Approximation Algorithms **Spring 2019**

Lecture 12: Optimality of Goemans-Williamson under the UGC

Scribe: Wilson Wu May 22, 2019

12.1 Introduction

It turns out sometimes doing hard things approximately is still hard. Perhaps this is appropriate, but we would like to console ourselves with theory about this hardness. Here we look at how the Unique Games Conjecture implies that the Goemans-Williamson approximation for MaxCut is optimal among polynomial-time algorithms. That is, assuming this conjecture, it is NP-hard to approximate MaxCut to any constant better than

$$\alpha_{\mathrm{GW}} = \min_{0 \leq \theta \leq \pi} \frac{\theta/\pi}{2(1-\cos\theta)} \approx 0.87856$$

in expectation. Along the way we will see ideas from social choice theory and the study of probabilistically checkable proofs.

12.2 The Goemans-Williamson Algorithm

The MaxCut problem is that of finding the maximum cut of a graph $G = (V, E)$. That is, the desired output is a cut $S \subseteq V$ maximizing the number of "cut" edges $|\{(i,j) \mid i \in S \wedge j \in V \setminus S\}|$. Recall that the Goemans-Williamson algorithm relaxes the exact integer programming formulation of this problem

$$\begin{aligned} \max_{(x_i)_{i \in V} \in \mathbb{R}} \quad & \frac{1}{2} \sum_{(i,j) \in E} (1 - x_i x_j) \\ \text{s.t.} \quad & \forall i \in V : x_i \in \{-1, 1\} \end{aligned}$$

to the following vector program:

$$\begin{aligned} \max_{(\mathbf{u}_i)_{i \in V} \in \mathbb{R}^{|V|}} \quad & \frac{1}{2} \sum_{(i,j) \in E} (1 - \mathbf{u}_i^\top \mathbf{u}_j) \\ \text{s.t.} \quad & \forall i \in V : \mathbf{u}_i^\top \mathbf{u}_i = 1. \end{aligned}$$

That is, we embed the graph in a sphere in $\mathbb{R}^n$ and maximize

$$\frac{1}{2} \sum_{(i,j) \in E} (1 - \mathbf{u}_i^\top \mathbf{u}_j) = \frac{1}{4} \sum_{(i,j) \in E} \|\mathbf{u}_i - \mathbf{u}_j\|^2$$

the summed pairwise ℓ_2 distance between vertices. Vector programs are equivalent to semidefinite programs, and can be solved in polynomial time. Recall also that a solution $(\mathbf{u}_i)_{i \in V}$ is rounded to a cut using a random hyperplane — that is, $\mathbf{g} \in \mathbb{R}^n$ is sampled from an uncorrelated n-dimensional Gaussian distribution and we create the cut S such that $i \in S$ if and only if $\mathbf{u}_i^\top g \geq 0$. Then, the probability of each edge $(i,j) \in E$ being cut is $\theta_{i,j}/\pi$, where $\theta_{i,j}$ is the angle between $\mathbf{u}_i$ and $\mathbf{u}_j$, and the contribution of each edge to the objective is

$$\frac{1}{2} \sum_{(i,j) \in E} (1 - \cos\theta_{i,j}).$$

It follows that this rounding procedure approximates any G with ratio at least

$$\alpha_{\mathrm{GW}} = \min_{0 \leq \theta \leq \pi} \frac{\theta/\pi}{1/2(1-\cos\theta)}.$$

That is, the cut found by the Goemans-Williamson algorithm is at worst α_{GW}OPT, where OPT is the size of the true maximum cut.

12.3 Majority is Stablest

Before we discuss the optimality of the α_{GW}-approximation of MAXCUT we will introduce some concepts from social choice theory and the analysis of Boolean functions.

Definition 12.1. *Given a function $f : \{-1,1\}^n \longrightarrow \mathbb{R}$, the* influence *of the coordinate i on f is*

$$\mathrm{Inf}_i(f) = \mathbb{E}_{x_j : j \neq i}[\mathrm{var}_{x_i}(f(\mathbf{x}))]$$

where $\mathbf{x}$ is sampled uniformly over $\{-1,1\}^n$.

