In our discussion of probabilistically checkable proofs (PCPs) last time, we saw two equivalent descriptions of this concept: One was stated in terms of proof verification, the other ones in terms of reduction to constraint satisfaction. For the proof of the PCP theorem it is more convenient to adopt the second perspective. Then a PCP of query complexity q and randomness complexity $\log m$ is a polynomial-time reduction that maps instances x of the problem to qCSP instances Ψ with m constraints such that

$$\begin{aligned} \text{If } x \in \mathit{YES}, &\quad \text{then } \Psi \text{ is satisfiable,} \\ \text{If } x \in \mathit{NO}, &\quad \text{then no assignment satisfies more than half of } \Psi\text{'s constraints.} \end{aligned}$$

The PCP theorem states that every problem $(\mathit{YES}, \mathit{NO})$ in NP has this property.

It will be useful to pay attention to two more parameters of the PCP which we fixed to constants in the above definition. One is the *soundness* which is the maximum fraction of satisfied constraint for *NO* instances, set to $1/2$ above. Another one is the *alphabet size*: In our definition of qCSPs we assumed that the variables take boolean values. Today we will also look at CSPs over variables that take values in some larger alphabet Σ. (In the proof verification view of PCPs this means the proof symbols come from alphabet Σ.)

1 The proof of the PCP theorem

The starting point of the proof of the PCP theorem is the fact that the NP-hardness of 3SAT can be viewed as an extremely weak hardness of approximation result: The *YES* instances are satisfiable CNFs, and the *NO* instances are those in which a $s = 1 - 1/m$ fraction of clauses cannot be simultaneously satisfied, where m is the number of clauses.

We will start from this statement and design a sequence of transformations that gradually improve the soundness parameter s, while leaving all the other parameters unchanged. It is more convenient to prove this reduction for general constraint satisfaction problems.

	Ψ (qCSP instance)	$\to \Psi'$ (qCSP instance)
number of constraints	m	$\to Cm$ (C is some constant)
completeness	Ψ is satisfiable	$\to \Psi'$ is satisfiable
soundness	$1 - \delta$	$\to 1 - 2\delta$ (for $\delta < 1/C$)

If we start with a 3SAT instance and repeat this reduction $O(\log m)$ times the soundness drops from $1 - 1/m$ to a fixed constant, while the size of the instance remains polynomial in m. This is exactly the PCP theorem. The reduction is implemented by composing several smaller reductions, where at each stage one of the parameters is improved, but at the expense of the others.

An important theme in the proof of the PCP theorem is to work, whenever possible, with CSPs over possibly larger alphabet but with only two variables per constraint. The presence of various constraints in such instances can be encoded by a *constraint graph*: This is the undirected graph whose vertices v correspond to variables x_v in the CSP and where for each constraint $\psi(x_u, x_v)$ in Ψ there is an edge (u, v) in the graph. The satisfiability of CSP instances will turn out to be closely connected to the expansion properties of this graph, so we take a detour to discuss graph expansion next.

2 Graph expansion

An expander is a graph that remains well connected even if some edges from the graph are removed. We will focus on d-regular n-vertex graphs, where we think of d as a constant and n as a growing quantity.

Definition 1. A d-regular graph is an *α-edge expander* if for every subset S of at most $n/2$ vertices

$$|E(S, \overline{S})| \geq \alpha d \cdot |S|$$

where $E(S, \overline{S})$ is the set of edges between S and $\overline{S}$.

As $d|S|$ is the largest possible number of edges that can come out of S, the definition postulates that the number of edges coming out of any set is within a constant factor of the maximum possible.

Edge expansion is closely related to a stochastic property of graphs called the spectral gap. A d-regular graph can be viewed as a Markov chain whose states are the vertices and whose transitions are moves to a uniformly random neighbor. If the graph is not bipartite then the uniform distribution is the single stationary distribution of this Markov chain.

