Inapproximability of MAX-3XOR

The high-level goal of this lecture is to see how *PCP ideas* translate into *hardness of approximation* for a concrete Boolean CSP, namely MAX-3XOR. A MAX-3XOR instance consists of XOR-constraints on triples of Boolean variables; a random assignment satisfies exactly half the constraints in expectation, so $1/2$ is the "trivial" baseline. The main message is that, assuming P $\neq$ NP, one cannot do substantially better than this baseline: it is NP-hard to distinguish instances that are almost satisfiable from instances where no assignment satisfies more than $1/2 + \varepsilon$ of the constraints.

The construction follows the standard PCP pipeline: we start from *Label Cover* (projection games), which serve as the "universal" starting point for many inapproximability reductions, and then encode labels using the *Long Code*. The verifier's checks are designed to be *linear* (XOR) constraints, so that the resulting PCP verifier can be viewed directly as a MAX-3XOR instance.

On the analysis side, the key technical tool is Fourier analysis on the Boolean cube. Noise operators allow us to reason about "low-level" Fourier mass, and Håstad's observation is that a fully fledged dictatorship test is not always necessary: it suffices to show that if the verifier accepts noticeably more than $1/2$, then one can *decode* a nontrivial Label Cover assignment from the Fourier coefficients.

10.1 Label Cover as the starting point

A convenient starting point for many inapproximability reductions is the following projection CSP.

Definition 10.1 (Label Cover (projection game)). A *Label Cover* instance is a tuple

$$G = (L, R, E, \Sigma_L, \Sigma_R, \{\pi_e\}_{e\in E}, \{S_v \subseteq \Sigma_R\}_{v\in R}),$$

where (L, R, E) is a bipartite graph, Σ_L, Σ_R are finite alphabets, and every edge $e = (u, v) \in E$ has an associated *projection map* $\pi_e : \Sigma_R \to \Sigma_L$. A *labeling* is a pair of maps

$$\ell_L : L \to \Sigma_L, \qquad \ell_R : R \to \Sigma_R.$$

An edge $e = (u, v)$ is *satisfied* by (ℓ_L, ℓ_R) if

$$\ell_L(u) \;=\; \pi_e(\ell_R(v)) \wedge \ell_r(v) \in S_v.$$

The *value* of the instance is

$$\mathrm{val}(G) \;:=\; \max_{(\ell_L,\ell_R)} \Pr_{e\sim E}[e \text{ is satisfied}].$$

In the special case appearing repeatedly in the notes, the right alphabet is a q-tuple over the left alphabet, $\Sigma_R = \Sigma_L^q$, and every projection π_e simply selects one coordinate of the tuple. Concretely, if $\ell_R(v) = (b_1, \ldots, b_q)$, then $\pi_e(\ell_R(v)) = b_{i(e)}$ for a fixed index $i(e) \in [q]$.

10.2 Gap hardness for Label Cover

FACT 10.2 (Gap Label Cover hardness (Raz; parallel repetition)). *For every $\delta > 0$, it is NP-hard to distinguish between the following two cases for a Label Cover instance G:*

$$\mathrm{val}(G) = 1 \qquad \textit{vs.} \qquad \mathrm{val}(G) \le \delta.$$

Equivalently, $(1, \delta)$-LC *is NP-hard for arbitrarily small δ.*

REMARK 10.3. This is a deep theorem; the standard route to arbitrarily small δ uses Raz's parallel repetition theorem for projection games.

A useful intuition for why Label Cover should have a constant gap is to reduce from a constant-query PCP. Suppose we have a q-CSP instance Ψ on Boolean variables with the promise that it is either fully satisfiable or at most an s fraction of constraints can be satisfied. Construct a Label Cover instance as follows:

- Left vertices correspond to variables; right vertices correspond to constraints.
- $\Sigma_L = \{0, 1\}$ labels a variable by its Boolean value.
- A constraint vertex v touches q variables; its alphabet $\Sigma_R(v) \subseteq \{0, 1\}^q$ is the set of *satisfying local assignments* to those q variables.
- The projection on an edge (u, v) reads the coordinate of the q-tuple corresponding to u.

