Lecture 22: Hardness of Max-E3Lin

April 3, 2008

Lecturer: Ryan O'Donnell *Scribe: Evan Danaher*

1 Overview

This lecture is beginning the proof of $1-\epsilon$ vs. $1/2+\epsilon$ hardness of Max-E3Lin. The basic idea is similar to previous reductions; reduce from Label-Cover using a gadget that creates 2^K variables corresponding to the key vertices and 2^L vertices corresponding to the label vertices, where they correspond in the usual way to $\{0,1\}^K$ and $\{0,1\}^L$.

We then want to select some subsets x, y, z of these strings to use as the constraints in the Max-E3Lin instance, and whether each constraint is equal to 0 or 1. We then want take the solution to the Max-E3Lin instance, and "decode" it into a good solution to the original Label-Cover problem, thereby showing that solving the Max-E3Lin instance was hard.

Completeness requires that the dictator function $f_i(x) = x_i$ pass with probability at least $1-\epsilon$, while soundness requires that any function which passes with probability at least $1/2+\epsilon$ should decode to some small and nonempty list $\mathrm{Sugg}(f)$ of coordinates in the Label-Cover.

Finally, to check these conditions, Fourier analysis will be useful. So instead of using $0, 1$ and $\oplus$ (exclusive or), we use $-1, 1$ and $\cdot$ (multiplication).

2 The Gadget

As noted above, we need the dictator functions $f_i(x) = x_i$ to pass each constraint with probability at least $1-\epsilon$. Ignoring the details of how this relates to label cover, consider what this means about the 3Lin conditions.

If $f(x)\cdot f(y)\cdot f(z) = 1$ is a constraint for some x, y, z, then we must have each $x_i \cdot y_i \cdot z_i = 1$ with high probability for the dictators to pass (completeness). However, if too many functions pass this with probability at least $1/2+\epsilon$, then we have to be able to decode these into good solutions for Label-Cover (soundness). So we need to make sure that the dictators all pass, but not too many other functions do.

2.1 Naive Attempt

One possibility is to design the triples as simply as possible while passing all the dictators; simply pick triples so that the parity of each bit is correct for $f_i(x)\cdot f_i(y)\cdot f_i(z) = 1$ for each dictator. This can be formulated in two equivalent ways:

- For each constraint, pick each triple (x_i, y_i, z_i) independently and uniformly from the four triples $\{(-1,-1,1),(-1,1,-1),(1,-1,-1),(1,1,1)\}$ which satisfy $x_i \cdot y_i \cdot z_i = 1$.
- For each constraint, pick $x, y \in \{-1,1\}^K$ independently and uniformly, then take $z = x \circ y$ (where $\circ$ is the bitwise multiplication operation).

The second formulation demonstrates that we can treat x and y as independent and z as dependent on them, which will be useful below.

This algorithm clearly satisfies completeness, since it is designed so that each dictator function is always satisfied.

Now consider soundness. To find the probability that an arbitrary function $f : \{-1,1\}^K \to \{-1,1\}$ passes, use the indicator $1/2 + 1/2 f(x) \cdot f(y) \cdot f(z)$. If f passes the condition, this is $1/2 + 1/2 \cdot 1 = 1$, and if it fails, it is $1/2 + 1/2(-1) = 0$. So

$$\begin{aligned}
\mathbf{Pr}[f \text{ passes}] &= \mathbf{E}_{x,y}[1/2 + 1/2 f(x) \cdot f(y) \cdot f(z)] \\
&= 1/2 + 1/2\mathbf{E}_{x,y}[f(x) \cdot f(y) \cdot f(z)] \\
&= \frac{1}{2} + \frac{1}{2}\mathbf{E}_{x,y}\left[\sum_{S \in [K]} \hat{f}(S)\chi_S(x) \cdot \sum_{T \in [K]} \hat{f}(T)\chi_T(y) \cdot \sum_{U \in [K]} \hat{f}(U)\chi_U(z)\right] \\
&= \frac{1}{2} + \frac{1}{2}\sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y}\left[\chi_S(x)\chi_T(y)\chi_U(z)\right] \\
&= \frac{1}{2} + \frac{1}{2}\sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y}\left[\chi_S(x)\chi_T(y)\chi_U(x \circ y)\right] \\
&= \frac{1}{2} + \frac{1}{2}\sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y}\left[\chi_{S \triangle U}(x)\chi_{T \triangle U}(y)\right] \\
&= \frac{1}{2} + \frac{1}{2}\sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y}[\chi_{S \triangle U}(x)]\mathbf{E}_{x,y}[\chi_{T \triangle U}(y)] \\
&= \frac{1}{2} + \frac{1}{2}\sum_{S \in [K]} \hat{f}(S)^3
\end{aligned}$$

