, \ \texttt{w.p.} \ \ \frac{1}{2} \\ -1, \ \texttt{w.p.} \ \ \frac{1}{2} \end{cases}$$

and $M_{ij} = M_{ji}$. We will have:

$$\mathbb{E}[Tr(M^2)] = \sum_i \sum_j \mathbb{E}M_{ij}M_{ji} = \sum_{i,j} \mathbb{E}[M_{ij}^2] = n^2$$

So we get that $\mathbb{E}[||M||] \leq n$. Can we do better? Let's try for $k = 4$:

$$\mathbb{E}[Tr(M^4)] = \sum_{ijkl} \mathbb{E}M_{ij}M_{jk}M_{kl}M_{li}$$

Note that when all i, j, k, l are distinct, their total contribution to the sum is to 0 as they cancel out. However, when we have that only 3 of them are distinct, for example when $l = j$, we get:

$$\sum_{ijk} \mathbb{E}M_{ij}M_{jk}M_{kj}M_{ji} = \sum_{ijk} \mathbb{E}M_{ij}^2 M_{jk}^2 = O(n^3) \implies \mathbb{E}||M|| \leq O(n^{3/4})$$

For general k we have:

$$\mathbb{E}[Tr(M^{2k})] = \sum_{i_1,\dots,i_{2k}} \mathbb{E}[\prod_j M_{i_j i_{j+1}}]$$

One thing that we can notice is that we need each edge $e_j = (i_j, i_{j+1})$ to appear an even number of times, in order for the term to be nonzero in expectation. In the case where it appears an odd number of times, then that means that the net contribution of the edge, to the sum is 0. To see this just note that the term $(e_1, e_2, ..., e_{j-1}, 1, e_{j+1}, ..., e_{2k})$ is exactly cancelling with $(e_1, e_2, ..., e_{j-1}, -1, e_{j+1}, ..., e_{2k})$. So we need each labelling $(i_1, ..., i_{2k})$ to be double covered, i.e. contain each pair (i_j, i_{j+1}) twice. Therefore for general k, the following claim is useful:

Claim: Any double covering labeling of length $2k$ has at most $k+1$ distinct labels.

Sketch: Start with a $2k$ vertex cycle-graph of vertices $i_1, i_2, ..., i_{2k}$, possibly with repeated vertices. Start collapsing vertices with the same labelling. Note that every time we collapse a vertex the graph remains connected. In the end we have at most k edges, therefore at most $k+1$ vertices, i.e. at most $k+1$ labels, as required.

For general k, the above claim gives us a bound for the trace:

$$\mathbb{E}[Tr(M^{2k})] \leq n^{k+1}(2k)^{k-1}$$

Which can give a bound

$$\mathbb{E}[||M||] \leq [n^{k+1}(2k)^{k-1}]^{\frac{1}{2k}} \approx \sqrt{n} \cdot (2k)^{\frac{1}{2}} \leq C\sqrt{n \log n}$$

3.3 Trace method for C_4

Let C_4 be a $n^2 \times n^2$ matrix, where:

$$C_{ijkl} = \begin{cases} 1 \texttt{ if ijkl is a 4-clique in G} \\ 0, \texttt{ o.w.} \end{cases}$$

Notice that if ij or kl are not edges in G, the whole row/column is going to be 0. Let us take the matrix D_4 which is of dimensions $|E| \times |E|$, where E is the number of edges of the random graph G. Then C_4 and $\bar{C}_4$, which is the 'centered' version of C_4, i.e. $\bar{C}_4 = C_4 - \mathbb{E}C_4$ can be written as:

$$C_4 = \begin{pmatrix} 0 & 0 \\ 0 & D_4 \end{pmatrix},$$

$$\bar{C}_4 = \begin{pmatrix} 0 & 0 \\ 0 & \bar{D}_4 \end{pmatrix}$$

, where $\bar{D}_4 = D_4 - \mathbb{E}D_4 = D_4 - \frac{J}{2^4}$, since we have the 4 indices i, j, k, l form a clique with probability $\frac{1}{2^4}$, conditioned on the fact that $i \sim j$ and $k \sim l$. We ultimately want to write a bound for the magnitude of $\bar{C}_4$. It is not hard to see that by bounding the magnitude of $\bar{D}_4$, we can get the same magnitude bound for $\bar{C}_4$. We have:

$$\mathbb{E}[Tr(\bar{D}_4^4)] = \sum_{i_1j_1, i_2j_2, i_3j_3, i_4j_4} \mathbb{E}[(1_{i_1j_1, i_2j_2} - 2^{-4})(1_{i_2j_2, i_3j_3} - 2^{-4})(1_{i_3j_3, i_4j_4} - 2^{-4})(1_{i_4j_4, i_1j_1} - 2^{-4})]$$

Note that the indicator variables 1_{xyzw} indicate whether $xyzw$ is a clique or not. As mentioned before, we have that each 4-tuple makes a clique with probability $\frac{1}{2^4}$, conditioned on the fact that edges i_a, j_a exist. Now, notice that the contribution of tuples $(i_1, i_2, i_3, i_4, j_1, j_2, j_3, j_4)$ where all of the labels are distinct is 0. Therefore we have that the terms that contribute to the overall sum, must have at most 7 distinct labels, and since there are at most $O(n^7)$ of such terms,

$$\mathbb{E}[Tr(\bar{D}_4^4)] \leq O(n^7)$$

Thus:

$$\mathbb{E}||\bar{D}_4|| \leq \mathbb{E}[Tr(\bar{D}_4^4)]^{\frac{1}{4}} \leq O(n^{7/4})$$

Which bounds $||\bar{D}_4||$, and subsequently $||\bar{C}_4||$ to $O(n^{\frac{7}{4}})$.