If Ψ is satisfiable then all edges can be satisfied (completeness 1). If only an s fraction of constraints of Ψ are satisfiable, then at an unsatisfied constraint vertex v no label in $\Sigma_R(v)$ can agree with *all* of its neighbors, so at most $q-1$ of the q incident edges can be satisfied. Hence

$$\mathrm{val}(G) \;\le\; s \cdot 1 + (1-s) \cdot \frac{q-1}{q} \;=\; 1 - \frac{1-s}{q}.$$

This shows how a (constant) PCP gap yields a (constant) Label Cover gap; parallel repetition amplifies it.

10.3 The Long Code

To turn Label Cover into a low-query PCP with linear constraints, we encode each label by its Long Code.

Definition 10.4 (Long Code (as a function)). Fix an alphabet $\Sigma = \{1, 2, \ldots, n\}$. The *Long Code* encodes a label $i \in \Sigma$ as the Boolean function

$$\mathrm{LC}(i) : \{0,1\}^n \to \{0,1\}, \qquad \mathrm{LC}(i)(x) \ = \ x_i.$$

Equivalently, the codeword is the truth table of the dictator function $x \mapsto x_i$, which has length 2^n.

It is often cleaner to work in $\{-1, 1\}$ notation: given $f : \{0,1\}^n \to \{0,1\}$, define the associated $\{-1,1\}$-valued function $F : \{0,1\}^n \to \{-1,1\}$ by $F(x) = (-1)^{f(x)}$. Then dictators become *characters*:

$$(-1)^{x_i} \ = \ \chi_{\{i\}}(x).$$

As a result, dictators have extremely simple Fourier expansions (supported on a single subset).

10.4 A "PCP" for Label Cover with linear constraints

Let $G = (L, R, E, \Sigma_L, \Sigma_R, \{\pi_e\})$ be a Label Cover instance. The prover is supposed to write down the Long Code encoding of a labeling:

- For each $u \in L$, an oracle/function $f_u : \{0,1\}^{\Sigma_L} \to \{0,1\}$.
- For each $v \in R$, an oracle/function $g_v : \{0,1\}^{\Sigma_R} \to \{0,1\}$.

(Here $\{0,1\}^{\Sigma}$ means strings indexed by Σ, i.e. functions $\Sigma \to \{0,1\}$.)

The intended proof is that f_u is the dictator for the left label $\ell_L(u)$ and g_v is the dictator for the right label $\ell_R(v)$.

A key feature is that the verifier will use only *linear* checks (XOR constraints), so the verifier itself can be viewed as producing an instance of MAX-3XOR whose variables are the queried proof bits.

Definition 10.5 (Extension map along an edge). Fix an edge $e = (u, v) \in E$ with projection $\pi_e : \Sigma_R \to \Sigma_L$. Given $x \in \{0,1\}^{\Sigma_L}$, define its *extension* $z = \mathrm{Ext}_e(x) \in \{0,1\}^{\Sigma_R}$ by

$$z_b \ := \ x_{\pi_e(b)} \qquad \text{for each } b \in \Sigma_R.$$

If g_v is a dictator at $b^\star \in \Sigma_R$, then $g_v(\mathrm{Ext}_e(x)) = x_{\pi_e(b^\star)}$, which should match the dictator $f_u(x)$ if $\pi_e(b^\star)$ is the correct left label.