(The final reduction uses the fact that $\mathbf{E}[\chi_{S \triangle U}(x)]$ is 0 if $S \triangle U$ is nonempty, and 1 otherwise; thus, the terms with $S \neq U$ disappear, while in the remaining terms $\mathbf{E}[\chi_{S \triangle U}(x) = 1$. Likewise, we must have $T = U$, so all three sets are equal.)

So any function with Fourier support of size 1 with coefficient 1 will pass with probability $1/2 + 1/2(1) = 1$. While this does include the dictator functions, it also includes several other functions, notably any parity function and the constant function.

This is unfortunate, since soundness will force us to decode these functions into solutions for the original Label-Cover instance. The parity function on all bits will be difficult to decode into small sets because of its global nature; there is no clear way to get small sets out of it. The constant functions is even worse, since it depends in no way on its inputs.

Thus, we look for a better test to excludes these troublesome cases.

2.2 Fixing the Constant Functions

First, we modify the test to remove the constant function from the set that needs to be decoded.

The problem is that all the constraints chosen had right-hand side 1, so setting all values on the left to 1 satisfies them all. By changing half of these to -1, we ensure that both constant function passes with probability $1/2$; since $1/2 < 1/2 + \epsilon$, we no longer have to decode them.

To do so, we have a new algorithm, which can again be expressed in two ways:

- For each constraint, pick $b \in \{0, 1\}$, then pick each triple (x_i, y_i, z_i) independently and uniformly from the four triples which satisfy $x_i \cdot y_i \cdot z_i = b$.
- For each constraint, pick $b \in \{0, 1\}$ and $x, y \in \{-1, 1\}^K$ independently and uniformly, then take $z = x \circ y \circ \overline{b}$ (where $\overline{b} = (b, b, \ldots b)$).

Dictators still pass, and now the constant functions only pass with probability $1/2$. The analysis for a general function goes roughly as before, except that we must also take b into account:

$$\begin{aligned}
\mathbf{Pr}[f \text{ passes}] &= \mathbf{E}_{x,y,b}[1/2 + 1/2b \cdot f(x) \cdot f(y) \cdot f(z)] \\
&= 1/2 + 1/2\mathbf{E}_{x,y,b}[b \cdot f(x) \cdot f(y) \cdot f(z)] \\
&= \frac{1}{2} + \frac{1}{2}\mathbf{E}_{x,y,b}\left[b \cdot \sum_{S \in [K]} \hat{f}(S)\chi_S(x) \cdot \sum_{T \in [K]} \hat{f}(T)\chi_T(y) \cdot \sum_{U \in [K]} \hat{f}(U)\chi_U(z)\right] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y,b}\left[b\chi_S(x)\chi_T(y)\chi_U(z)\right] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y,b}\left[b\chi_S(x)\chi_T(y)\chi_U(x \circ y \circ \overline{b})\right] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y,b}\left[b\chi_{S \triangle U}(x)\chi_{T \triangle U}(y)\chi_U(\overline{b})\right] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y,b}[\chi_{S \triangle U}(x)]\mathbf{E}_{x,y,b}[\chi_{T \triangle U}(y)]\mathbf{E}_{x,y,b}[b\chi_U(\overline{b})] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S \in [K],\, |S| \text{ odd}} \hat{f}(S)^3
\end{aligned}$$

(The final reduction again uses that we must have $S = T = U$; further, $b\chi_U(\overline{b}) = (-1)^{|U|+1} + 1^{|U|+1}$, so is 0 if $|U|$ is even and 1 if $|U|$ is odd.)

This is still reasonable, since the dictators f_i have $\hat{f}(S) = 1$ on $S = \{i\}$, which has odd size, while the constant functions have no odd terms in their Fourier support, so now pass with probability $1/2$.

However, parity functions on an odd number of variables still pass, and in particular, either the parity of all bits or the parity of all but one will still pass; these are still too large to decode reasonably. We thus need one more revision to eliminate these cases.

2.3 Removing Large Parity Functions

In the constructions so far, the dictator functions have passed with probability 1. This corresponds to satisfying every constraint, and it is known how to explicitly find such solutions. Thus, decoding these solutions into good Label-Cover solutions would show Label-Cover to be easy, not Max-E3Lin to be hard.

