CS265, Fall 2022

Class 18: Agenda and Questions

1 Warm-Up

Suppose that X is a k-source on $\{0,1\}^n$. Let $N = 2^n$. Let $\sigma \in \mathbb{R}^N$ be the "vectorized" version of the pmf of X. That is,

$$\sigma_i = \Pr[X = i] \qquad \forall i \in \{0, \dots, N-1\},$$

where we associate a number $i < N$ with its binary expansion in $\{0,1\}^n$.

1. Why is $\|\sigma\|_\infty \le 2^{-k}$?
2. Argue that $\|\sigma\|_2 \le 2^{-k/2}$.
   Hint: *Use the fact that for any vector x, $\|x\|_2^2 \le \|x\|_\infty \|x\|_1$ (why is this true?).*

2 Announcements

- HW8 (THE LAST ONE!) is out now, due Friday!
- Thursday will be research talks! No mini-lectures to watch.
- FINAL EXAM: Thursday 12/15, 12:15-3:15pm, Room 420-040.
  - Practice exam out soon.

3 Questions?

Any questions from the minilectures and/or the quiz and/or the warm-up? (Expanders, extractors?)

4 Recap

Recall the definition of a (k, ε)-extractor:

Definition 1. *A function* $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ *is a* (k, ε) *extractor if, for all* k*-sources* X *on* $\{0,1\}^n$, $\|\mathsf{Ext}(X, U_d) - U_m\| \le \varepsilon$.

Above, $\|\cdot\|$ is the total variation distance, and U_d refers to the uniform distribution on d bits.

Suppose that $G = (V, E)$ is an (undirected, unweighted) degree-D expander graph with $|V| = N$, and with expansion parameters $\lambda(G) \leq 1/2$. Recall that

$$\lambda(G) = \max\{\lambda_2, |\lambda_N|\},$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ are the eigenvalues of A, where A is the *normalized adjacency matrix* of G. (aka, A_{ij} is $1/D$ if $\{i, j\} \in E$ and is zero otherwise).

5 Extractors *from* expanders

At this point, there will be some slides illustrating a construction of an extractor. A description is below for reference.

Let $\varepsilon > 0$. Let $N = 2^n$ and fix some arbitrary bijection between $\{0,1\}^n$ and V, where V is the vertex set of G above. Fix any $k \leq n$. Let $d = \log(D) \cdot \ell$, where

$$\ell = (n - k)/2 + \log(1/\varepsilon),$$

.

Consider the following function $\mathsf{Ext} : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^n$.

On input $x, s \in \{0,1\}^n \times \{0,1\}^d$:

- Treat $x \in \{0,1\}^n$ as an element of V.
- Treat $s \in \{0,1\}^d$ as a string of ℓ numbers in $\{0, 1, \ldots, D\}$. (That is, break up s into ℓ chunks, each $\log(D)$ bits long). Say these numbers are $a_1, a_2, \ldots, a_\ell$.
- Consider the following walk on G: let $x^{(0)} = x$. For $i = 1, 2, \ldots, \ell$, get from $x^{(i-1)}$ to $x^{(i)}$ by choosing the a_i'th neighbor of $x^{(i-1)}$.
- output $x^{(\ell)} \in V$, which we treat as an element of $\{0,1\}^n$.

That is, we use the source x to tell us where to start a walk, we use the seed s to tell us how to take a random walk (notice that if $s \sim U_d$ is uniform, then we really are taking a random walk), and after we've walked ℓ steps, we output whatever vertex we happen to be on.

6 Group work: this is a good extractor!

In this section, you'll show that Ext is a (k, ε) extractor.

Group Work

1. Let σ be the "vectorized" pmf of X (as in the warm-up). Explain why the distribution of $\mathsf{Ext}(X, U_d)$ is given by $A^\ell \cdot \sigma$. (Recall that A is the normalized adjacency

matrix of G).

