

Gap Amplification

Instructor: Manindra Agrawal

Scribe: Ramprasad Saptharishi

Last class we concluded by showing that an instance of G_k can be converted into another instance of G_k with just a constant factor loss in the gap but making the underlying graph a constant degree expander. This class we shall look at the third step, which is Gap Amplification.

1 Gap Amplification

Instead of directly constructing the reduction, we shall look at it with a different perspective. We shall think of the constraints as being checked by a probabilistic verifier. The verifier will perform certain structured random walks to decide whether or not a certain edge is to be accepted. We shall then see that this can be converted to a new graph where each walk would correspond to an edge.

The instance G of G_k is a (n, d, λ) expander. Take any vertex $v \in G$, the number of vertices reachable from v within t steps is at most d^{t+1} since the degree of each vertex is d . We shall increase the alphabet size to make every vertex hold the colour of all vertices that are reachable within t steps from it. The alphabet size increases from $[1..k]$ to $[1..k]^{d^{t+1}}$ where the colours of all vertices reachable within t steps (including itself) is present in its colour. The number of vertices reachable within t steps could be very well under d^{t+1} in which case the superfluous colours are arbitrary.

One thing to note is that a vertex v could be reachable from two vertices u_1 and u_2 within t steps. Both these vertices might assign a certain colour to v and in general they may not coincide. We shall refer to the colour that u_1 assigns to v as u_1 's opinion of v .

Now for the acceptance criterion. We shall say whether a certain random walk is accepted or not based on a certain assignment; this shall be translated to whether a certain edge is accepted or not in the modified graph as each random would correspond to an edge in the new graph. The verifier performs a random walk as follows:

1. Pick a random vertex $a \in V$.

2. Pick a random neighbour of a .
3. With probability $1/t$, stop.
4. Move to the neighbour and go to step 2.

The verifier performs such a random walk. If the length of the walk is larger than ct for a constant c that shall be specified later, then he accepts.

Let a_1, a_2, \dots, a_l be the vertices in the walk. The verifier reads the colours of a_1 and a_l . If a_1 and a_l gives different colours to any a_i , the verifier rejects. And for any intermediate edge (a_i, a_{i+1}) , if the opinions of the two end points of the walk violates the constraint in the original graph, the verifier rejects. If none of the above happened, then he accepts.

We need to argue that the verifier will now see a reasonably large gap. For this purpose, we shall estimate the probability that the verifier rejects. As usual, we shall pick a colouring assignment to this new instance and construct a colouring assignment of the old graph G . This is done by a process similar to the most popular colour, namely the most popular opinion. However, rather than doing a naive majority, we shall take a weighted average with the probabilities. To formalize this notion, we shall define the colour of a vertex u as follows:

1. Start at the vertex u .
2. Pick a random neighbour
3. With probability $1/t$, stop.
4. If length walk equal to t , stop.
5. Move to the neighbour and go to step 2.

The final vertex now has an associated probability of reaching it. The most popular colour is now defined as the most probably opinion.

Let F be the set of violated edges by this colour assignment. What we know is that $|F| \geq \text{gap}|E|$. The first thing we need to do is estimate the number of occurrences of a certain edge e in a random walk.

The expected length of the random walk is t and the probability that e occurs as the first edge is $1/|E|$ since the initial vertex is chosen at random by the verifier. Let X_i was the indicator random variable that takes 1 if the

i -th edge is e and 0 otherwise

$$\begin{aligned} \Pr[e \text{ is the } i\text{-th edge in the walk}] &= \frac{1}{|E|} \\ \implies \mathbf{E}[X_i] &= \frac{1}{|E|} \\ \implies \mathbf{E}\left[\sum X_i\right] &= \frac{\mathbf{E}[\text{length of the walk}]}{|E|} \end{aligned}$$

and hence the expected number of occurrences of an edge e in the walk is $\frac{t}{|E|}$. We shall use N to denote the length of the walk, N_F to denote the number of times a violated edge occurs in walk. We want to estimate N_F . Let $N_{F,e}$ denote the random variables that counts the number of times the edge e , whose constraint should be violated, appears in the walk. Again, by linearity of expectation:

$$\mathbf{E}[N_F] = \mathbf{E}\left[\sum N_{F,e}\right]$$

And by the definition of expectation

$$\begin{aligned} \mathbf{E}[N_{F,e}] &= \sum_{r=1}^{\infty} r \Pr[e \text{ is violated and number of times } e \text{ occurs in walk is } r] \\ &= \sum_{r=1}^{\infty} r \Pr[e \text{ is violated} \mid \text{occurs } r \text{ times}] \Pr[e \text{ occurs } r \text{ times}] \end{aligned}$$

Firstly, the number of times e occurs in the walk is irrelevant to whether or not e is violated. Therefore,

$$\mathbf{E}[N_{F,e}] = \sum_{r=1}^{\infty} r \Pr[e \text{ is violated}] \Pr[e \text{ occurs } r \text{ times}]$$

(I'm not too clear from this point on, needs some cleaning up. Lack of clarity begins)

Now to estimate the probability that an edge e is violated, let $e = (u, v)$. Now look at the walk up till u , reversing this walk gives us a random walk just as before. So one could think of it as a walk being performed starting at u and similarly one starting at v . The probability that these walks have length more than t is $(1 - \frac{1}{t})^t$ which is certainly less than half. Therefore, the probability that the vertex u gets the most popular colour is at least $1/2k$ (why?). And similarly, for vertex v as well.

The next thing to note is that both the vertices u and v are independently assigned colours. Therefore, the probability that the constraint (u, v) is violated is at least $\frac{1}{4k^2}$.

(Lack of clarity ends)

Therefore,

$$\begin{aligned}
 \mathbf{E}[N_{F,e}] &= \sum_{r=1}^{\infty} r \Pr[e \text{ is violated}] \Pr[e \text{ occurs } r \text{ times}] \\
 &\geq \frac{1}{4k^2} \sum_{r=1}^{\infty} r \Pr[e \text{ occurs } r \text{ times}] \\
 &= \frac{1}{4k^2} \mathbf{E}[\text{number of times } e \text{ occurs in the walk}] \\
 &= \frac{1}{4k^2} \frac{t}{|E|} \\
 \implies \mathbf{E}[N_F] &\geq \frac{t}{4k^2} \frac{|F|}{|E|}
 \end{aligned}$$

Thus, in a sense, the verifier sees a new gap of about $\frac{t}{4k^2}$ on expectation. To argue the probability bounds, we need a handle on the variance as well. We shall do that in the next class.