Set $t \approx \log_2(m/\varepsilon_0)$ so that $2^t/m \in [\varepsilon_0/2, \varepsilon_0]$

Done?

what is size?

$$|R^{(t)}(\varphi)| \leq C \cdot |R^{(t-1)}(\varphi)| \leq \ldots \leq C^t \cdot |\varphi|$$
$$= \mathrm{poly}(m)|\varphi|$$

$\square$

Def: CL-reduction ("complete linear-blowup reduction")

R from Max q-CSP_W to Max q'-$CSP_{W'}$:

* Perfect completeness (val=1 to val=1)
* poly-time
* Constant factor size blowup (∃ universal const. C s.t. $|R(\varphi)| \leq C \cdot |\varphi|$

Lemma 22.5

$\forall \ell$ $\exists W, \varepsilon_0 > 0$ and CL-reduction R_ℓ^+ from Max 2-CSP_3 to Max 2-CSP_W s.t. if $\mathrm{val}(\varphi) \leq 1 - \varepsilon$ for $\varepsilon \leq \varepsilon_0$ then

$$\mathrm{val}(R_\ell^+(\varphi)) \leq 1 - \ell \cdot \varepsilon$$

Lemma 22.6

For all W there is a CL-reduction R_W^- from Max 2-CSP$_W$ to Max 2-CSP$_3$ s.t. if val$(\varphi) \leq 1-\varepsilon$ then val$(\varphi) \leq 1 - \varepsilon/3$

<u>Note</u>: while we lose a bit, the loss is <u>independent of W!</u>

Proof of Lemma 22.4 from 22.5 + 22.6

Given Max 2-CSP$_3$ instance φ, apply R_6^+ then R_W^- (where W is alph. size of R_6^+)

	val	size	alphabet
φ	$1-\varepsilon$	$\lvert\varphi\rvert$	3
$\downarrow$			
$R_6^+(\varphi)$	$1-6\varepsilon$	$C_1 \lvert\varphi\rvert$	$W(6)$
$\downarrow$			
$R_W^-(R_6^+(\varphi))$	$1-\frac{6\varepsilon}{3}$ $=1-2\varepsilon$	$C_2 \cdot C_1 \cdot \lvert\varphi\rvert$	3

$\square$

Proof ideas in Lemma 22.6

* constraint on 2 W-ary variables can be thought of as a constraint on $2 \cdot \log_2 W$ binary variables

* Think of each constraint as a little mini-instance on $2 \log \cancel{W}$ variables.
* We would like to encode it in such a way that we can check it while reading only a small number of bits (indep. of W)
  (That would get us to Max q-CSP_2 instead of Max 2-CSP_3 but the last step is a simple gadget.)

Does this sound familiar?
Isn't this pretty much the PCP thm?
Circularity alert?!

Way out:

2 log W is maybe large but it is a constant indep. of $|\varphi|$, so we can afford something pretty inefficient like an exponential-sized PCP.

(Not quite the whole picture, need to ensure consistency between constraints)

Back to Lemma 22.5

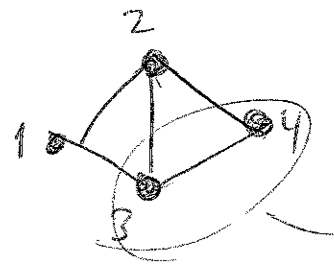
We're going to need some extra properties of φ:

1. Assume φ is d-regular for some constant d
2. Assume φ is a good expander graph (won't get sufficiently deep into details to see how this is used so will skip defn.)

Prop can assume props hold WLOG.

Key operation in Lemma 22.5: graph powering

Constraint graph:



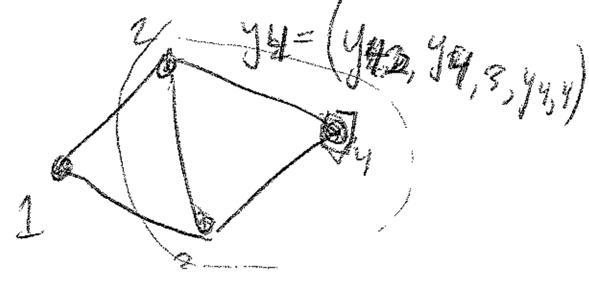
constraints are edges

Constraint on x_3, x_4

t'th power of φ, φ^t:

- Same #variables, $y_1, \dots, y_n$ over larger alphabet. The variable y_i has an "opinion" about the value of x_j for all j at distance $\leq t + \sqrt{t}$ from i in φ.

i.e. value of y_i is a list of values from $[3]$ for all such j



Alphabet size?