Definition 12.2. *Given $\mathbf{x} \in \{-1,1\}^n$, and $\rho \in \mathbb{R}$ such that $-1 \leq \rho \leq 1$, a ρ*-correlated copy *of $\mathbf{x}$ is a string $\mathbf{y} \in \{-1,1\}^n$ sampled as*

$$y_i = \begin{cases} x_i & w.p.\ 1/2(1+\rho) \\ -x_i & w.p.\ 1/2(1-\rho) \end{cases}$$

Note then that $\mathbf{x}$ and $\mathbf{y}$ are indeed ρ-correlated, since if $\mathbf{x}$ is sampled uniformly then $\forall i : \mathrm{cov}(x_i, y_i) = \mathbb{E}[x_i y_i] = \rho$.

Definition 12.3. *Given $f : \{-1,1\}^n \longrightarrow \mathbb{R}$ and $\rho \in \mathbb{R}$ such that $-1 \leq \rho \leq 1$, the* noise stability *of f at ρ is*

$$\mathrm{St}_\rho(f) = \mathbb{E}_{\mathbf{x},\mathbf{y}}[f(\mathbf{x})f(\mathbf{y})]$$

for $\mathbf{x}$ sampled uniformly from $\{-1,1\}^n$ and $\mathbf{y}$ a ρ-correlated copy of $\mathbf{x}$.

Intuitively, the influence of coordinate i on a function f is how sensitive f is to a perturbation at i. Meanwhile, the stability of f is its robustness to perturbations anywhere in its input.

In voting theory, we generally consider $f : \{-1,1\}^n \longrightarrow \{-1,1\}$, interpreted as a voting rule — each of n people vote for one of the two candidates -1 and 1, and the winner is determined by f. In this case, it follows readily from the definitions that

$$\mathrm{Inf}_i(f) = \Pr_{\mathbf{x}}[f(x_1, \ldots, x_n) \neq f(x_1, \ldots, -x_i, \ldots, x_n)]$$

and

$$\mathrm{St}_\rho(f) = 2 \Pr_{\mathbf{x},\mathbf{y}}(f(x) = f(y)) - 1$$

for uniformly sampled $\mathbf{x} \in \{-1, 1\}^n$ and a ρ-correlated copy $\mathbf{y}$. Typically, we would like to have a rule that is robust to errors in vote counting — that is, we want a function f such that $\mathrm{St}_\rho(f)$ is high for $\rho > 0$. It turns out that *dictator functions*, defined as $d_i(\mathbf{x}) = x_i$ are very stable:

$$\begin{aligned}\mathrm{St}_\rho(d_i) &= 2 \Pr_{\mathbf{x},\mathbf{y}}(f(x) = f(y)) - 1 \\ &= 2 \Pr_{\mathbf{x},\mathbf{y}}(x_i = y_i) - 1 \\ &= \rho.\end{aligned}$$

This is intuitively rather obvious, as in dictatorships the accuracy of vote-counting systems should not matter too much. However, dictatorships nowadays are somewhat rare — people tend to prefer voting rules where no one individual has too much influence. Consider the majority voting rule:

$$\mathrm{Maj}(\mathbf{x}) = \mathrm{sgn} \sum_{i=1}^{n} x_i.$$

Note that under the majority rule, an individual's vote only matters if *exactly* half of the population vote for one candidate. Thus, using Stirling's formula,

$$\forall i : \mathrm{Inf}_i(\mathrm{Maj}) \approx \binom{n}{n/2} \frac{1}{2^n} \approx \sqrt{\frac{2}{\pi n}}$$

which tends to zero as n approaches infinity. Moreover, the stability is still pretty good:

$$\begin{aligned}\mathrm{St}_\rho(\mathrm{Maj}) &= 1 - 2 \Pr_{\mathbf{x},\mathbf{y}}(\mathrm{Maj}(x) \neq \mathrm{Maj}(y)) \\ &= 1 - 2 \Pr_{\mathbf{x},\mathbf{y}} \left(\mathrm{sgn} \frac{\sum_{i=1}^n x_i}{\sqrt{n}} \neq \mathrm{sgn} \frac{\sum_{i=1}^n y_i}{\sqrt{n}} \right)\end{aligned}$$