The transition matrix A of this Markov chain is the adjacency matrix of the graph scaled by $1/d$. This is a symmetric matrix so all its eigenvalues are real. If $\mathbf{p}$ is a probability distribution over the vertices then we can represent $\mathbf{p}$ as

$$\mathbf{p} = \alpha_1 \mathbf{v}_1 + \cdots + \alpha_n \mathbf{v}_n$$

where $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is an orthonormal basis of eigenvectors with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ where we call the largest one λ_1. After one step of the Markov chain the distribution becomes

$$\mathbf{p}A = \lambda_1 \alpha_1 \mathbf{v}_1 + \cdots + \lambda_n \alpha_n \mathbf{v}_n$$

and after t steps

$$\mathbf{p}A^t = \lambda_1^t \alpha_1 \mathbf{v}_1 + \cdots + \lambda_n^t \alpha_n \mathbf{v}_n.$$

Since this must eventually converge to the uniform distribution, it is not difficult to deduce that $\lambda_1 = 1$, $|\lambda_i| < 1$ for all other i and $\alpha_1 \mathbf{v}_1$ is the uniform distribution $\mathbf{u}$. It then follows that

$$\|\mathbf{p}A^t - \mathbf{u}\| \leq \max_{i \geq 2} |\lambda_i|^t$$

so the value $\lambda = \max_{i \geq 2} |\lambda_i| \in [0, 1)$ is a bound on the speed of convergence to the uniform distribution. If λ is bounded away from 1 then convergence takes time logarithmic in the number of vertices. The value $1 - \lambda$ is called the *spectral gap* of the graph.

Now suppose $\mathbf{p}$ puts all its probability mass at a single vertex. In order for the distribution to converge to uniform in a logarithmic number of steps we would expect the number of vertices reached by the Markov Chain to grow by a constant fraction at each step. This is guaranteed by the following lemma, which says that if the spectral expansion of a graph is large then so is its edge expansion.

Lemma 2. *The edge expansion of a graph is at least half as large as its spectral gap.*

For every $\lambda > 0$ there exists a sufficiently large d such that for every n there exists a d-regular graph on n vertices with spectral gap at least $1 - \lambda$, and therefore edge expansion at least $(1 - \lambda)/2$. Random graphs that are d-regular have this property with high probability, but there are also "explicit" constructions that produce such graphs in time polynomial in the number of vertices.

3 PCP transformations

The proof of the PCP theorem performs the following sequence of reductions on CSP instances:

reduction	size	$\lvert\Sigma\rvert$	queries	soundness gap	degree	spectral gap
	m	2	$q = C$	δ		
query reduction	$\times C$	2^q	2	$\div C$	large	
degree reduction	$\times C$	2^q	2	$\div C$	C	small
expanderizing	$\times C$	2^q	2	$\div 2$	$d = 2C$	1/4
gap amplification	$\times d^t$	2^{qd^t}	2	$\times t/\lvert\Sigma\rvert^4$		
alphabet size reduction	$\times \exp(2^{qd^t})$	2	q	$\div 2$		

Here size refers to the number of constraints, and degree and spectral gap refer to the underlying constraint graph.

The only quantities that are non-constant throughout this sequence of transformations are the size (initially m) and the soundness gap (initially δ). When the parameters are chosen appropriately (t is a sufficiently large constant in terms of C), the soundness gap is doubled, while the size of the instance increases only by a constant factor, giving the desired conclusion.

3.1 Query reduction

The goal of this transformation is to turn a qCSP Ψ (for some constant q) into a 2CSP Ψ'. Suppose Ψ has n variables $x_1, \ldots, x_n$ and m constraints over alphabet $\{0,1\}$. Ψ' is specified as follows:

- **Variables of Ψ':** The instance Ψ' has variables $x_1, \ldots, x_n, y_1, \ldots, y_m$, where x_i takes values in $\{0,1\}$ and y_i takes values in $\Sigma = \{0,1\}^q$. The intended value of y_i is $x_{i_1} \ldots x_{i_q}$, where $x_{i_1}, \ldots, x_{i_q}$ are the variables participating in the constraint ψ_i of Ψ.
- **Constraint graph of Ψ':** For each constraint ψ_i of Ψ and each variable x_{i_j} that participates in ψ_i, there is a constraint-edge (x_{i_j}, y_i). The corresponding constraint is satisfied if y_i satisfies ψ_i and the jth entry in y_i equals x_{i_j}.

