Reduction from Planted Clique to Robust Sparse Mean Estimation

Sam Hopkins

December 4, 2023

These notes have not been subjected to the usual scrutiny of a formal publication. Caveat emptor!

In these lectures (22b and 23) we cover a reduction from a variant of planted clique to robust sparse mean estimation due to Brennan and Bresler. The majority of the lectures cover section 8.3.1 of the Diakonikolas-Kane book. Read that section first.

Our presentation diverges from the Diakonikolas-Kane version in the construction of the matrix A in the proof of Lemma 8.28. Here's the proof we presented in class. Translating from the notation in Diakonikolas-Kane Lemma 8.28, we have taken $\varepsilon = 0.01, \delta = 1$.

Lemma 0.1. *Let $s, n \in \mathbb{N}$ and $\eta > 0$. Suppose that $k \geq C(\log s \cdot s + n)/\eta^2$ for a sufficiently-large constant C. Then there's a matrix A with s rows and n columns where each column has $0.99s$ entries equal to $\frac{1}{\eta\sqrt{k}}$ and $\|A\| \leq 1$.*

Proof. Let A be a random matrix with independent columns $w_1, \ldots, w_n$, where each column has a random subset of $0.99s$ entries equal to $1/\eta\sqrt{k}$, and the other $0.01s$ entries equal to $(0.99/0.01) \cdot (-1/\eta\sqrt{k})$. Here we chose $(0.99/0.01)$ so that the column-sums are 0. By construction, each column of A has $0.99s$ entries equal to $1/\eta\sqrt{k}$.

Now we would like to calculate $\mathbf{E} \sum_{i \leq n} w_i w_i^\top$ and use Matrix Bernstein to bound $\mathbf{E} \| \sum_{i \leq n} w_i w_i^\top - \mathbf{E} \sum_{i \leq n} w_i w_i^\top \|$ – we can use Matrix Bernstein here because each term $w_i w_i^\top - \mathbf{E}\, w_i w_i^\top$ in the sum above is independent. This would be easy if the entries of each w_i were independent of each other, but they are not, because we constructed each to have $0.99s$ entries equal to $1/\sqrt{k}$. We will first show that the difference between our choice of random vector w and a similar one w' where we choose the entries independently is minimal.

Claim 0.2. *Let $a, b \leq s$. Then $\mathbf{E}\, w(a)w(b) = \mathbf{E}\, w'(a)w'(b) \pm O(1/(s\eta^2 k))$, where w' is a random vector where each entry is $1/\eta\sqrt{k}$ with probability 0.99 and $(0.99/0.01) \cdot (-1/\eta\sqrt{k})$ with probability 0.01, independently.*

Proof of claim. Follows by explicitly calculating $\mathbb{P}(w(a) = 1/\sqrt{k}, w(b) = 1/\eta\sqrt{k})$, and so on. □

This claim implies that $\| \mathbf{E}\, w_i w_i^\top - \frac{1}{\eta^2 k} I \| \leq \frac{1}{\eta^2 k}$, and hence $\| \mathbf{E} \sum_{i \leq n} w_i w_i^\top \| \leq \frac{n}{\eta^2 k}$. This is at most 0.00001 by hypothesis.

Using the claim again, together with the observation $\|w_i\|^2 \leq \frac{s}{\eta^2 k}$ with probability 1, we can bound the variance term for matrix Bernstein as

$$\left\| \mathbf{E} \sum_{i \leq n} \|w_i\|^2 w_i w_i^\top \right\| \leq O(n) \cdot \frac{s}{\eta^2 k} \cdot \frac{1}{\eta^2 k} \,.$$

From Matrix Bernstein we can then argue

$$\mathbf{E} \left\| \sum_{i \leq n} w_i w_i^\top - \mathbf{E}\, w_i w_i^\top \right\| \leq O(\sqrt{\log s}) \cdot (\frac{ns}{\eta^4 k^2}) + O(\log s) \frac{s}{\eta^2 k} \,.$$

Putting this together with our earlier bound on $\| \mathbf{E} \sum_{i \leq n} w_i w_i^\top \|$ and using our assumed bound on k finishes the proof.

$\square$