with $\mathbf{x}$ and $\mathbf{y}$ being ρ-correlated. By the Central Limit Theorem, this is equivalent to two ρ-correlated zero-mean unit Gaussian random variables z_x and z_y having the same sign. We can construct such random variables as $z_x = \mathbf{g}^\top \mathbf{a}$ and $z_y = \mathbf{g}^\top \mathbf{b}$ for some unit vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^k$ (choose an arbitrary $k \geq 2$) such that $\mathbf{a}^\top \mathbf{b} = \rho$, and $\mathbf{g}$ sampled from an uncorrelated multivariate Gaussian. Indeed, we then have

$$\begin{aligned}\mathbb{E}[z_x] &= \sum_{i=1}^{k} a_i \mathbb{E}[g_i] = 0 \\ \mathrm{var}(z_x) &= \sum_{i=1}^{k} a_i^2 \mathbb{E}[g_i^2] = \|a_i\|_2 = 1 \\ \mathrm{cov}(z_x, z_y) &= \sum_{i=1}^{k} \sum_{j=1}^{k} a_i b_j \,\mathrm{cov}(g_i, g_j) = \sum_{i=1}^{k} a_i b_j = \rho.\end{aligned}$$

Then,

$$\Pr(\mathrm{sgn}\, z_x \neq \mathrm{sgn} -z_y) = \Pr(\mathrm{sgn}\, \mathbf{g}^\top \mathbf{a} \neq \mathrm{sgn}\, \mathbf{g}^\top \mathbf{b}).$$

Since the uncorrelated multivariate Gaussian is rotationally symmetric, this is equal to θ/π where $\theta = \arccos(\rho)$ is the angle between $\mathbf{a}$ and $\mathbf{b}$. Thus,

$$\mathrm{St}_\rho(\mathrm{Maj}) = 1 - \frac{2}{\pi} \arccos \rho + o(1)$$

with the $o(1)$ error coming from the Central Limit Theorem.

In fact, in a democratic world it is impossible to do much better than majority in terms of stability. Formally,

Theorem 12.4. ***(Majority is Stablest)*** *For any $\rho \in [0, 1)$ and $\epsilon > 0$ there exists a constant δ such that for any $f : \{-1, 1\}^n \longrightarrow [-1, 1]$ with $\mathbb{E}_{\mathbf{x}}[f(\mathbf{x})] = 0$ and $\forall i : \mathrm{Inf}_i(f) \le \delta$ we have*

$$\mathrm{St}_\rho(f) \le 1 - \frac{2}{\pi} \arccos \rho + \epsilon$$

We will not go into the proof here, but it involves the Invariance Principle, which roughly states that multilinear polynomials behave similarly over the uniform distribution on the Boolean hypercube and the Gaussian, and a Gaussian isoperimetric inequality due to Borell. The idea of this result is that among all sets of Gaussian volume $1/2$, the half-space has minimal "surface area". That is, intuitively, Majority is Stablest is closely related to the fact that hyperplanes are a good way to cut ρ-correlated vectors on a hypersphere.

12.3.1 Analysis of Boolean Functions

The statement of Theorem 12.4 above is in a somewhat different form than the one we will end up using. To state it in a more convenient manner, we will use some basic Fourier analysis.

Consider functions $f : \{-1, 1\}^n \longrightarrow \mathbb{R}$ and view the domain $\{-1, 1\}^n$ as a uniform distribution over the Boolean hypercube. These functions form an inner product space

$$\langle f, g \rangle = \mathbb{E}[fg]$$

with $\|f\|_2 = \sqrt{\mathbb{E}[f^2]}$. It is a fact that the parity functions $\chi_S(\mathbf{x}) = \prod_{i \in S} x_i$ form an orthonormal basis of this space. That is, we can rewrite f as

$$f = \sum_{S \subseteq [n]} \hat{f}(S) \chi_S$$

where $\hat{f}(S) = \langle f, \chi_S \rangle$ are referred to as the *Fourier coefficients* of f. *Plancherel's identity* states

$$\langle f, g \rangle = \sum_{S \subseteq [n]} \hat{f}(S) \hat{g}(S).$$

Note then that for $f : \{-1, 1\} \longrightarrow [-1, 1]$ we have

$$\|f\|_2^2 = \sum_{S \subseteq [n]} \hat{f}(S)^2 \le 1$$

with equality if f has range $\{-1, 1\}$.