Clearly, if Ψ is satisfiable, so is Ψ' (we just assign $y_i = x_{i_1} \ldots x_{i_q}$ for every i). On the other hand, if every assignment violates a δ-fraction of constraint in Ψ, then every assignment will violate a δ/q-fraction of constraints in Ψ'. To prove this, assume that some assignment (x, y) violates less than a δ/q-fraction of constraints in Ψ'. Since every y_i is involved in q constraints, it means that all constraints involving y_i are satisfied for at least a $1 - \delta$ fraction of y_is. But if all constraints involving y_i are satisfied, it must be that x satisfies ψ_i in Ψ, so x satisfies a $1 - \delta$ fraction of constraints in Ψ.

3.2 Degree reduction

After applying query reduction, the constraint graph may have large degree. The goal of the degree reduction step is to make the degree constant (independent of the instance size), while losing only a constant factor in the soundness gap.

Let G be the constraint graph of Ψ. We create a new CSP Ψ' by replacing every vertex i in G of degree n_i by a cloud of n_i vertices. So each variable x_i of Ψ will give rise to d_i variables $x'_{i1}, \ldots, x'_{in_i}$

in Ψ'. Each constraint in Ψ gives rise to $d/2$ parallel constraints in Ψ' between unique vertices in the corresponding clouds. Within each cloud, we interconnect the vertices by a 1/8-edge expander and make each expander constraint an equality constraint (i.e. requiring that variables get the same value). Notice that if Ψ has m constraints, then Ψ' will have m variables and dm constraints.

Clearly, if Ψ has a satisfying assignment x, we can obtain a satisfying assignment for Ψ' by setting $x'_{i1} = \cdots = x'_{in_i} = x_i$ for every i.

Now we prove soundness. The fact that the soundness gap goes down by at most a constant factor in this transformation is a consequence of the following claim:

Claim 3. *If some assignment x' the violates at most an ε-fraction of constraints in Ψ', then there exists an assignment x that violates at most a 34ε fraction of constraints in Ψ.*

The assignment x is obtained from x' as follows: Within each cloud, let x_i be the plurality value (i.e., the most representative value) among $x'_{i1}, \ldots, x'_{in_i}$. Let ε_i be the fraction of constraints violated in cloud i. Then $\sum_{i=1}^{n} \varepsilon_i \cdot (dn_i/4) \leq \varepsilon \cdot (dm/2)$, the total number of violated constraints.

Let S_i be the set of variables x'_{ij} that agree with the plurality value x_i. Let ε_i be the fraction of the $dn_i/4$ equality constraints for i violated by the assignment x'. We will argue that $|\overline{S_i}| \leq 8\varepsilon_i n_i$:

- If $|S_i| > n_i/2$, then $|E(S_i, \overline{S}_i)| \geq d|\overline{S}_i|/8$. Since all the equality constraints for i between S_i and $\overline{S}_i$ are violated by x', $\varepsilon_i(dn_i/4) \geq |E(S_i, \overline{S}_i)|$, so $|\overline{S_i}| \leq 2\varepsilon_i n_i$.
- If $n_i/4 \leq |S_i| \leq n_i/2$, then $|E(S_i, \overline{S}_i)| \geq d|S_i|/8 \geq dn_i/32$. Since all the equality constraints for i between S_i and $\overline{S}_i$ are violated by x', it follows that $\varepsilon_i \geq 1/8$, so $|\overline{S_i}| \leq n_i \leq 8\varepsilon_i n_i$.
- If $|S_i| < n_i/4$, then no value in Σ is taken by more than a 1/4-fraction of the x'_{ij}s, so there must exist some subset of values $\Sigma' \subseteq \Sigma$ so the number of x'_{ij} taking values in Σ' is between $n_i/4$ and $n_i/2$. Just like in the previous case, we get $|\overline{S_i}| \leq n_i \leq 8\varepsilon_i n_i$.

Now consider what happens in Φ' when we replace the assignment x' with the plurality assignment $x'_{\text{plur } ij} = x_i$ for every j. Replacing x' by x'_{plur} may cause the violation of at most $(d/2)|\overline{S_i}|$ non-equality constraints for every i. If x' violates εdm constraints, x'_{plur} will then violate at most

$$\varepsilon dm + \sum_{i=1}^{n} (d/2)|\overline{S_i}| \leq \varepsilon dm + \sum_{i=1}^{n} (d/2)(8\varepsilon_i n_i) = \varepsilon dm + 16 \sum_{i=1}^{n} \varepsilon_i dn_i/4 \leq 17\varepsilon dm$$

constraints of Ψ'. This is a 17ε-fraction of all the constraints in Ψ'. Since exactly half the constraints in Ψ' are equality constraints, x cannot violate more than a 34ε fraction of constraints in Ψ.

