In this homework we will consider the following variant of the Unique Games Conjecture, sometimes called the 2-to-2 Conjecture. It will be more convenient to use the following variant of the Label Cover problem, where we may not assign colors to all the vertices:

**Definition 1 (Label Cover).**

**Input:** A directed graph $G = (V, E)$, an alphabet $\Sigma$, and a constraint $\phi_e \subset \Sigma \times \Sigma$ for every $e \in E$.

**Output:** A set $S \subseteq V$ and an assignment $A : S \to \Sigma$ that satisfies the constraints on $(u,v) \in (S \times S)$. The goal is to (simultaneously) maximize the size of $S$ and the fraction of constraints on $S \times S$ satisfied.

We say that a constraint is 2-to-2 if for every assignment to variable $u$ there are (exactly) two assignments to its neighbor $v$ that would satisfy the constraint.

**Definition 2 (2-to-2 constraint).**

We say that a constraint $\phi \subset \Sigma \times \Sigma$ is 2-to-2 if $|\Sigma| = 2k$ is even and there exists permutations $\pi(u,v), \sigma(u,v) : \Sigma \to \Sigma$ such that:

$$\phi_{(u,v)} = \left\{ (\pi(2i - 1), \sigma(2i - 1)), (\pi(2i), \sigma(2i)), (\pi(2i), \sigma(2i - 1)), (\pi(2i - 1), \sigma(2i)) \right\}_{i=1}^{k}.$$ 

**Conjecture 1 (2-to-2 Conjecture).** The following holds for every constant $\epsilon > 0$. Given an instance of Label Cover with 2-to-2 constraints, it is NP-hard to distinguish between:

**Completeness** There exists an assignment to all variables that satisfies every constraint; and

**Soundness** any assignment to $\epsilon$-fraction of the variables satisfies at most $\epsilon$-fraction of the constraints between them.

1. **(Warmup)** Suppose we replace the 2-to-2 constraints in Conjecture 1 with unique (i.e. 1-to-1) constraints. Explain\(^1\) why this would be different from the Unique Games Conjecture and false assuming $P \neq NP$.

\[^1\] Here and thereafter, *explain* means "write a couple of hand-wavy sentences".
2. (Guided) Prove that assuming Conjecture 1, the following holds for every constant integer \( c > 4 \).
Given a graph \( G' = (V', E') \), it is \( \text{NP} \)-hard to distinguish between:

**Completeness** \( G' \) is 4-colorable\(^2\)

**Soundness** \( G' \) is not \( c \)-colorable.

(a) **Defining** \( V' \): Similar to Unique-Games hardness \( \text{Max-Cut} \), we will construct a “cloud” cloud\((v)\) of \( G' \)-vertices for every variable \( v \in V \). But now the cloud will correspond to \( \{0, 1, 2, 3\}^\Sigma \) (instead of \( \{0, 1\}^\Sigma \) as in \( \text{Max-Cut} \)). For \( x \in \{0, 1, 2, 3\}^\Sigma \) and permutation \( \pi : \Sigma \to \Sigma \), we abuse notation and let \( \pi(x) \in \{0, 1, 2, 3\}^\Sigma \) denote the vector whose \( i \)-th coordinate is \( x_{\pi(i)} \).

**Do:** Explain why we should use \( \{0, 1, 2, 3\} \) instead of binary \( \{0, 1\} \).

(b) **Defining** \( E' \): For \( u, v \in V \) and \( x, y \in \{0, 1, 2, 3\}^\Sigma \), we construct an edge \( ((u, x), (v, y)) \in E' \) iff for every \( 2i, 2i + 1 \in \Sigma \),

\[
\{x_{\pi(u,v)}(2i) \textbf{-} x_{\pi(u,v)}(2i+1)} \cap \{y_{\pi(u,v)}(2i) \textbf{-} y_{\pi(u,v)}(2i+1)} = \emptyset.
\]

(In words: for every pair \( (i, j) \in \Sigma \times \Sigma \) that satisfies constraint \( \phi_{(u,v)} \), the colors in the respective coordinates must be different.)