Definition 10.6 (Håstad's noisy 3-query projection test). Fix a noise parameter $\rho \in (0, 1)$. The verifier performs:

1. Sample a random edge $e = (u, v) \in E$.

2. Sample $x \in \{0,1\}^{\Sigma_L}$ and $y \in \{0,1\}^{\Sigma_R}$ uniformly at random. Let $z = \mathrm{Ext}_e(x) \in \{0,1\}^{\Sigma_R}$.
3. Sample $y' \sim N_\rho(y)$, where N_ρ is the standard ρ-noise operator (defined formally in the next section).
4. Query the three bits $f_u(x)$, $g_v(y')$, and $g_v(y+z)$ and *accept* iff

$$f_u(x) \;=\; g_v(y') \oplus g_v(y+z). \tag{10.1}$$

Equation (10.1) is a 3-variable XOR constraint. Thus, if we create one Boolean variable for each possible query $f_u(x)$ and $g_v(y)$, and add one constraint for each random choice made by the verifier, we obtain a MAX-3XOR instance.

10.5 Fourier analysis and the noise operator

We briefly collect the Fourier facts used in the soundness analysis.

Definition 10.7 (Characters and Fourier coefficients). For $S \subseteq [n]$ define the character

$$\chi_S(x) \;:=\; (-1)^{\sum_{i\in S} x_i}, \qquad x \in \{0,1\}^n.$$

Every function $F : \{0,1\}^n \to \mathbb{R}$ has a unique Fourier expansion

$$F(x) \;=\; \sum_{S\subseteq[n]} \widehat{F}(S)\,\chi_S(x), \qquad \widehat{F}(S) \;:=\; \mathbb{E}_x[F(x)\chi_S(x)].$$

If $F : \{0,1\}^n \to \{-1,1\}$, then Parseval gives $\sum_S \widehat{F}(S)^2 = 1$.

Definition 10.8 (Noise operator). For $\rho \in [-1,1]$, the noise distribution $N_\rho(x)$ is obtained by independently, for each coordinate, setting

$$y_i \;=\; \begin{cases} x_i & \text{with probability } \frac{1+\rho}{2}, \\ 1-x_i & \text{with probability } \frac{1-\rho}{2}. \end{cases}$$

The corresponding noise operator is

$$(T_\rho F)(x) \;:=\; \mathbb{E}_{y\sim N_\rho(x)}[F(y)].$$

Claim 10.9 (Noise scales Fourier levels). *For every $S \subseteq [n]$,*

$$\widehat{T_\rho F}(S) \;=\; \rho^{|S|}\,\widehat{F}(S).$$

Proof sketch. Check on $F = \chi_S$:

$$(T_\rho\chi_S)(x) = \mathbb{E}_{y\sim N_\rho(x)}\left[(-1)^{\sum_{i\in S} y_i}\right] = \prod_{i\in S}\mathbb{E}[(-1)^{y_i}] = \prod_{i\in S}\rho(-1)^{x_i} = \rho^{|S|}\chi_S(x).$$

Linearity gives the general case. ∎

Corollary 10.10 (Noise stability identity). *If $F : \{0,1\}^n \to \{-1,1\}$, then*

$$\mathrm{Stab}_\rho(F) \;:=\; \mathbb{E}_{x,\,y\sim N_\rho(x)}[F(x)F(y)] \;=\; \sum_{S\subseteq[n]} \rho^{|S|}\,\widehat{F}(S)^2.$$

10.6 Reduction to MAX-3XOR and completeness

Using the noisy projection test, we can reduce Label Cover to MAX-3XOR.

THEOREM 10.11 (Informal reduction statement). *Fix $\varepsilon > 0$ and choose a corresponding noise parameter $\rho = 1 - \Theta(\varepsilon)$. There is a polynomial-time reduction that maps a Label Cover instance G to a* MAX-3XOR *instance Φ such that:*

1. *(Completeness) If $\mathrm{val}(G) = 1$ then $\mathrm{OPT}(\Phi) \geq 1 - \varepsilon$.*
2. *(Soundness) If $\mathrm{OPT}(\Phi) > 1/2 + \varepsilon$ then $\mathrm{val}(G) \geq \Omega\big(\varepsilon^5/\log^2(1/\varepsilon)\big)$.*

Consequently, if $\mathrm{val}(G) \leq \delta$ for $\delta \ll \varepsilon^5/\log^2(1/\varepsilon)$, then $\mathrm{OPT}(\Phi) \leq 1/2 + \varepsilon$.

