We continue from where we left in the previous lecture.

3 Gap amplification of nice $2\text{CSP}_w$ instances

We have seen in the last lecture that a $2\text{CSP}$ instance $\varphi$ is said to be nice if and only if:

$\bullet$ $G(\varphi) = (V, E)$ is a $d$-regular graph, where $G(\varphi)$ denotes the constraint graph corresponding to $\varphi$ with variables as vertices and constraints as edges (see Definition 20, Lecture 18)

$\bullet$ For every vertex $v \in V$, at least half of the incident edges are self-loops

$\bullet$ $G(\varphi)$ is an expander

Since one can reduce arbitrary CSP instances to nice $2\text{CSP}$ instances (Section 2.6, Lecture 18), it is sufficient to prove Dinur’s gap amplification lemma for nice $2\text{CSP}$ instances.

Lemma 27 (Powering lemma) Given a nice $2\text{CSP}_w$ instance $\psi$ and a positive integer $t$ we can produce a $2\text{CSP}_{w^t}$ instance of $\psi^t$ in time $\text{poly}(n, w^{5t})$ – where $d$ is the degree of $\psi$’s constraint graph – such that:

$\bullet$ the alphabet size of $w'$ is $\leq |w|^{d^t}$ and the number of constraints in $\psi^t$ is $\leq d^{t+2} \cdot n$

$\bullet$ $\psi$ is satisfiable $\Rightarrow \psi^t$ is satisfiable

$\bullet$ for every $\epsilon < \frac{1}{d^{t+1}}$, $\text{val}(\psi) \leq 1 - \epsilon \Rightarrow \text{val}(\psi^t) \leq 1 - \frac{\sqrt{t}}{w^{5t} \cdot d^t} \cdot \epsilon$ (gap amplification)

Proof

Let us first argue the way $\psi^t$ is constructed:

Variables: Let $\psi^t$ be a nice $2\text{CSP}$ instance such that it has the same number of variables as $\psi$. In other words, $\psi^t$ has a “new” variable $y_i$ for each “old” variable $u_i$ in $\psi$. We shall think of the domain of the new variables to be a $d^t$-tuple over $\{0, 1, ..., w-1\}$, i.e., of size $|w|^{d^t}$.

The intuition behind the new variables is that $y_i$ claims a value of each old variable $u_j$ if $j$ lies in the ball centered in $i$ and with radius $t + \sqrt{t}$ in $G$. The number of such vertices is $d^{t+\sqrt{t}+1} \leq d^t$ since $G$ is a $d$-regular graph and therefore the alphabet size $w'$ is sufficiently large for encoding this information.

Constraints: For each $(2t + 1)$-step path $p_i = \langle i_1, ..., i_{2t+2} \rangle$ there is a constraint $C_p$ in $\psi$ that depends only on $y_{i_1}$ and $y_{i_{2t+2}}$ (so it is a $2\text{CSP}_{w^t}$) and outputs false iff $\exists j \in \{1, ..., 2t + 1\}$ such that:

1. $y_{i_1}$ claims value $\omega$ of $i_j$

2. $y_{i_{2t+2}}$ claims a value $\omega'$ of $i_{j+1}$

3. $(\omega, \omega')$ does not satisfy the original constraint in $\psi$

Notice that constructing $\psi^t$ takes $\text{poly}(n, w^{d^t})$, $w' \leq |w|^{d^t}$ and the number of paths/constraints are at most $d^{t+2} \cdot n$ (since the constraint graph of $\psi$ is $d$-regular). Also, an assignment to the variables $u_1, ..., u_n$ in $\psi$ is naturally “expanded” to $y_1, y_2, ..., y_n$ by assigning to each $d^t$-tuple $y_i$ all values of $u_j$ that lie in the ball centered in $i$ and with radius $t + \sqrt{t}$ in $G(\varphi)$. In this way, if $u_j$ is a satisfying assignment for $\psi$, then so is $y_i$ for $\psi^t$ since it satisfies all constraints encountered on paths of length...
\[ \leq 2t+1. \] Thus, the only property of the Powering lemma that we need to prove is the soundness property (i.e., the gap amplification).

In what follows we will argue how to transform an assignment \( y \) of \( \psi^t \) into an assignment \( u \) of \( \psi \) such that if \( u \) violates a few constraints, then \( y \) violates a fraction \( \Omega(\frac{\sqrt{t}}{10^d\epsilon_d}) \) of \( \psi^t \)'s constraints.