2. Let $\pi = \frac{1}{N}\mathbf{1}$ be the vector that corresponds to the uniform distribution. Explain why
$$\|U_n - \mathsf{Ext}(X, U_d)\| = \|\pi - A^\ell \cdot \sigma\| \leq \frac{\sqrt{N}}{2}\lambda(G)^\ell \|\pi - \sigma\|_2.$$
Hint: *Mimic a computation that we did in the Expanders minilecture to show that random walks mix quickly when $\lambda(G)$ is small.*

3. Argue that $\|\pi - \sigma\|_2 \leq 2 \cdot 2^{-k/2}$.
Hint: *By the triangle inequality, $\|\pi-\sigma\|_2 \leq \|\pi\|_2 + \|\sigma\|_2$.* ***Hint***: *Use the warm-up!*

4. Assume that G is a good enough expander that $\lambda(G) \leq 1/2$. (It turns out that these exist for large enough degrees D). Conclude that $\|U_n - \mathsf{Ext}(X, U_d)\| \leq \varepsilon$ and thus Ext is a (k, ε) extractor.

Group Work: Solutions

1. To see why $Ext(X, U_d) \sim A^\ell \sigma$, notice that by definition, $Ext(X, U_d)$ is the outcome of ℓ steps of a uniformly random walk on G, if we start from the distribution X on the vertices of G. As we saw in the Markov chain unit, this is given by $A^\ell \sigma$.
(Notice that I've switched from left-multiplying to right-multiplying. In general it doesn't matter as long as we're consistent, but in this case it really doesn't matter since A is symmetric).

2. To bound $\|U_n - Ext(X, U_d)\|_{TV}$, we first note that
$$\|U_n - Y\|_{TV} = \frac{1}{2}\|\pi - A^\ell \sigma\|_1$$
by the def. of total variation distance. Then we can write
$$\begin{aligned} \|\pi - A^\ell \sigma\|_1 &= \|A^\ell(\pi - \sigma)\|_1 && \text{since } A\pi = \pi \\ &\leq \sqrt{N}\|A^\ell(\pi - \sigma)\|_2 && \text{Cauchy-Schwarz} \\ &\leq \sqrt{N}\lambda(G)^\ell \|\pi - \sigma\|_2 \end{aligned}$$
where the last line follows since $(\pi - \sigma) \perp \pi$ and π is the top eigenvector of A. To see that $(\pi - \sigma) \perp \pi$, note that
$$\sum_i \pi_i(\pi_i - \sigma_i) = \frac{1}{N}\sum_i(\pi_i - \sigma_i) = \frac{1}{N}(\sum_i \pi_i - \sum_i \sigma_i) = \frac{1}{N}(1-1) = 0.$$

3. Following the hints,

$$\begin{aligned}\|\pi - \sigma\|_2 &\leq \|\pi\|_2 + \|\sigma\|_2 \\ &\leq 2^{-n/2} + 2^{-k/2} \\ &\leq 2^{-k/2+1}\end{aligned}$$

using the warm-up to bound $\|\sigma\|_2$ and the fact that $n \geq k$ in the final line.

4. To show that Ext is a (k, ε) extractor, we need to show that $\|U_n - Ext(X, U_d)\|_{TV} \leq \varepsilon$. To do that, we put together all the pieces:

$$\begin{aligned}\|U_n - Ext(X, U_d)\|_{TV} &\leq \frac{\sqrt{N}}{2}\lambda(G)^\ell \|\pi - \sigma\|_2 \\ &\leq \frac{\sqrt{N}}{2} 2^{-\ell} \cdot 2^{1-k/2} \\ &= 2^{n/2-1} \cdot 2^{-\ell} 2^{1-k/2} \\ &= 2^{(n-k)/2-\ell} \\ &= 2^{\log(1/\varepsilon)} \\ &= \varepsilon,\end{aligned}$$

where we used the choice of $\ell = (n-k)/2 + \log(1/\varepsilon)$ in the final line.