$\le d^{\ell}$ neighbours at distance ℓ

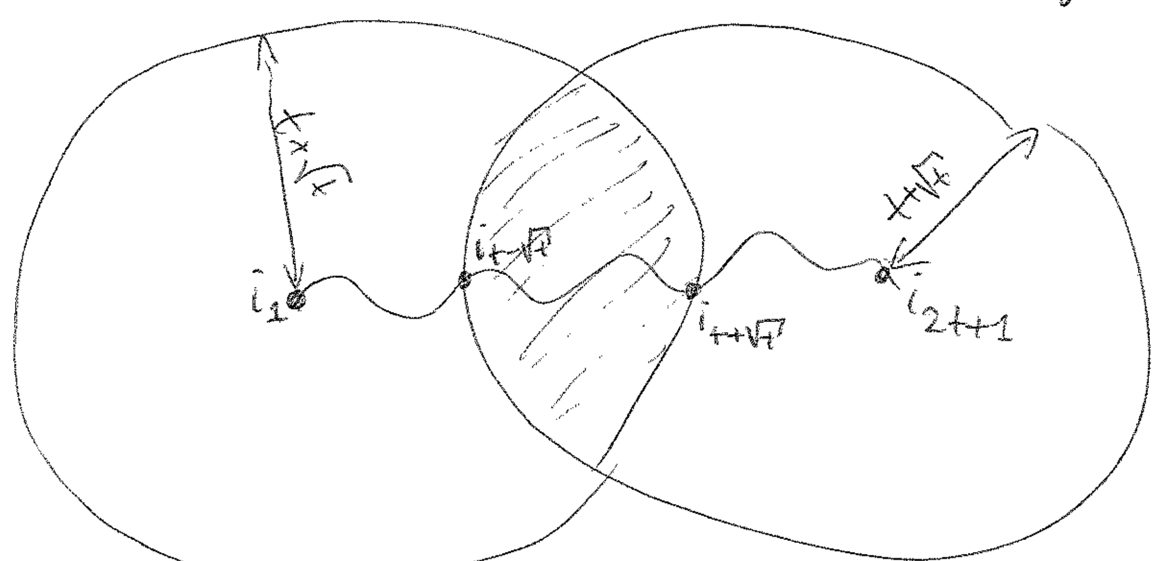
$\Rightarrow \le d^{t+\sqrt{t}+1}$ neighs at dist $\le t+\sqrt{t}+1$

$\Rightarrow$ alphabet size $W = 3^{d^{t+\sqrt{t}+1}}$

Constraints in φ^t:

Here we really needed regularity! (And low degree)

Let $i_1, \dots, i_{2t+1}$ be path of length $2t$



y_{i_1} and $y_{i_{2t+1}}$ both contain (opinions of) values of the variables in . Add constraint on $(y_{i_1}, y_{i_{2t+1}})$ which is sat if opinions of $y_{i_1}, y_{i_{2t+1}}$ are consistent and satisfy all constraints in

Proposition

There are constants $\gamma, \delta > 0$ s.t.
if $val(\varphi) \le 1-\varepsilon$ and $t < \gamma/\varepsilon^2$ then
$$val(\varphi^t) \le 1 - \delta \cdot \varepsilon \cdot \sqrt{t}$$

(This is where expander property is needed)

How this gives Lemma 22.5:

Set $t = (l/\delta)^2$ (const. depending only on l)

If $(l/\delta)^2 < \gamma/\varepsilon^2$ (i.e. $\varepsilon < \sqrt{\gamma} \cdot \delta / l := \varepsilon_0$)

~~then~~ and $val(\varphi) \le 1-\varepsilon$ then $val(\varphi^t) \le 1 - \delta\varepsilon\sqrt{t} = 1 - l\cdot\varepsilon$

alphabet size $W = 3^{d^{t+\sqrt{t}+1}} \le 3^{d^{2(l/\delta)^2}}$

which is a (huge) constant depending only on l (d, δ are universal)

$\square$

Thm [Håstad '97]:

$\forall \varepsilon > 0$, given MAX 3-SAT instance φ it is NP-hard distinguish between

$$\mathrm{val}(\varphi) = 1 \quad \text{and} \quad \mathrm{val}(\varphi) \leq 7/8 + \varepsilon$$

Perspective:

given MAX 3-SAT instance φ, consider the completely mindless algorithm which just picks a random assignment to the variables.

Claim: in expectation, this algorithm satisfies $\frac{7}{8}$ of the clauses

Håstad's Thm: The mindless algorithm is optimal!

(MAX 3-SAT is approximation resistant)

Another problem

MAX 3-LIN:

given system of linear equations of form $x_i \oplus x_j \oplus x_k = b$

Håstad '97:

MAX 3-LIN is $(1-\varepsilon)$-vs-$(\frac{1}{2}+\varepsilon)$-NP-hard

~~Goal: prove weak version of~~ this.

Key parts in proof:

* Encoding label cover variables as collection of binary variables
* Fourier Analysis!

Repetition of PCP's:

t-fold sequential rep:

new verifier V':

- repeat t times: (with independent random choices)
  run V on proof

$$PCP_{c,s}[r,q,\Sigma] \subseteq PCP_{c^t,s^t}[t\cdot r, t\cdot q]$$

Parallel repetition

t-fold parallel rep.:

pick t checks

$$(i_1^1, \ldots, i_1^q), C_1$$
$$(i_2^1, \ldots, i_2^q), C_2$$
$$\vdots$$
$$(i_t^1, \ldots, i_t^q), C_t$$

Suppose $|\pi| = L$.