Completeness sketch. Assume $\mathrm{val}(G) = 1$ and fix a satisfying labeling (ℓ_L, ℓ_R). For each $u \in L$ set f_u to be the dictator for $\ell_L(u)$, and for each $v \in R$ set g_v to be the dictator for $\ell_R(v)$.

Fix an edge $e = (u, v)$ and let $b^\star = \ell_R(v) \in \Sigma_R$, $a^\star = \ell_L(u) = \pi_e(b^\star)$. Under the test, the left query returns $f_u(x) = x_{a^\star}$. The two right queries return $g_v(y') = y'_{b^\star}$ and $g_v(y + z) = y_{b^\star} \oplus z_{b^\star}$. Since $z_{b^\star} = x_{\pi_e(b^\star)} = x_{a^\star}$, the check $f_u(x) = g_v(y') \oplus g_v(y + z)$ becomes $x_{a^\star} = y'_{b^\star} \oplus y_{b^\star} \oplus x_{a^\star}$, which simplifies to $y'_{b^\star} = y_{b^\star}$. This holds with probability $(1 + \rho)/2 = 1 - O(\varepsilon)$. Hence completeness $1 - \varepsilon$ follows. ∎

The remainder of the lecture is devoted to the soundness direction in Theorem 10.11, following the Fourier-analytic argument sketched in the handwritten notes.

10.7 Soundness: decoding a Label Cover assignment

We prove the key contrapositive statement used for soundness.

CLAIM 10.12. *If the verifier based on Håstad's test accepts with probability greater than $1/2 + \varepsilon$, then the underlying Label Cover instance G satisfies*

$$\mathrm{val}(G) \;\geq\; \Omega\!\left(\frac{\varepsilon^5}{\log^2(1/\varepsilon)}\right).$$

Proof sketch (expanded from the notes). Call an edge $e = (u, v)$ *good* if, restricted to that edge, the test accepts with probability at least $1/2 + \varepsilon/2$. Let $p = \Pr_{e \sim E}[e \text{ is good}]$. Averaging shows $p \geq \varepsilon$: even if every non-good edge had acceptance probability as large as $1/2 + \varepsilon/2$, we would have

$$p \cdot 1 + (1 - p)\left(\frac{1}{2} + \frac{\varepsilon}{2}\right) \;\geq\; \frac{1}{2} + \varepsilon \qquad \Longrightarrow \qquad p \geq \varepsilon.$$

Fix a good edge $e = (u, v)$ and abbreviate $f = f_u$ and $g = g_v$. Switch to $\{-1, 1\}$ notation:

$$F(x) \;:=\; (-1)^{f(x)}, \qquad G(y) \;:=\; (-1)^{g(y)}.$$

The XOR check $f(x) = g(y') \oplus g(y+z)$ is equivalent to the product condition $F(x)\, G(y')\, G(y+z) = 1$. Thus acceptance probability $1/2 + \varepsilon/2$ implies the correlation bound

$$\varepsilon \;\leq\; \mathbb{E}\big[F(x)\, G(y')\, G(y + z)\big], \tag{10.2}$$

where the expectation is over the test's random choices on edge e.

Step 1: Fourier expansion and the projection map on sets. For a fixed shift z, define

$$H(z) \;:=\; \mathbb{E}_{y,\, y' \sim N_\rho(y)}\big[G(y')\, G(y+z)\big].$$

A standard Fourier calculation (autocorrelation + noise) gives

$$H(z) \;=\; \sum_{S \subseteq \Sigma_R} \rho^{|S|}\, \widehat{G}(S)^2\, \chi_S(z). \tag{10.3}$$

(Compare this to the stability identity; here we correlate a noisy copy of G with a shifted copy.)