3.4 Degree $d = 4$ sketch

Back to the planted clique for $d = 4$, recall that we want to construct a pseudo-expectation $\tilde{\mathbb{E}}$ for which we want:

$$\tilde{\mathbb{E}}1 = 1$$

$$\tilde{\mathbb{E}}x_i = \frac{k}{n}, \forall i \in [n]$$

$$\tilde{\mathbb{E}}x_i^2 x_j^2 = \tilde{\mathbb{E}}x_i x_j \begin{cases} \lambda_2, i \sim j \\ 0, \ \texttt{o.w.} \end{cases}$$

$$\tilde{\mathbb{E}}x_i x_j x_k = \begin{cases} \lambda_3, \ \texttt{if ijk clique} \\ 0, \ \texttt{o.w.} \end{cases}$$

$$\tilde{\mathbb{E}}x_i x_j x_k x_l = \begin{cases} \lambda_4, \ \texttt{if ijkl clique} \\ 0, \ \texttt{o.w.} \end{cases}$$

We also have from Cauchy Schwarz:

$$(\tilde{\mathbb{E}} \sum_i x_i)^4 \leq \tilde{\mathbb{E}}(\sum_i x_i)^4 = \sum_{ijkl} \tilde{\mathbb{E}}x_i x_j x_k x_l$$

The pseudo-expectation matrix is going to look like:

$$M_4 = \begin{pmatrix} A_0 & A_1 & A_2 \\ A_2^T & A_3 & A_4 \\ A_3^T & A_4^T & A_5 \end{pmatrix} = \begin{pmatrix} 1 & \tilde{\mathbb{E}}x & \tilde{\mathbb{E}}x \otimes x \\ \tilde{\mathbb{E}}x & \tilde{\mathbb{E}}xx^T & \tilde{\mathbb{E}}x(x \otimes x)^T \\ \tilde{\mathbb{E}}x \otimes x & \tilde{\mathbb{E}}x(x \otimes x)^T & \tilde{\mathbb{E}}(x \otimes x)(x \otimes x)^T \end{pmatrix}$$

The goal here is to assign the parameters so that the above matrix is PSD. We are going to show how we can get the block matrix A_5 with dimensions $n^2 \times n^2$ to be PSD. Once the other block matrices are shown to be PSD, we can apply the Schur complement to get that the whole matrix M_4 is PSD.

Similarly with before, given that $i \not\sim j$, then the whole row/column will be just zeros for that entry. Thus we consider the matrix A_5', which is just the matrix of dimensions $|E|\times|E|$, where each row/column represents one edge in the graph G. For A_5', each entry in the diagonal (ij, ij) needs to have $\tilde{\mathbb{E}}x_i^2x_j^2 = \tilde{\mathbb{E}}x_ix_j = \lambda_2$: this can be written as $\lambda_2 I$. Moreover, for (ij, kl) where only 3 indices are distinct, we will have $\tilde{\mathbb{E}}x_ix_jx_k = \lambda_3$ if these three indices form a clique in the original graph otherwise we have 0; this can be described by the matrix $\lambda_3 D_3$, where D_3 is the $n^2 \times n^2$ matrix with 1 on (ij, kl) if amongst i, j, k, l we have three distinct indices and they form a clique. Finally if all 4 indices are different $\tilde{\mathbb{E}}x_ix_jx_kx_l = \lambda_4$ otherwise the entry is 0; this can be described by the matrix $\lambda_4 D_4$, where D_4 is the $n^2 \times n^2$ matrix with 1 on (ij, kl) if amongst i, j, k, l we have four distinct indices and they form a clique (this D_4 is the same matrix as the D_4 used in section 3.3). We can then write A_5' as:

$$A_5' \approx \lambda_2 I + \lambda_3 D_3 + \lambda_4 D_4$$

In combination to what we found in section 3.3 we get

$$A_5' \approx \lambda_2 I + \lambda_3 D_3 + \lambda_4(\bar{D}_4 + J \cdot 2^{-4}) \succeq [\lambda_2 - \lambda_4 O(n^{\frac{7}{4}})]I + \lambda_3 D_3 + J \cdot 2^{-4}$$