3.3 Expanderizing

The expanderizing transformation starts with a CSP Ψ with two variables per constraint and (sufficiently large) constant degree d and creates a new CSP Ψ' with two variables per constraint, degree $2d$, and the property that the constraint graph is an expander with spectral gap at least 1/4.

Suppose the constraint graph G of Ψ has n vertices. Let Z be an expander on n vertices with degree d and $\lambda \leq 1/2$. The variables of Ψ' are be the same as the variables of Ψ. The constraints of Ψ' include all the constraints of Ψ. In addition, for every edge in Z, we add a "dummy" constraint in Ψ' that is satisfied for any assignment to its variables.

Clearly if Ψ is satisfiable, then Ψ' is satisfiable by the same assignment. On the other hand, if x fails to satisfy a δ fraction of the constraints in Ψ, then it will fail to satisfy the same constraints in Ψ', which form a $\delta/2$-fraction of all the constraints.

Let G' be the constraint graph of Ψ'. We now show that if G and Z are regular graph of the same degree, then $\lambda_{G'} \leq (\lambda_G + \lambda_Z)/2 \leq 3/4$. Notice that the adjacency matrices satisfy the relation $A_{G'} = (A_G + A_Z)/2$. Then for every $\mathbf{v} \perp \mathbf{u}$, we have

$$\|\mathbf{v}A_{G'}\| \leq \tfrac{1}{2}\big(\|\mathbf{v}A_G\| + \|\mathbf{v}A_Z\|\big) \leq \tfrac{1}{2}\big(\lambda_G\|\mathbf{v}\| + \lambda_Z\|\mathbf{v}\|\big) = \tfrac{1}{2}(\lambda_G + \lambda_Z)\|\mathbf{v}\|.$$

4 Gap amplification

Fix a constant t. The gap amplification step is a transformation from a 2CSP Ψ with degree d and $\lambda = 3/4$ to a 2CSP Ψ' with the following parameters:

	Ψ	$\to \Psi'$
size	m	$\to (\|\Sigma\|d)^{5t}m$
alphabet	Σ	$\to \Sigma^{1+d+d^2+\cdots+d^t}$
completeness	Psi is satsifiable	$\to \Psi'$ is satisfiable
soundness	$1-\delta$	$\to 1 - \Omega(t\delta/\|\Sigma\|^4)$

Let G be the constraint graph of Ψ. We now define the instance Ψ'. We specify the constraints as a probability distribution with the interpretation that each constraint is included a number of times proportional to its probability.

- **Variables of Ψ':** For each variable x_v of Ψ, there is a corresponding variable x'_v of Ψ'.
- **Values of x'_v:** The value of x'_v is a collection (tuple) of values in Σ, one corresponding to every vertex u at distance $\leq t$ from v in G. We write $x'_v(u)$ for the component of x'_v corresponding to u.
- **Distribution over constraints of Ψ':** The constraints ψ'_p of Ψ' correspond to paths p of length at most $5t\ln|\Sigma|$ in G. (We will identify constraints and the paths they represent.) The paths are generated from the following distribution:
  1. Choose a starting vertex v_0 of p. Set $i = 0$
  2. Repeat for at most $5t\ln|\Sigma|$ times: (1) Set v_{i+1} to be a random neighbor of v_i and increment i (2) With probability $1/t$, stop the repetition.
  3. Output the path $v_0, v_1, \ldots, v_i$.
- **Constraints of Ψ':** Let (u', v') be the endpoints of a path p. The constraint $\psi'_p(x'_{u'}, x'_{v'})$ is satisfied if all of the following hold:
  1. For every edge (u, v) in G such that u and v are both within distance t of u', the constraint $\psi_{(u,v)}$ is satisfied.
  2. For every edge (u, v) in G such that u and v are both within distance t of v', the constraint $\psi_{(u,v)}$ is satisfied.
  3. For every vertex v that is within distance t from both u' and v', $x'_{u'}(v) = x'_{v'}(v)$.