Next, recall that in the test we always set $z = \mathrm{Ext}_e(x)$ for the extension map along e. For a set $S \subseteq \Sigma_R$, define its *projection* $\mathrm{proj}_e(S) \subseteq \Sigma_L$ by

$$\mathrm{proj}_e(S) \;:=\; \{a \in \Sigma_L : \big|\{b \in S : \pi_e(b) = a\}\big| \text{ is odd}\}.$$

Equivalently, $\mathrm{proj}_e(S)$ records which left labels appear an odd number of times when projecting the right labels in S. One verifies the character identity

$$\chi_S(\mathrm{Ext}_e(x)) \;=\; \chi_{\mathrm{proj}_e(S)}(x). \tag{10.4}$$

Plugging (10.3) and (10.4) into (10.2) yields

$$\begin{aligned}
\varepsilon &\le \mathbb{E}_x\big[F(x)\, H(\mathrm{Ext}_e(x))\big] \\
&= \mathbb{E}_x\left[F(x) \sum_{S \subseteq \Sigma_R} \rho^{|S|} \widehat{G}(S)^2 \chi_{\mathrm{proj}_e(S)}(x)\right] \\
&= \sum_{S \subseteq \Sigma_R} \rho^{|S|} \widehat{G}(S)^2\, \widehat{F}(\mathrm{proj}_e(S)).
\end{aligned} \tag{10.5}$$

Step 2: Cauchy–Schwarz and truncating to small Fourier sets. Apply Cauchy–Schwarz to (10.5):

$$\varepsilon \le \left(\sum_S \widehat{G}(S)^2\right)^{1/2} \left(\sum_S \rho^{2|S|} \widehat{F}(\mathrm{proj}_e(S))^2\, \widehat{G}(S)^2\right)^{1/2}.$$

Since $\sum_S \widehat{G}(S)^2 = 1$ (Parseval for $\{-1,1\}$-valued G), we get

$$\varepsilon^2 \;\le\; \sum_{S \subseteq \Sigma_R} \rho^{2|S|}\, \widehat{F}(\mathrm{proj}_e(S))^2\, \widehat{G}(S)^2. \tag{10.6}$$

Now choose a cutoff

$$\theta \;:=\; \frac{10}{\varepsilon} \log \frac{1}{\varepsilon},$$

so that $\rho^{2\theta} \le 0.1\, \varepsilon^2$ (for $\rho = 1 - \Theta(\varepsilon)$, this is a standard calculus estimate). Split the RHS of (10.6) into the contributions of $|S| < \theta$ and $|S| \ge \theta$. For the large sets we use $\rho^{2|S|} \le \rho^{2\theta}$ and Parseval:

$$\sum_{|S| \ge \theta} \rho^{2|S|} \widehat{F}(\mathrm{proj}_e(S))^2\, \widehat{G}(S)^2 \le \rho^{2\theta} \sum_S \widehat{G}(S)^2 \le 0.1\, \varepsilon^2.$$

Therefore the small sets contribute at least $0.9\varepsilon^2$, and since $\rho^{2|S|} \le 1$,

$$0.9\,\varepsilon^2 \;\le\; \sum_{\substack{S \subseteq \Sigma_R \\ |S| < \theta}} \widehat{F}(\mathrm{proj}_e(S))^2\, \widehat{G}(S)^2. \tag{10.7}$$

Step 3: A randomized decoding rule. For each left vertex u, define a distribution on subsets $T \subseteq \Sigma_L$ by

$$\mathbb{P}[T] \;=\; \widehat{F}_u(T)^2,$$

and then decode a label by sampling T from this distribution and outputting a uniformly random $a \in T$.