We can also center D_3 by writing $\bar{D}_3 = D_3 - \frac{J}{2}$ and using that $||\bar{D}_3||$ can be bounded by $\sqrt{n}$, the above becomes:

$$A_5' \succeq [\lambda_2 - \lambda_4 O(n^{\frac{7}{4}}) - \lambda_3 O(\sqrt{n})]I + \lambda_4 J \cdot 2^{-4} + \lambda_3 \frac{J}{2} \succeq [\lambda_2 - \lambda_4 O(n^{\frac{7}{4}}) - \lambda_3 O(\sqrt{n})]I$$

Thus we need:

$$\lambda_2 \approx \frac{k^2}{n^2} \gg (\frac{k^4}{n^4})n^{\frac{7}{4}}$$

and

$$\lambda_2 \approx \frac{k^2}{n^2} \gg (\frac{k^3}{n^3})\sqrt{n}$$

which is satisfied when $k \ll n^{\frac{1}{8}}$. We have now that A_5 is PSD when $k \ll n^{\frac{1}{8}}$. to finish it, one can determine the conditions that give the other block matrices to be PSD, do the Schur complement, and get a lower bound for k.

3.5 Kelner's counter-example

In the previous sections, we have seen a 'naive' construction of the pseudoexpectation. The naive construction of a degree d pseudoexpectation uses $\lambda_1, ..., \lambda_d$, for each of which we have $\lambda_i = (\frac{k}{n})^i$, and sets for every $S \subset [n], |S| \leq d$:

$$\tilde{\mathbb{E}}[x_S] = \begin{cases} \lambda_{|S|}, & \texttt{if S is a clique} \\ 0, & \texttt{o.w.} \end{cases}$$

It worked for $d = 2$ and we did a sketch of how it can be done for $d = 4$. Therefore, a natural question is whether the 'naively' constructed pseudoexpectation, is PSD, and if so, for which k. Kelner provided a negative result and showed that for degree $d = 4$ the naive construction gives a matrix for the pseudo distribution that is not PSD, for a specific regime of k - more specifically for $k \gg n^{\frac{1}{3}}$. The construction goes as follows: Let $r_i(x) = \sum_{j=1}^{n}(1_{i\sim j} - \frac{1}{2})x_j$ and $P(x) = \sum r_i(x)^4$. Let also $r_{ij} = 1_{i\sim j} - \frac{1}{2}$. Suppose we have constructed $\tilde{\mathbb{E}}$, then we have:

$$\tilde{\mathbb{E}}P(x) = \tilde{\mathbb{E}}\sum r_i(x)^4 \geq \tilde{\mathbb{E}}\sum x_i r_i(x)^4 = \sum_{i,j,k,l,s} r_{ij}r_{ik}r_{il}r_{is}\tilde{\mathbb{E}}x_i x_j x_k x_l x_s = \sum_{i,j,k,l,s} \frac{1}{2^4}\tilde{\mathbb{E}}x_i x_j x_k x_l x_s =$$

$$= \frac{1}{2^4}\tilde{\mathbb{E}}[(\sum x_i)^5] \geq \Omega(k^5)$$

On the other hand now, building the pseudo distribution naively gives us:

$$\tilde{\mathbb{E}}P(x) = \tilde{\mathbb{E}}\sum r_i(x)^4 = \sum_{i,j,k,l,s} r_{ij}r_{ik}r_{il}r_{is}\tilde{\mathbb{E}}x_j x_k x_l x_s$$

We have the following cases:

- If in the set of variables $\{x_j, x_k, x_l, x_s\}$ we have at least three different variables, we have that the sum of all these terms, in expectation over G, is going to be 0.
- If we have exactly 2 distinct variables in $\{x_j, x_k, x_l, x_s\}$ then we have n^2 such choices of pairs, and also n choices for i. For each of these choices, the value of the pseudo-expectation is needed to be $\frac{k^2}{n^2}$ (as also seen in the above section, for λ_2). This gives us value $\frac{k^2}{n^2} \cdot n^2 \cdot n = nk^2$.
- If we have exactly 1 distinct variables in $\{x_j, x_k, x_l, x_s\}$ then we have n such choices of the variables, and also n choices for i. For each of these choices, the value of the pseudo-expectation is needed to be $\frac{k}{n}$ (as also seen in the above section, for λ_1). This gives us value $\frac{k}{n} \cdot n \cdot n = nk$.

The above leads to the inequality:

$$\mathbb{E}\tilde{\mathbb{E}}\sum_i r_i(x)^4 \leq O(nk) + O(nk^2) \leq O(nk^2)$$

Clearly we have $\tilde{\mathbb{E}}P(x) \geq k^5$ but also we have now that $\tilde{\mathbb{E}}P(x) \leq nk^2$. Therefore we require that

$$nk^2 \geq k^5$$

which if we plug in $k \gg n^{\frac{1}{3}}$ does not hold.
So what went wrong here? We had constructed $P(G, x)$ such that:

$$\tilde{\mathbb{E}}_{naive}P(G, x) \neq \tilde{\mathbb{E}}_{x,G\sim planted}P(G, x)$$

That was actually an SoS proof that:

$$\mathcal{C}_6 \vdash P(G, x) \geq \Omega(\tilde{\mathbb{E}}_{x,G\sim planted}P(G, x))$$