Consider the following linear operator:

Definition 12.5. *The* Bonami-Beckner operator $T_\rho \in \mathrm{End}(\{-1, 1\}^n \longrightarrow \mathbb{R})$ *is defined as*

$$T_\rho(f)(\mathbf{x}) = \mathbb{E}[f(\mathbf{y})]$$

where $\mathbf{y}$ is a ρ-correlated copy of $\mathbf{x}$.

Since T_ρ is linear, we see that

$$\begin{aligned} T_\rho(f)(\mathbf{x}) &= T_\rho\left(\sum_{S\subseteq[n]} \hat{f}(S)\chi_S(\mathbf{x})\right) \\ &= \sum_{S\subseteq[n]} \hat{f}(S)T_\rho(\chi_S(\mathbf{x})) \\ &= \sum_{S\subseteq[n]} \hat{f}(S)\mathbb{E}[\chi_S(\mathbf{y})]. \end{aligned}$$

Also, since the coordinates of $\mathbf{y}$ are independent,

$$\begin{aligned} \mathbb{E}[\chi_S(\mathbf{y})] &= \mathbb{E}\left[\prod_{i\in S} y_i\right] \\ &= \prod_{i\in S}\mathbb{E}[y_i] \\ &= \prod_{i\in S}\left(\frac{1}{2}(1+\rho)x_i - \frac{1}{2}(1-\rho)x_i\right) \\ &= \prod_{i\in S}\rho x_i \\ &= \rho^{|S|}\chi_S(x_i) \end{aligned}$$

from which it follows

$$T_\rho(f) = \sum_{S\subseteq[n]} \rho^{|S|}\hat{f}(S)\chi_S.$$

Thus,

$$\mathrm{St}_\rho(f) = \langle f, T_\rho f\rangle = \sum_{S\subseteq[n]} \rho^{|S|}\hat{f}(S)^2.$$

Moreover,

$$\begin{aligned} \mathrm{Inf}_i(f) &= \mathbb{E}_{x_j:j\neq i}\left[\mathrm{var}_{x_i}\left(\sum_{S\subseteq[n]} \hat{f}(S)\chi_S(\mathbf{x})\right)\right] \\ &= \mathbb{E}_{x_j:j\neq i}\left[\sum_{S\subseteq[n]} \hat{f}(S)^2\mathbb{1}\{i\in S\}\right] \\ &= \sum_{S\subseteq[n]:i\in S} \hat{f}(S)^2 \end{aligned}$$

We can now look at the contribution to the influence from sets at most a given size:

Definition 12.6. *For $f:\{-1,1\}^n \longrightarrow \mathbb{R}$, the k*-degree influence of coordinate i on f *is*

$$\mathrm{Inf}_i^{\leq k}(f) = \sum_{S\subseteq[n]:i\in S\wedge|S|\leq k} \hat{f}(S)^2$$

The Majority is Stablest theorem has the following alternate forms:

Corollary 12.7. *For any $\rho\in[0,1)$ and $\epsilon>0$ there exist constants δ, k such that for any $f:\{-1,1\}^n \longrightarrow [-1,1]$ with $\mathbb{E}_{\mathbf{x}}[f(\mathbf{x})]=0$ and $\forall i: \mathrm{Inf}_i^{\leq k}(f)\leq\delta$ we have*