Similarly, for each right vertex v, sample a set $S \subseteq \Sigma_R$ with probability $\widehat{G}_v(S)^2$ and output a uniformly random $b \in S$.

Step 4: Satisfying a good edge with nontrivial probability. Fix a good edge $e = (u, v)$ and apply the above decoding independently to u and v. By (10.7), the contribution to the RHS from pairs (T, S) with $T = \mathrm{proj}_e(S)$ and $|S| < \theta$ is at least $0.9\varepsilon^2$. Since the decoding samples T with probability $\widehat{F}_u(T)^2$ and S with probability $\widehat{G}_v(S)^2$, the probability of drawing such a pair satisfies

$$\mathbb{P}[T = \mathrm{proj}_e(S),\ |S| < \theta] \;\ge\; 0.9\varepsilon^2.$$

Condition on the event that $T = \mathrm{proj}_e(S)$ and $|S| < \theta$. At this point there is a small gap in the analysis: S could be empty (or could yield $T = \emptyset$), making the "pick a random element" rule ill-defined. Håstad addresses this using the *folding trick*, which enforces $\widehat{G}(S) = 0$ for all even $|S|$, in particular $\widehat{G}(\emptyset) = 0$. Thus we may assume the sampled S is *odd* and nonempty.

Since S is odd, there exists some $b^\star \in S$ such that $\pi_e(b^\star) \in T$ (since $T = \mathrm{proj}_e(S)$ consists exactly of the projected labels that occur an odd number of times). Pick b uniformly from S and a uniformly from T. Then

$$\mathbb{P}[a = \pi_e(b) \mid T = \mathrm{proj}_e(S),\, |S| < \theta] \;\ge\; \mathbb{P}[b = b^\star] \cdot \mathbb{P}[a = \pi_e(b^\star)] \;\ge\; \frac{1}{|S|} \cdot \frac{1}{|T|} \;\ge\; \frac{1}{\theta^2},$$

since $|T| \le |S| < \theta$.

For a good edge e we satisfy e with probability at least $0.9\varepsilon^2 \cdot (1/\theta^2)$ under the decoding, and a random edge is good with probability at least ε. Therefore

$$\mathbb{E}[\text{fraction of satisfied edges}] \;\ge\; \varepsilon \cdot 0.9\varepsilon^2 \cdot \frac{1}{\theta^2} \;=\; \Omega\!\left(\frac{\varepsilon^5}{\log^2(1/\varepsilon)}\right).$$

By averaging, there exists a deterministic labeling achieving at least this value, i.e. $\mathrm{val}(G)$ is at least the same quantity. This proves the claim. ∎

10.8 Remark: folding and "odd" Fourier support

The notes highlight an important technical cleanup:

- It can happen that the sampled Fourier set S is empty.

- More generally, one would like to rule out even-sized S so that the projection $T = \mathrm{proj}_e(S)$ is nonempty.

Håstad's *folding* trick enforces an "oddness" symmetry on the proof tables, e.g. in $\{0, 1\}$ notation one can enforce a constraint of the form

$$G(x) \;=\; -G(\bar{x}), \qquad \bar{x} := 1 - x,$$

which implies that $\widehat{G}(S) = 0$ for all even $|S|$. Equivalently, defining $G^{\mathrm{odd}}(x) = G(x) - G(\bar{x})$ removes all even Fourier components:

$$G^{\mathrm{odd}}(x) \;=\; \sum_{S:\,|S| \text{ odd}} \widehat{G}(S)\,\chi_S(x).$$

This odd-support property is what is used in the decoding step to guarantee that T is nonempty and to lower bound $\mathbb{P}[a = \pi_e(b)]$ by $1/\theta^2$.

There is a "second issue" in the full proof; in a complete treatment one tracks several additional technicalities (e.g. how folding interacts with the encoding and how to ensure all sampled distributions are well-defined). The core Fourier-analytic decoding argument is captured above.