The size and alphabet size of Ψ' are easy to check. We need to argue completeness and soundness. By design, the transformation has perfect completeness. Suppose x is a satisfying assignment of Ψ. Now consider the assignment x' of Ψ' given by $x'_u(v) = x_v$. This satisfies all the constraints of Ψ'.

The (relatively) difficult part is to argue soundness. To do this, we must show that for every x' that satisfies $1 - \Omega_\Sigma(t\delta)$ constraints of Ψ', there is an x that satisfies $1 - \delta$ constraints of Ψ.

The assignment x is constructed from x' via the following procedure. For every vertex v,

1. Define the following distribution D_v on vertices. Initially, set $v' = v$. Now repeat the following experiment: With probability $1/t$ stop, and with the remaining probability, set $v' =$ a random neighbor of v'.
2. Set x_v to equal the plurality value of $x'_{v'}(v)$, when v' is chosen from D_v, among those v' that are within distance t of v.

We now need to argue that if x' satisfies $1 - \Omega(t\delta/|\Sigma|^4)$ constraints of Ψ', then x satisfies $1 - \delta$ constraints of Ψ. In fact, we will argue the contrapositive:

Claim 4. *Assume $t\delta < 1$. If x violates δ constraints of Ψ, then x' violates $\Omega(t\delta/|\Sigma|^4)$ constraints of Ψ'.*

Before we prove the claim, let us make one simplification. We will modify the distribution over constraints of Ψ' so that the path p is not truncated after $5t \ln|\Sigma|$ steps (see step 2), but can be of any length. Intuitively, this simplification should not make a difference because long paths are unlikely. Formally, we will analyze the effect of this simplification later.

Now let's explain the intuition behind this claim. Let F be the set of constraints of Ψ (which we also think of as edges of G) that are violated by x (so $|F| = \delta m$). Now take a random constraint ψ' of Ψ'. What are the chances that this constraint is violated by x'?

$$\Pr[x' \text{ violates } \psi'] \geq \Pr[\psi' \text{ intersects } F] \cdot \Pr[x' \text{ violates } \psi' \mid \psi' \text{ intersects } F].$$

Let's try to estimate both of these quantities. We expect ψ' to have about t edges; since $|F| = \delta m$, we expect ψ' to contain about δt edges of F. Since δ is fairly small, we might expect that most ψ' which intersect F intersect only a single edge of F. If this is the case, then $\Pr[\psi' \text{ intersects } F]$ should be about $t\delta$. This is where the soundness amplification happens.

What about the other probability? Let's now fix an edge $(u, v) \in F$ that is contained in ψ'. Now consider the distribution of the endpoints u' and v' of the path ψ'. Since the endpoints of the path are determined by a Poisson process, it follows that conditioned on (u, v) being in ψ', the endpoint v' is determined by the following distribution: Start from v and at each step (1) with probability $1/t$ stop and (2) with the remaining probability move to a random neighbor of v and continue. This is exactly the distribution D_v. Ignoring for now the fact that the path could be too long, we reason as follows. Since the value x_v was defined as the plurality value $x_{v'}(v)$, the two should match with probability at least $1/|\Sigma|$. For the same reason, $x_{u'}(u)$ and x_u should match with probability $1/|\Sigma|$. But since the constraint $\psi(x_u, x_v)$ is violated, $\psi'(x'_{u'}, x'_{v'})$ is then also violated.

To summarize, we expect that the probability that a random constraint of Ψ' is violated is about $t\delta/|\Sigma|^2$. In our estimates we made two overly simplifying assumptions. The actual analysis will have to address the following additional possibilities:

- What happens when ψ' intersect multiple edges of F?

- What happens when ψ' contains more than t edges? In this case, it may happen that ψ' contains a "bad" edge, but this edge cannot be "seen" from its endpoints.

The analysis will show that both events (1) and (2) happen rarely: Event (1) owing to the expansion of G and event (2) owing to the small probability assigned to long paths.

4.1 Analysis of gap amplification

Now let us do the actual analysis. Call an edge (u, v) *faulty* (with respect to ψ', x', x) if (1) $(u, v) \in F$, (2) $d(u', u), d(v, v') < t$, and (3) $x'_{u'}(u) = x_u$ and $x'_{v'}(v) = x_v$, where u', v' are the endpoints of ψ'. If some edge in ψ' is faulty, then ψ' is violated as the inconsistency between x_u and x_v can be seen either by $x'_{u'}$ or by $x'_{v'}$.