$$\mathrm{St}_\rho(f) \leq 1 - \frac{2}{\pi}\arccos\rho + \epsilon$$

Proof. For a given $\rho \in [0,1)$ and $\epsilon > 0$, by Theorem 12.4 there exists $\delta > 0$ such that for any $g : \{-1,1\}^n \longrightarrow [-1,1]$ if $\forall i : \mathrm{Inf}_i(g) \leq \delta$ then $\mathrm{St}_\rho(g) \leq 1 - 2/\pi \arccos \rho + \epsilon/4$. Choose γ such that $\rho^{k'}(1-(1-\gamma)^{2k'}) < \epsilon/4$ for all positive k', and k such that $(1-\gamma)^k < \delta/2$. Let $f : \{-1,1\}^n \longrightarrow [-1,1]$ be a function with $\forall i : \mathrm{Inf}_i^{\leq k} \leq \delta/2$ and $g = T_{1-\gamma} f$. Then, for all i

$$\begin{aligned}\mathrm{Inf}_i(g) &= \sum_{S \subseteq [n] : i \in S} (1-\gamma)^{|S|} \hat{f}(S)^2 \\ &\leq \sum_{S \subseteq [n] : i \in S, |S| \leq k} \hat{f}(S)^2 + (1-\gamma)^k \sum_{S \subseteq [n] : i \in S, |S| > k} \hat{f}(S)^2 \\ &\leq \frac{\delta}{2} + \frac{\delta}{2} = \delta.\end{aligned}$$

Thus $\mathrm{St}_\rho(g) \leq 1 - 2/\pi \arccos \rho + \epsilon/4$ so

$$\mathrm{St}_\rho(f) = \mathrm{St}_\rho(g) + \sum_{S \subseteq [n]} (\rho^{|S|}(1-(1-\gamma)^{|S|}))\hat{f}(S)^2 \leq 1 - \frac{2}{\pi}\arccos \rho + \frac{3\epsilon}{4}$$

□

Corollary 12.8. *For any $\rho \in (-1,0]$ and $\epsilon > 0$ there exists a constant δ such that for any $f : \{-1,1\}^n \longrightarrow [-1,1]$ with $\forall i : \mathrm{Inf}_i(f) \leq \delta$ we have*

$$\mathrm{St}_\rho(f) \geq 1 - \frac{2}{\pi}\arccos \rho - \epsilon$$

Proof. Consider $f : \{-1,1\}^n \longrightarrow [-1,1]$ such that $\forall i : \mathrm{Inf}_i(f) \leq \delta$. Define

$$\begin{aligned}g(x) &= \frac{f(x) - f(-x)}{2} \\ &= \frac{\sum_{S \subseteq [n]} (\hat{f}(S)\chi_S(x) - \hat{f}(S)\chi_S(-x))}{2} \\ &= \sum_{S \subseteq [n] : |S| \text{ odd}} \hat{f}(S)\chi_S\end{aligned}$$

where the last equality follows from noticing

$$\chi_S(-x) = \begin{cases} \chi_S(x) & \text{if } |S| \equiv 0 \mod 2 \\ -\chi_S(x) & \text{otherwise.} \end{cases}$$

Note $\mathbb{E}[g] = 0$ and $\forall i : \mathrm{Inf}_i(g) \leq \mathrm{Inf}_i(f)$ so

$$\mathrm{St}_\rho(f) \geq \mathrm{St}_\rho(g) = -\mathrm{St}_{-\rho}(g) \geq -\left(1 - \frac{2}{\pi}\arccos -\rho + \epsilon\right) = 1 - \frac{2}{\pi}\arccos \rho - \epsilon$$

where the last equality follows from $\arccos(\rho) + \arccos(-\rho) = \pi$. □

Corollary 12.9. *For any $\rho \in (-1,0]$ and $\epsilon > 0$ there exist constants δ, k such that for any $f : \{-1,1\}^n \longrightarrow [-1,1]$ with $\forall i : \mathrm{Inf}_i^{\leq k}(f) \leq \delta$ we have*

$$\mathrm{St}_\rho(f) \geq 1 - \frac{2}{\pi}\arccos \rho - \epsilon$$

Proof. Combination of Corollaries 12.7 and 12.8 □

12.4 PCP and Unique Games

One natural way to think of NP is as the class of problems which can be efficiently verified. That is if $L \in \mathsf{NP}$, then there exists a polynomial-time verifier algorithm V such that for any problem instance x,

- $x \in L \longrightarrow \exists \pi : V(x, \pi) = 1$.
- $x \notin L \longrightarrow \forall \tilde{\pi} : V(x, \tilde{\pi}) = 0$.