Expect as proof string π' over alphabet Σ^t indexed by L^t where

$$\pi'(i_1, \ldots, i_t) = (\pi(i_1), \pi(i_2), \ldots \pi(i_t))$$

$$\begin{array}{ccc} (i_1^1, & \dots, & i_1^q) \\ (i_2^1, & \dots, & i_2^q) \\ \vdots & & \vdots \\ (i_t^1, & \dots, & i_t^q) \\ \Downarrow & & \Downarrow \\ I_1 & \dots & I_q \end{array}$$

Ask for $\pi'(I_1), \dots, \pi'(I_q)$

get values for $\{\pi(i_\cdot^\cdot)\}$

check $C_1(i_1^1, \dots, i_1^q)$

$\vdots$

$C_t(i_t^1, \dots, i_t^q)$

satisfied

Parameters of V':

$$PCP_{?,?}[t \cdot r, q, \Sigma^t]$$

Completeness c' of V': $c' \geq c^t$

Soundness s' of V': $s' \leq s^t$?

No

Raz': If $s \leq 1-\varepsilon$ then $s' \leq (1-\mathrm{poly}(\varepsilon))^t$

Label Cover

PCP Theorem: $PCP_{1,1-\varepsilon}[O(\log n), 2, 3] \ni SAT$

Add parallel repetition: $\forall \varepsilon\ \exists L$ s.t.

$$SAT \in PCP_{1,\varepsilon}[O(\log n), 2, L]$$

Def An instance Ψ of Label Cover is a tuple $(U, V, E, \Sigma, \{\pi_e\})$ where

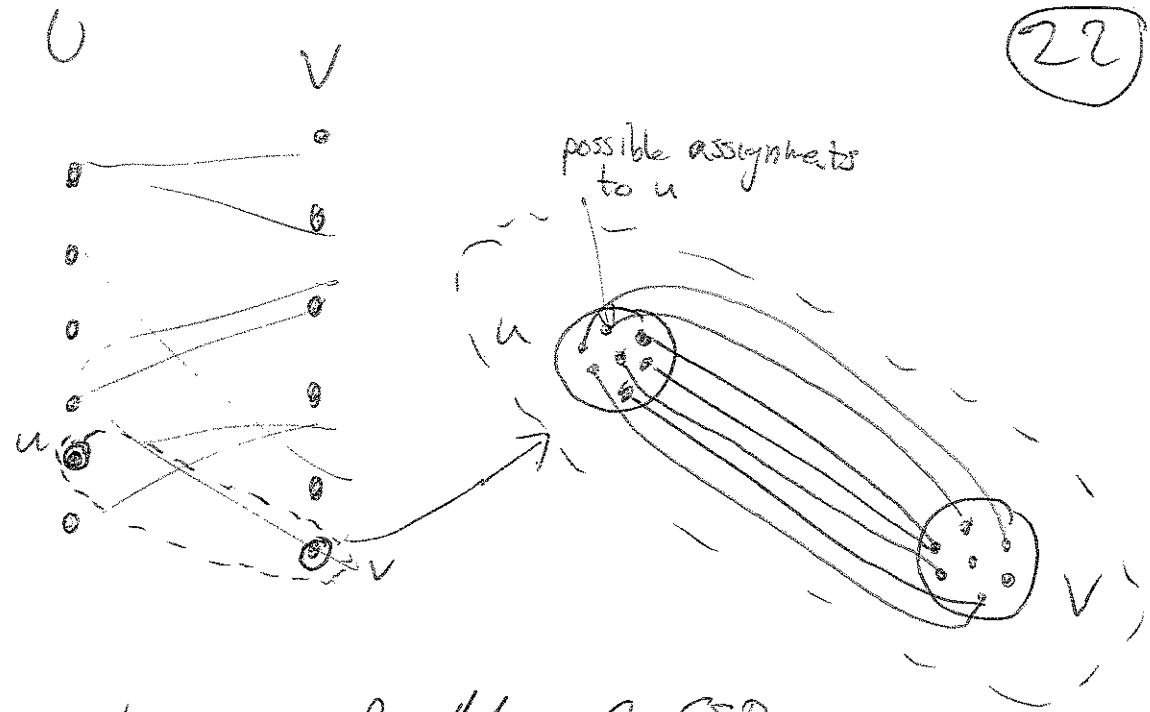
(U, V, E) is a bipartite graph

Σ is alphabet

For each edge $e \in E$ there is a constraint $\pi_e : \Sigma \to \Sigma$

A labelling is an assignment $\alpha : U \cup V \to \Sigma$ of labels to the vertices.

Edge $e = (u,v)$ is satisfied by labelling if $\alpha(v) = \pi_e(\alpha(u))$



A special case of Max 2-CSP$_\Sigma$

<u>PCP Thm + Parallel Rep.:</u>

$\forall \varepsilon \; \exists \Sigma$ s.t. given Label Cover instance φ on alphabet size Σ, it is NP-hard to distinguish between
$$\text{val}(\varphi)=1 \quad \text{and} \quad \text{val}(\varphi) \le \varepsilon.$$

WHY?

Turns out to be exceptionally useful starting point to prove results for problems we care about!