Let N denote the number of faulty edges of ψ', where ψ' is chosen at random. By the Paley-Zygmund inequality,

$$\Pr[\psi' \text{ is violated}] \geq \Pr[N > 0] \geq \mathrm{E}[N]^2 / \mathrm{E}[N^2]. \tag{1}$$

The first moment. We first estimate $\mathrm{E}[N]$. For $f \in F$, let I_f denote the number of occurrences of f in ψ', and let $N_f = I_f$ is f is faulty, and 0 otherwise. Then:

$$\mathrm{E}[N] = \sum_{f \in F} \mathrm{E}[N_f] = \sum_{f \in F} \sum_{k=1}^{\infty} \Pr[N_f \geq k] = \sum_{f \in F} \sum_{k=1}^{\infty} k \cdot \Pr[I_f \geq k] \cdot \Pr[f \text{ is faulty} \mid I_f \geq k].$$

Let us analyze the probability that f is faulty conditioned on $I_f \geq k > 0$. Fix an arbitrary collection of k occurrences of f in ψ and let u be the left endpoint of the first occurrence and v be the right endpoint of the last occurrence. As discussed above, u' follows the distribution D_u, and v' independently follows the distribution D_v. In this distribution, the probability that u' is at distance more than t from u is $\leq (1-1/t)^t < 1/2$. Conditioned on this distance being at most t, the distribution on u' is exactly the one used to define the plurality assignment x_u, so the probability that $x'_{u'}(u) = x_u$ is at least $1/|\Sigma|$. As the same is true for v and v' independently, for any $k > 0$

$$\Pr[f \text{ is faulty} \mid I_f \geq k] \geq \left(\frac{1}{2} \cdot \frac{1}{|\Sigma|}\right)^2$$

and therefore

$$\mathrm{E}[N] \geq \frac{1}{4|\Sigma|^2} \cdot \sum_{f \in F} \sum_{k=1}^{\infty} \Pr[I_f \geq k] = \frac{1}{4|\Sigma|^2} \cdot \sum_{f \in F} \mathrm{E}[I_f] = \frac{\delta t}{4|\Sigma|^2},$$

because the expected number of occurrences of any particular edge in ψ' is $1/m$ times the expected length of ψ', which is t.

The second moment. We now upper bound $\mathrm{E}[N^2]$. To do so, let N' be the number of edges in F that intersect ψ'. Obviously $N \leq N'$ (since N counts the number of such edges that are also faulty). So we will bound $\mathrm{E}[N'^2]$ instead. To do so, let Z_i be a random variable that indicates if the ith edge of ψ' is in F (if ψ' has fewer than i edges, then $Z_i = 0$). Then

$$\mathrm{E}[N'^2] = \sum_{i=1}^{\infty} \mathrm{E}[Z_i] + 2 \sum_{1 \leq i < j} \mathrm{E}[Z_i Z_j]. \tag{2}$$

It is easily seen that $\mathrm{E}[Z_i] = \delta \cdot (1 - 1/t)^i$, so the first summation is at most $t\delta$.

For the second summation, notice that $\mathrm{E}[Z_i Z_j]$ is the probability that both edges i and j are present in the path and faulty. The probability they are both present is $(1 - 1/t)^j$. Conditioned on them being both present, the probability they are both faulty is bounded using the following lemma.

Lemma 5. *Let G be a d-regular graph with spectral gap $1-\lambda$ and F be a subset consisting of a δ fraction of the edges of G. The probability that both the first and the last edge of a random walk of G of length $\ell \geq 2$ are in F is at most $\delta^2 + \delta\lambda^{\ell-2}$.*