**Do:** Prove completeness, i.e. that if the original 2-to-2 instance \( G \) is satisfiable, then \( G' \) is 4-colorable.

(c) **Assigning probabilities to the edges:** In the reduction to \( \text{Max-Cut} \) the probability of an edge corresponded to negating every coordinate of the corresponding Boolean vector, and adding noise. Here, we will modify this strategy in a few ways:

- The actual graph \( G' \) is unweighted (for 4-coloring, weights are meaningless); we already defined the edges above, and will only use the weights in the analysis.
- Because of the 2-to-2 constraints, we have to update the coordinates in pairs (i.e. they are no longer independent). Let \( Q := \{0, 1, 2, 3\}^2 \). Below we define a Markov operator \( T : Q \to Q \) that acts on each pair of coordinates and determines the weight on the edges.
- We do not add noise.

**Do:** Explain why not adding noise is a good idea for 4-coloring.

(d) **Defining the Operator** \( T \): We now want to define the operator \( T : Q \to Q \). We have the following desiderata:

- We want the operator to be *reversible*, i.e. \( T^{-1} = T \). In other words, we need to define a weighted, *undirected* graph on \( Q \).
- To avoid trivialities like the identity operator, we want the operator to mix well. Formally, we require that the *spectral gap* of \( T \) is bounded away from 1. Since in our case \( Q \) is of constant size, it is enough to show that \( T \) (as a graph) is connected and non-bipartite.
- Finally, we want \( T \) to assign positive probability only to the edges in \( E' \). This means that for \( (x^1, x^2), (y^1, y^2) \in Q \), we have an edge from \((x^1, x^2)\) to \((y^1, y^2)\) iff \( x^1, x^2 \notin \{y^1, y^2\} \). In other words, let \( T^{\otimes k} : Q^k \to Q^k \) be the operator that applies \( T \) independently to every pair of coordinates in its inputs. Then \( ((u, x), (v, y)) \in E' \) if and only if \( \Pr[T^{\otimes k}(\pi(u,v)x) \leftrightarrow \sigma(u,v)(y)] > 0 \).

**Do (optional):** Construct such an operator \( T \).

**Hint:** Assign probability 1 to each different case:\( (x^1, x^2) \leftrightarrow (y^1, y^2) \) for some vector \( \tilde{x} \).

\(^2t\)-colorable means that we can color the vertices of the graph with \( t \) colors such that the endpoints of every edge have different colors.
(e) **Skipping the analysis of Boolean functions:** Like Unique-Games hardness results (and many other results in TCS), our analysis requires a theorem from the analysis of Boolean functions.

**Theorem 1** (Informal). *For every constant \( \varepsilon > 0 \) there exists a constant \( \ell = \ell(\varepsilon) > 0 \) such that the following holds. Suppose that a pair of functions \( f, g : \{0, 1, 2, 3\}^\Sigma \rightarrow \{0, 1\} \) satisfies:

\[
\mathbb{E}[f(x)], \mathbb{E}[g(x)] \geq \varepsilon \quad \text{AND} \quad \mathbb{E}[f(x) \cdot g(T(x))] = 0.
\]

Then each of \( f, g \) has a small set \( S_f, S_g \subseteq \Sigma \) of influential coordinates (\(|S_f|, |S_g| \leq \ell\)), and there is a pair \((2i - 1, 2i)\) such that both \( S_f \) and \( S_g \) contain at least one of the pair (but maybe not the same one).

(The definition of “influential” is irrelevant for our purposes — we only need that there is some such set of variables.)

**Do: nothing!**

(f) **Soundness — do:** Prove that if the original 2-to-2 instance does not have an assignment for \( \varepsilon \)-fraction of the variables satisfying at least an \( 1/\ell(\varepsilon)^2 \)-fraction of the constraints, then \( G' \) is not \( 1/(2\varepsilon) \)-colorable.

**Hint:** Assume by contradiction that \( G' \) has an independent set \( I \) of size \( |I| \geq \varepsilon |V| \). Consider the indicator function of this independent set. Use the set of influential coordinates of this indicator function to assign a value to this variable.