So now we relax this, and let the dictator functions occasionally fail a constraint. The larger parity functions, however, depend on more bits, so are more likely to accumulate errors and fail additional constraints. We make each bit of each constraint wrong with probability δ, by using a δ-biased distribution: one where each $\lambda_i = 1$ with probability $1 - \delta$ and $\lambda_i = -1$ with probability δ.

Specifically, we use the following algorithm to generate constraints:

- For each constraint, pick $b \in \{0, 1\}$ and $x, y \in \{-1, 1\}^K$ independently and uniformly, and pick $\lambda \in \{-1, 1\}^K$ from the δ-biased distribution. Then take $z = x \circ y \circ \overline{b} \circ \lambda$.

We can again compute the probability that an arbitrary function passes one of these constraints:

$$\begin{aligned}
\mathbf{Pr}[f \text{ passes}] &= \mathbf{E}_{x,y,b,\lambda}[1/2 + 1/2b \cdot f(x) \cdot f(y) \cdot f(z)] \\
&= 1/2 + 1/2\mathbf{E}_{x,y,b,\lambda}[b \cdot f(x) \cdot f(y) \cdot f(z)] \\
&= \frac{1}{2} + \frac{1}{2}\mathbf{E}_{x,y,b,\lambda}\left[b \cdot \sum_{S \in [K]} \hat{f}(S)\chi_S(x) \cdot \sum_{T \in [K]} \hat{f}(T)\chi_T(y) \cdot \sum_{U \in [K]} \hat{f}(U)\chi_U(z)\right] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y,b,\lambda}\left[b\chi_S(x)\chi_T(y)\chi_U(z)\right] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y,b,\lambda}\left[b\chi_S(x)\chi_T(y)\chi_U(x \circ y \circ \overline{b} \circ \lambda)\right] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y,b,\lambda}\left[b\chi_{S \,\triangle\, U}(x)\chi_{T \,\triangle\, U}(y)\chi_U(\overline{b})\chi_U(\lambda)\right] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S,T,U \in [K]} \hat{f}(S)\hat{f}(T)\hat{f}(U)\mathbf{E}_{x,y,b,\lambda}[\chi_{S \,\triangle\, U}(x)]\mathbf{E}_{x,y,b,\lambda}[\chi_{T \,\triangle\, U}(y)] \cdot \\
&\qquad\qquad \mathbf{E}_{x,y,b,\lambda}[b\chi_U(\overline{b})]\mathbf{E}_{x,y,b,\lambda}[\chi_U(\lambda)] \\
&= \frac{1}{2} + \frac{1}{2} \sum_{S \in [K],\, |S| \text{ odd}} \hat{f}(S)^3(1 - 2\delta)^{|S|}
\end{aligned}$$

Note that the last step computed $\mathbf{E}[\chi_U(\lambda)]$ as follows:

$$\mathbf{E}_{x,y,b,\lambda}[\chi_U(\lambda)] = \mathbf{E}_\lambda[\chi_U(\lambda)] = \mathbf{E}[\prod_{i \in U} \lambda_i] = \prod_{i \in U} \mathbf{E}[\lambda_i] = \prod_{i \in U}(1(1 - \delta) + (-1)(\delta)) = (1 - 2\delta)^{|U|}$$

This is now a much more useful test: A dictator passes with probability $1/2 + 1/2(1-2\delta) = 1-\delta$, which is close to 1, while a larger parity function on t variables will be "penalized" by an exponential factor of $(1-2\delta)^t$, so will pass with probability close to $1/2$, and thus will not need to be decoded. Smaller parities may still pass with high probability, but these will still be small enough to be decoded.

3 Soundness

To show soundness, we show that we can choose a small number of Fourier coefficients of small sets, and decode these small sets into labels.