It follows that $\mathrm{E}[Z_i Z_j] \leq (1-1/t)^j \cdot \delta \cdot (\delta + \lambda^{j-i-1})$. Plugging this in (2) we have

$$\begin{aligned}
\mathrm{E}[N'^2] &\leq \delta t + 2\delta \sum_{1\leq i<j} \mathrm{E}[Z_i Z_j] \\
&\leq \delta t + 2\delta \sum_{i=1}^{\infty} (1-1/t)^i \sum_{j=i+1}^{\infty} (1-1/t)^{j-i} \cdot (\delta + \lambda^{j-i-1}) \\
&\leq \delta t + 2\delta \sum_{i=1}^{\infty} (1-1/t)^i (\delta t + 1/(1-\lambda)) \\
&\leq \delta t + 2\delta t(\delta t + 4) \\
&= 9\delta t + 2(\delta t)^2.
\end{aligned}$$

Second moment calculation. Finally, from (1) we have:

$$\Pr[N > 0] \geq \frac{\mathrm{E}[N]^2}{\mathrm{E}[N^2]} \geq \frac{(\delta t/4|\Sigma|)^2}{9\delta t + 2(\delta t)^2} = \Omega(\delta t/|\Sigma|^4).$$

The effect of truncation. This calculation was done in the idealized setting where ψ' can be arbitrarily long, while it is actually restricted to have length at most $5t\ln|\Sigma|$. It is not hard to see that these long paths contribute little to N. In particular, the contribution from the long paths can be bounded by

$$\sum_{\ell=5t\ln|\Sigma|}^{\infty} \mathrm{E}[N \mid \psi' \text{ has length } \ell] \Pr[\psi' \text{ has length } \ell] \leq \sum_{\ell=5t\ln|\Sigma|}^{\infty} (\delta\ell) \cdot (1-1/t)^{\ell} < \mathrm{E}[N]/2$$

For the calculation of $\mathrm{E}[N^2]$, the truncation of long paths only improves this quantity, so the lower bound on the probability that $N > 0$ is only affected by a constant.

Proof of Lemma 5. Let A be the (normalized) adjacency matrix of G and A' be the adjacency matrix of a graph representing $\ell - 2$ steps of a random walk on G. Then $A' = A^{\ell-2}$ and $\lambda' = \lambda^{\ell-2}$.

If we write $A' = (1-\lambda')J + E$ then for every vector $\mathbf{v}$, $\|\mathbf{v}E\| \leq \lambda'\|\mathbf{v}\|$. Here J is the adjacency matrix of the complete graph on n vertices with self-loops. Now we write

$$\frac{1}{2n}|\mathbf{v}A'\mathbf{v}^{\mathrm{T}}| \leq (1-\lambda')\frac{1}{2n}|\mathbf{v}J\mathbf{v}^{\mathrm{T}}| + \frac{1}{2n}|\mathbf{v}E\mathbf{v}^{\mathrm{T}}|.$$

Let $\mathbf{v}$ be the vector such that $\mathbf{v}(u)$ equals the fraction of edges incident to u that are in F. Then $(\mathbf{v}A'\mathbf{v}^{\mathrm{T}})/2n$ equals exactly the fraction of paths with the first and last edge in F, and $(\mathbf{v}J\mathbf{v}^{\mathrm{T}})/2n$ equals $\mathrm{E}_u[\mathbf{v}(u)]^2 = \delta^2$. For the last term we have

$$\frac{1}{2n}|\mathbf{v}E\mathbf{v}^{\mathrm{T}}| \leq \frac{1}{2n}\|\mathbf{v}E\| \cdot \|\mathbf{v}\| \leq \lambda'\frac{1}{2n}\|\mathbf{v}\|^2 \leq \lambda'\frac{1}{2n}\sum_u \mathbf{v}(u) = \delta\lambda',$$

so the desired quantity is at most $(1-\lambda')\delta^2 + \delta\lambda' \leq \delta^2 + \delta\lambda^{\ell-2}$. □

5 Alphabet size reduction

The purpose of alphabet size reduction is to transform a 2-query PCP Ψ over large (but constant) alphabet Σ into a q-query PCP Ψ' over alphabet $\{0,1\}$, where q is independent of the size of Σ. We want to preserve completeness and lose only a constant factor (independent of Σ) in the soundness gap. On the other hand, the size of the instance is allowed to increase by a constant factor, which may depend on Σ.

Without loss of generality, we can think of Σ as $\{0,1\}^\sigma$ for some constant c. Then we can think of every variable y_i of Ψ as taking values in $\{0,1\}^\sigma$ and we can view every constraint $\psi(y_i, y_j)$ of Ψ as a function from $\{0,1\}^{2\sigma}$ to $\{0,1\}$.