The analysis has the following outline:

1. Consider an arbitrary assignment \( y \) and decode it to get assignment \( u \) to \( \psi \)
2. Since \( u \) leaves \( m\epsilon \) constraints unsatisfied this implies that \( y \) leaves a \( \epsilon' \) fraction of constraints unsatisfied where \( \epsilon' = \frac{\sqrt{t}}{10^d\epsilon_d}\epsilon \)

**Plurality assignment.** For every \( y_i \) consider a random variable \( Z_i \) which represents the outcome of the following process:

1. Start from vertex \( i \) and take a \( t \)-step random walk to reach a vertex \( k \)
2. Output the value that \( y_k \)'s claims for \( u_i \) (this is intuitive since we want to determine the value of \( i \))

Let \( z_i \) be the most likely outcome of \( Z_i \), thus the plurality assignment of the variables \( y_1,\ldots,y_n \) is \( z_1,\ldots,z_n \). Since \( val(\psi) \leq 1 - \epsilon \) there exists a subset \( F \) of constraints in \( \psi \) such that \( |F| = m\epsilon = \frac{nd\epsilon}{2} \) and the plurality assignment violates these constraints.

Consider a \((2t+1)\)-step path \( \langle i_1,\ldots,i_{2t+1} \rangle \in G(\varphi) \) (or constraint \( \in \psi^t \)) picked uniformly at random.

An edge \( e = (i_j,i_{j+1}) \), \( j \in \{1,2,\ldots,2t+1\} \) is truthful if \( i_1 \) claims the plurality value of \( i_j \) and \( i_{2t+2} \) claims the plurality value of \( i_{j+1} \). Thus, if \( e \in F \) is truthful, then the constraint corresponding to this path is not satisfied. Through the next sequence of claims we show that at least a \( \frac{\sqrt{t}}{10^d\epsilon_d}\epsilon \) fraction of the paths have such truthful edges.

**Claim 28 (Random edge)** Let \( G = (E,V) \) be an arbitrary graph. Then it follows that for each edge \( e \in E \) and each \( j \in \{1,\ldots,2t+1\} \)

\[
Pr[e \text{ is the } j \text{'th edge on path } p] = \frac{1}{|E|}
\]

where the probability is taken over a random walk.

**Proof** We prove this using an example.

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
1 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
2 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
3 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
4 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\
\end{array}
\]
Start from 1. With probability $1/2$ we end up on 2 or on 4. Then $G \cdot (1 \ 0 \ 0 \ 0)^T$ is the 1-step probability distribution.

For the general case, $G \cdot (1/n \ldots \ 1/n)^T = (1/n \ldots \ 1/n)^T$ where $n = |V|$.

The probability that the $j$-th edge in the path belongs to $F$ is $|F|/|E|$. Thus, a claim is truthful if it agrees with the plurality assignment. If an edge lies roughly in the middle of a path (i.e. within $d \in \{t - \delta \sqrt{t}, \ldots, t + \delta \sqrt{t}\}$ steps), then it is very likely to be truthful.

**Claim 29 (Middle-truthful)** Let $\delta < \frac{1}{100w}$. For each $e \in E$, each $j \in \{t, t + 1, \ldots, t + \delta \sqrt{t}\}$. Then

$$\gamma = Pr[ j \text{'th edge is truthful} | e \text{ is j} \text{'th edge}] \geq \frac{1}{2w^2}$$

**Proof** Let $e = (i_j, i_{j+1})$. The key idea is to launch two independent random walks – one $j$-step walk from $i_j$ and one $(2t - j)$-step walk from $i_{j+1}$.

If $j = t$ then start a $t$-step random walk from $i_j$ and a $t$-step random walk from $i_{j+1}$; this is exactly the process for defining plurality. Finally, since the two walks are independent $\gamma \geq \frac{1}{w}$.

However, if $j$ varies in $\{t, t + 1, \ldots, t + \delta \sqrt{t}\}$ we produce a $\ell$-step random walk where $\ell \in \{t - \delta \sqrt{t}, t - \delta \sqrt{t} + 1, \ldots, t + \delta \sqrt{t}\}$ as follows

- Flip $\ell$ coins and let $S_\ell$ be the number of times it turned out heads
- Walk $S_\ell$ "real" edges (no self-loops)

$S_\ell$ follows a binomial distribution and thus similar $S_\ell$ since $\ell$ differs from $t$ by an additive factor $\delta \sqrt{t}$ which is small ($\delta \leq \frac{1}{100w}$). It can be shown that

$$\frac{1}{2} \sum \alpha \frac{1}{2} |Pr[S_\ell = \alpha] - Pr[S_\ell = \alpha]| \leq 10\delta$$