3.1 Smallness

Using the fact that $\hat{f}(S)^3(1-2\delta)^{|S|} = \hat{f}(S)(1-2\delta)^{|S|}\hat{f}(S)^2 \le |\hat{f}(S)|(1-2\delta)^{|S|}\hat{f}(S)^2$, we can bound the probability of a function passing the constraints above:

$$\begin{aligned}
\Pr[f \text{ passes}] &= \frac{1}{2} + \frac{1}{2}\sum_{|S|\,\mathrm{odd}} \hat{f}(S)^3(1-2\delta)^{|S|} \\
&\le \frac{1}{2} + \frac{1}{2}\sum_{|S|\,\mathrm{odd}} |\hat{f}(S)|(1-2\delta)^{|S|}\hat{f}(S)^2 \\
&\le \frac{1}{2} + \frac{1}{2}\max_{|T|\,\mathrm{odd}} |\hat{f}(T)|(1-2\delta)^{|T|}\sum_{|S|\,\mathrm{odd}} \hat{f}(S)^2 \\
&\le \frac{1}{2} + \frac{1}{2}\max_{|T|\,\mathrm{odd}} |\hat{f}(T)|(1-2\delta)^{|T|}
\end{aligned}$$

(The last line follows from Parseval's theorem.)

We only need to be able to decode functions which pass with probability at least $1/2+\epsilon$, which requires that $\max_{|T|\,\mathrm{odd}} |\hat{f}(S)|(1-2\delta)^{|T|} \ge 2\epsilon$. So any function must have some nonempty T^* such that $|\hat{f}(T^*)| \ge 2\epsilon$ and $|T^*| \le \frac{\ln(1/2\epsilon)}{2\delta}$.

This T^* exists since only nonempty sets are in the maximum, and $(1-2\delta)^{|T|} \le 1$, so we must have $|\hat{f}(S)| \ge 2\epsilon$. Finally, $|T^*| \le \frac{\ln(1/2\epsilon)}{2\delta}$, since $|\hat{f}(T^*)| \le 1$, so we must have $(1-2\delta)^{|T^*|} \ge 2\epsilon$.

We can thus decode T^* reasonably, since its size depends only on δ and ϵ, and not on K. Further, since $|\hat{f}(T^*)| \ge 2\epsilon$, $\hat{f}(T^*)^2 \ge 4\epsilon^2$, so there can be at most $1/(4\epsilon^2)$ such T^*, so we can take the union of all such sets, and still be able to decode them.

3.2 Decoding

Definition 3.1 (Suggestions)**.** Let $\mathcal{T} = \{T \subseteq [K] \mid |T| \le \frac{\ln(1/(2\epsilon))}{2\delta}, \hat{f}(T)^2 \ge 4\epsilon^2\}$. Then $\mathrm{Sugg}(f) = \bigcup \mathcal{T}$.

As described above, there are at most $1/4\epsilon^2$ subsets $T \subseteq [K]$ which satisfy $\hat{f}(T)^2 \ge 4\epsilon^2$, and each such set in $\mathrm{Sugg}(f)$ has size at most $\frac{\ln(1/(2\epsilon))}{2\delta}$, so $|\mathrm{Sugg}(f)| \le \frac{\ln(1/(2\epsilon))}{8\delta\epsilon^2}$

Theorem 3.2. *Let $f : \{-1,1\}^K \to \{-1,1\}$ be given. Pick $x, y \in \{-1,1\}^K$ and $b \in \{-1,1\}$ uniformly, and $\lambda \in \{-1,1\}^K$ δ-biased. Set $z = b \cdot (x \circ y \circ \lambda)$. Then if we test whether $f(x) \cdot f(y) \cdot f(z) = b$, if f is a dictator, it will pass with probability $1 - \delta$, and if f passed with probability at least $1/2 + \epsilon$, $\mathrm{Sugg}(f) \neq \emptyset$.*

Proof. The proof of dictators passing comes from section (2.3), and we just showed that if f passed with probability at least $1/2 + \epsilon$, then $\mathrm{Sugg}(f) \neq \emptyset$. □

Thus, we have successfully found a way to decode a solution to Max-E3Lin.

4 Remaining Work

We have a scheme which creates constraints on 2^K elements such that dictator sets are well-satisfied, while any set satisfied significantly more than half the time can be decoded into a set of suggestions.

Unfortunately, the connection to Label-Cover is not obvious. There is no relation to the constraints of the Label-Cover problem, or even that there are two sets of vertices, one for the label and one for the key.

Next lecture, this will be extended roughly so that each edge (u, v) has a constraint $f_u(x)f_v(y)f_v(z) = b$ for some f_u, f_v, and b. This will be chosen such that if both functions are dictators, then the constraint is satisfied with probability at least $1/\delta$, and if we choose functions such that at least $1/2 + 1/2\epsilon$ constraints pass in the Max-E3Lin instance, then $\mathrm{Sugg}(f_u) \neq \emptyset$ and $\mathrm{Sugg}(f_v) \neq \emptyset$, and they have a compatible suggestion.