From the proof verification perspective, Ψ is described by the following proof system. The proof is a string of length $\{0,1\}^{\sigma n}$ which for every variable y_i of Ψ contains all the bits of y_i. The verifier chooses a random constraint $\psi(y_i, y_j)$, reads the bits of y_i and y_j, and accepts if the constraint is satisfied. The query complexity of this PCP is 2σ; we would like it to be a constant independent of σ.

To achieve this effect we apply the PCP from the last lecture to each one of the constraints ψ using shared encodings to their variables. The PCP then certifies that a random constraint is satisfied. Since the encodings to the variables are shared among the constraints, they must be consistent with a single assignment. The randomness complexity of the "inner PCP" is polynomial in the size of the assignment to ψ, which is 2σ bits long, so it is just a constant depending on $|\Sigma|$.

More formally, to implement the PCP construction from last lecture, we want to transform each constraint $\psi(y_i, y_j)$ into an equivalent system of quadratic equations Q. Recall that the system Q will have at most $O(2^{2\sigma})$) equations, which in addition to the variables y_i and y_j involve $O(2^{2\sigma})$) auxiliary variables z_{ij}.

The proof π in the PCP Ψ' will now consist of two parts:

1. For each y_i taking values in $\{0,1\}^\sigma$, an encoding $C_i \in \{0,1\}^{2^\sigma}$ which is supposed to equal $C_i(a) = \langle a, y_i \rangle$ for every $a \in \{0,1\}^\sigma$.

2. For every constraint $\psi(y_i, y_j)$ consider the corresponding quadratic system $Q(y_i, y_j, z_{ij})$ where z_{ij} takes values in $\{0,1\}^{O(2^{2\sigma})}$. Provide an encoding C_{ij} for z_{ij} where $C_{ij}(a)$ is supposed to equal $\langle a, z_{ij} \rangle$, as well as an encoding D_{ij} of Q, where for every linear combination b of quadratic terms in the variables y_i, y_j, z_{ij}, $D_{ij}(b)$ is supposed to equal the value of this combination.

The verifier of Ψ' chooses a random constraint $\psi(y_i, y_j)$ in Ψ and runs the PCP from last lecture on the part of the proof that contains the encodings C_i, C_j, C_{ij}, D_{ij} to verify that the constraint is satisfied.[1]

Clearly if Ψ is satisfiable, the verifier of Ψ' will accept with probability 1. Now we argue that if Ψ' rejects with probability at most $\delta/2$, then some sassignment violates at most a δ-fraction of constraints in Ψ.

Assume Ψ' rejects with probability at most $\delta/2$. Let y_i be the most likely assignment encoded by C_i (i.e. the one such that the encoding of y_i and C_i differ in the smallest number of places, breaking ties arbitrarily). Then for at least a $1-\delta$ fraction of the constraints ψ, when ψ is chosen Ψ' accepts with probability at least $1/2$. By the analysis from last time if this is the case, then all of C_i, C_j

[1] This is not exactly the same PCP. In the last lecture C_i, C_j, and C_{ij} were grouped into a single chunk, while here they are separate. However we can run the linearity test and local decoding procedures on each part separately with the same effect.

and C_{ij} must be 1/8-close to encodings of some assignments y'_i, y'_j and z_{ij} so that $Q(y'_i, y'_j, z_{ij})$ is satisfied and therefore $\psi(y'_i, y'_j) = 1$. Since y_i is the most likely assignment encoded by C_i, it must be that the encodings of y_i and y'_i differ in at most a 1/4-fraction of places. But any two distinct linear functions differ on at least half the outputs, so it must be that $y_i = y'_i$. Similarly $y_j = y'_j$. Therefore y satisfies the constraint ψ, so it satisfies a $1 - \delta$ fraction of constraints of Ψ.

References

This proof of the PCP theorem is from Irit Dinur's *The PCP Theorem by Gap Amplification* with some simplifications by Jaikumar Radhakrishnan and others. Lemma 2 is a slightly weaker version of the easy direction of Cheeger's inequality. Its proof is not difficult and can be found in many places including the survey *Expander graphs and their applications* by Hoory, Linial, and Wigderson. That survey also describes several constructions of expander graphs.