In other words, the distribution of a $t$-step walk is statistically close to both the distribution of a $(t - \delta \sqrt{t})$-step random walk and the distribution of a $(t + \delta \sqrt{t})$-step random walk. Therefore, considering the definition of a truthful edge,

$$\gamma = Pr[y_{i_j} \text{ claims the plurality value of } i_j \wedge \text{ claims the plurality value of } i_{j+1}] \geq \left(1 - \frac{10\delta}{w} - 10\delta\right) \geq \frac{1}{2w^2}$$

Let $V$ be the set of truthful edges in the middle of the path that belong to $F$. The probability that $j \in \{t, t + 1, \ldots, t + \delta \sqrt{t}\}$ edge is in $F$ and truthful is $\frac{|F|}{|E|} \cdot \frac{1}{2w^2}$. Using the linearity of expectation

$$E[V] = \frac{|F|}{|E|} \cdot \frac{1}{2w^2} \cdot \delta \sqrt{t} = \frac{e\delta \sqrt{t}}{2w^2}$$

It is expected to have many truthful edges, however some paths may contain only violated constraints. Note that $Pr[V > 0] \geq \frac{E[V]^2}{E[V]}$. Our final goal is to prove that this probability is at least $\frac{\sqrt{t}}{100w} \cdot \epsilon$ (which implies that $\psi^e$ satisfies a fraction of at most $1 - \frac{\sqrt{t}}{100w} \cdot \epsilon$ constraints). Claim 29 shows that $E[V]$ is high, however $Pr[V > 0]$ could potentially still be 0 with a large probability. To bound $Pr[V > 0]$ we upper bound the variance of $V$ (actually $E[V^2]$). This is the only step where we use that $G(\varphi)$ is an expander.
Claim 30 \( E[V^2] \leq 25\epsilon d \sqrt{t} \)

Proof Let \( V' \) be the number of middle edges (i.e., edges within \( t \pm \delta \sqrt{t} \) steps) that are in \( F \). Clearly \( V' \geq V \) since \( V \) counts the number of edges in \( F \) that are truthful and so \( E[V'^2] \geq E[V^2] \).

Let \( I_j \) be the indicator function for \( F \).

\[
V' = \sum_{j \in \{t, \ldots, t + \delta \sqrt{t}\}} I_j 
\implies E[V'] = \sum_{i,j} E[I_i \cdot I_j] = \sum_{i} E[I_i^2] + 2 \sum_{i < j} E[I_i I_j] 
\leq \epsilon \delta \sqrt{t} \text{ since } E[I_i] = \epsilon
\]

Also,

\[
E[I_i \cdot I_j] = Pr[\text{the } i^{th} \text{ edge is in } F \text{ and } j^{th} \text{ edge is in } F] 
\leq \frac{|S|}{|G|} \left( \frac{|S|}{|G|} + \lambda(G)^{j-1} \right),
\]

where \( S \) is the set of endpoints of an edge in \( F \). Clearly we have \( |F| = \frac{\epsilon n d}{\omega} \) and \( |S| \leq \epsilon nd \). Thus,

\[
E[V^2] \leq \sum_{i \leq j} E[I_i \cdot I_j] \leq \sum_{i \leq j} \left( \frac{|S|}{|V|} \right)^2 + \frac{|S|}{|V|} \lambda(G)^{j-1} \leq \sum_{i \leq j} (\epsilon d)^2 + \sum_{i \leq j} \epsilon d \lambda(G)^{j-1} 
\leq \epsilon^2 d^2 \delta^2 t + \sigma \delta d \sqrt{t} \sum \lambda(G)^{k} \leq \epsilon^2 d^2 \delta^2 t + 10 \delta \sqrt{t} ed \leq 25 \epsilon d \delta \sqrt{t} \text{ (by assumption } \epsilon < \frac{1}{d \sqrt{t}})\]

where \( * \) holds since we use a good enough expander such that \( \lambda(G) \leq \frac{1}{17} \).

Finally using the fact that \( Pr[V > 0] \geq \frac{(E[V])^2}{E[V^2]} \geq \frac{(E[V])^2}{E[V^2]} = \frac{\epsilon^2 d^2 \delta^2 t}{4w^4} \cdot \frac{1}{25 \epsilon d \delta \sqrt{t}} = \frac{\delta \sqrt{t}}{100 w^4 \epsilon} \). Hence, \( \psi^t \) satisfies at most \( 1 - \frac{\delta \sqrt{t}}{100 w^4 \epsilon} = \frac{1}{100 w^4} \epsilon \) and so Lemma 27 follows.