The class PCP is similar, except now V has oracle access to the proof π and queries at a small number of random locations to make its decision. Formally,

Definition 12.10. *We say $L \in \mathsf{PCP}_{c,s}[l, r, q]_\Sigma$ if when ρ is sampled uniformly from $\{-1, 1\}^r$*

- $x \in L \longrightarrow \exists \pi : \Pr_\rho(V^\pi(x) = 1) \geq c$
- $x \notin L \longrightarrow \forall \tilde{\pi} : \Pr_\rho(V^{\tilde{\pi}}(x) = 1) \leq s$

and V^π makes no more than q queries to π consisting of characters from alphabet Σ with $|\pi| \leq l$. We then say V has completeness c and soundness error s.

The following PCP theorem is a seminal result in complexity theory.

Theorem 12.11. ***(PCP Theorem)*** $\mathsf{NP} = \mathsf{PCP}_{c,s}[\mathrm{poly}(n), O(\log n), O(1)]_{\{-1,1\}}$ *for any constant $c, s > 0$.*

It is conjectured that in fact much more powerful PCPs for NP exist:

Conjecture 12.12. ***(Unique Games Conjecture)*** *For any constant $c, s > 0$, there exists an alphabet Σ of constant size such that $\mathsf{NP} = \mathsf{PCP}_{c,s}[\mathrm{poly}(n), O(\log n), 2]_\Sigma$. Moreover, the verifier queries q_1, q_2 are nonadaptive (not dependent on values returned from π) and from disjoint sets, such that for each $\pi|_{q_1}$ there exists a unique value in Σ that $\pi|_{q_2}$ such that the verifier will accept, and vice versa.*

We will now see that this conjecture implies the optimality of the Goemans-Williamson approximation among polynomial-time algorithms.

12.5 Goemans-Williamson is Tight

Theorem 12.13. *Assuming the Unique Games Conjecture, it is NP-hard to approximate* MaxCut *to any factor greater than α_{GW}.*

Proof. We will show that, assuming an $(\alpha_{\mathrm{GW}} + \epsilon)$-approximation for MaxCut with $\epsilon > 0$, it is possible to decide any language in NP. We do this using a technique known as *proof composition*: we construct an "inner" PCP which proves that the "outer" PCP posited by the Unique Games Conjecture must accept some proof, then we use a $(\alpha_{\text{GW}} + \epsilon)$-MaxCut approximation to decide whether there exists a proof our inner PCP accepts. Let us start by constructing the inner PCP.

Let V_{out} be the outer PCP verifier. For each instance the outer PCP must decide, consider the bipartite graph $G = (L, R, E)$ where L is the set of all possible first queries, R is the set of all possible second queries, and E consists of edges between possible query pairs (q_1, q_2). Each proof π consists of assignments $L \longrightarrow \Sigma$ and $R \longrightarrow \Sigma$. By the Unique Games Conjecture, for each value a vertex in L takes on, there is exactly one accepting value its neighbors in R can take on, and vice versa. Equivalently, we can assign permutations σ_e to each edge $e \in E$, such that for any query pair the values i, j assigned to the vertices incident to its corresponding edge e must satisfy $j = \sigma(i)$. For accepting instances, more than $c|E|$ of the edges will be satisfied, while for rejecting cases, less than $s|E|$ of them will. Thus, we need our inner PCP to differentiate between these two cases. To do this, it will need *long codes*:

Definition 12.14. *The* Long Code *over Σ encodes $i \in \Sigma$ as the evaluation table of the dictator function $d_i : \{-1,1\}^\Sigma \longrightarrow \{-1,1\}$, where $d_i(\mathbf{x}) = x_i$.*

In our protocol, the honest proof will consist of the long code encoding of the correct character from Σ for each vertex that will make the outer PCP accept with probability $\geq c$. It is known that we can assume without loss of generality that L is regular, i.e. each vertex in L has the same degree. Thus, to sample the edges uniformly at random it suffices to sample a vertex in L then sample a random incident edge. Consider the following verifier, parameterized by $\rho \in (-1, 0)$:

1. Sample $u \in L$ then two incident edges $(u, v), (u, v') \in R$. Let σ and σ' be the permutations corresponding to (u, v) and (u, v') respectively.
2. Query f_v and $f_{v'}$ the long codes assigned to v and v' respectively.
3. Sample $\mathbf{x} \in \{-1,1\}^\Sigma$ uniformly at random
4. Sample $\mathbf{y}$ where each y_i is independently 1 with probability $1/2(1+\rho)$ and -1 otherwise.
5. Accept if and only if $f_v(\sigma^{-1}(\mathbf{x})) \neq f_{v'}(\sigma'^{-1}(\mathbf{x})\mathbf{y})$ where we denote $\sigma(x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma(n)})$.

- **Completeness:** If the proof is honest, it will be an encoding of a labeling where both (u, v) and (u, v') will be accepted with probability at least $1 - 2(1-c) = 2c - 1$. Let the labels of u, v, v' be $i, j, j' \in \Sigma$ respectively. With probability at least $2c - 1$, we have $\sigma(i) = j$ and $\sigma'(i) = j'$. Then,

$$f_v(\sigma^{-1}(\mathbf{x})) = x_{\sigma^{-1}(j)} = x_i$$
$$f'_v(\sigma'^{-1}(\mathbf{x})\mathbf{y}) = x_{\sigma'^{-1}(j')} y_{j'} = x_i y_{j'}$$

with $\Pr(x_i \neq x_i y_{j'}) = \Pr(y_{j'} = -1) = 1/2(1-\rho)$. Thus, the total completeness is

$$c' = (2c-1)\left(\frac{1}{2}(1-\rho)\right).$$

- **Soundness:** We will show that the soundness error is no more than $\dfrac{\arccos \rho}{\pi}$. Suppose some proof π

caused the verifier to accept with probability $\frac{\arccos \rho}{\pi} + \epsilon$. Note

$$\begin{aligned}
\Pr(\text{accept}) &= \mathbb{E}_{u,v,v',\mathbf{x},\mathbf{y}}\left[\frac{1}{2} - \frac{1}{2} f_v(\sigma^{-1}(\mathbf{x})) f_{v'}(\sigma'^{-1}(\mathbf{x})\mathbf{y})\right] \\
&= \frac{1}{2} - \frac{1}{2}\mathbb{E}_{u,\mathbf{x},\mathbf{y}}\left[\mathbb{E}_v[f_v(\sigma^{-1}(\mathbf{x}))]\mathbb{E}_{v'}[f_{v'}(\sigma'^{-1}(\mathbf{xy}))]\right] \\
&= \frac{1}{2} - \frac{1}{2}\mathbb{E}_{u,\mathbf{x},\mathbf{y}}[g_u(\mathbf{x})g_u(\mathbf{xy})] \\
&= \frac{1}{2} - \frac{1}{2}\mathbb{E}_u[\mathrm{St}_\rho(g_u)]
\end{aligned}$$

where $g_u(\mathbf{z}) = \mathbb{E}_{v:(u,v)\in E}[f_v({\sigma^{-1}}_{(u,v)}(\mathbf{z}))]$. Suppose

$$\Pr(\text{accept}) = \frac{1}{2} - \frac{1}{2}\mathbb{E}_u[\mathrm{St}_\rho(g_u)] \geq \frac{\arccos \rho}{\pi} + \epsilon$$

Then,

$$\mathbb{E}_u[\mathrm{St}_\rho(g_u)] \leq 1 - \frac{2\arccos\rho}{\pi} - 2\epsilon$$

so for at least an $\epsilon/2$ fraction of $u \in L$,

$$\mathrm{St}_\rho(g_u) \leq 1 - \frac{2\arccos\rho}{\pi} - \epsilon$$

as otherwise

$$\mathbb{E}_u[\mathrm{St}_\rho(g_u)] \geq (1 - \epsilon/2)\left(1 - \frac{2\arccos\rho}{\pi} - \epsilon\right) \geq 1 - \frac{2\arccos\rho}{\pi} - \frac{3\epsilon}{2}$$

a contradiction. Call such a vertex u "good". By Corollary 12.9 of the Majority is Stablest Theorem we know each good u has a coordinate, which we will call h, such that

$$\mathrm{Inf}_h^{\leq k}(g_u) \geq \delta$$

for some constant k and δ. Then,

$$\begin{aligned}
\delta &\leq \sum_{S\subseteq\Sigma: h\in S \wedge |S|\leq k} \hat{g_u}(S)^2 \\
&= \sum_{S\subseteq\Sigma: h\in S \wedge |S|\leq k} \mathbb{E}_v[\hat{f}_v({\sigma_{u,v}}^{-1}(S))]^2 \\
&\leq \sum_{S\subseteq\Sigma: h\in S \wedge |S|\leq k} \mathbb{E}_v[\hat{f}_v({\sigma_{u,v}}^{-1}(S))^2] \\
&= \mathbb{E}_v[\mathrm{Inf}^{\leq k}_{{\sigma_{u,v}}^{-1}(h)}(f_v)].
\end{aligned}$$

This implies that at least a $\delta/2$ fraction of neighbors v of u have

$$\mathrm{Inf}^{\leq k}_{{\sigma_{u,v}}^{-1}(h)}(f_v) \geq \frac{\delta}{2}$$

or otherwise

$$\mathbb{E}_v[\mathrm{Inf}^{\leq k}_{{\sigma_{u,v}}^{-1}(h)}(f_v)] \leq \left(1 - \frac{\delta}{2}\right)\frac{\delta}{2} + \frac{\delta}{2} \leq \delta$$

contradicting the above. Let $C_v = \{i \in \Sigma : \mathrm{Inf}_i^{\leq k}(f_v) \geq \delta/2\}$ be the candidate values for v. Then, at least $\delta/2$ of u's neighbors v have $h \in C_v$. Moreover, since

$$\sum_{i\in\Sigma} \mathrm{Inf}_i^{\leq k}(f_v) = \sum_{i\in\Sigma}\sum_{S\subseteq\Sigma: i\in S\wedge|S|\leq k} \hat{f}_v(S)^2 \leq k$$

we know $|C_v| \leq {}^{2k}/_{\delta}$. Thus, if we choose h as the value of u and a random $\ell \in C_v$ for each neighbor v, then a random edge incident to each good u has probability $({}^{\delta}/_2)({}^{\delta}/_{2k})$ of being accepting. This results in a constant $({}^{\epsilon}/_2)({}^{\delta}/_2)({}^{\delta}/_{2k})$ of the edges being satisfied, which does not depend on $|\Sigma|$. Thus, by the Unique Games Conjecture, we can choose $|\Sigma|$ large enough that this exceeds the outer soundness error of s, meaning it accepts. By the contrapositive, the soundness error s' of the inner PCP is no more than

$$\frac{\arccos(\rho)}{\pi} + \epsilon$$

for any $\epsilon > 0$.

Note now our inner PCP has completeness c' and soundness s' with ratio that can be made to be

$$\frac{c'}{s'} = \min_{-1<\rho<0} \frac{\arccos(\rho)/\pi}{{}^1/_2(1-\rho)(2c-1)} + \epsilon = \alpha_{\text{GW}} + \epsilon'$$

for any ϵ', by choosing ρ, c, ϵ appropriately. Also, our inner verifier has a special query structure — it queries at two uniform random points and accepts only if returned bits are not equal. For each instance for the outer PCP, consider the corresponding graph $G_{\text{in}} = (V_{\text{in}}, E_{\text{in}})$ where V_{in} is the set of query locations for the verifier and E_{in} are possible query pairs. Since the alphabet of the inner verifier is $\{-1, 1\}$, each possible proof corresponds to a cut of G_{in}, with the size of the cut proportional to the probability the verifier accepts this proof. If we have a ${}^c/_s$-approximation algorithm for MaxCut where ${}^c/_s = \alpha_{\text{GW}} + \epsilon$, we can differentiate between accepting and rejecting instances, and thus decide any language in NP. $\square$