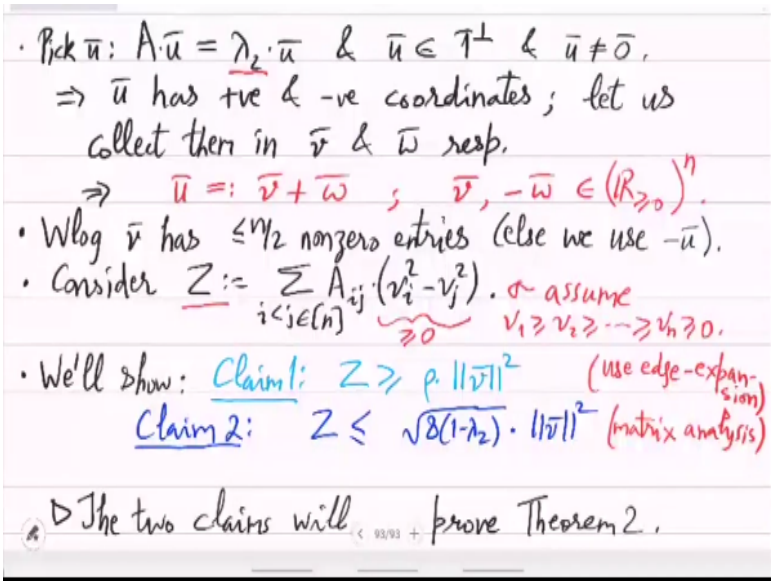
Randomized Methods in Complexity
Prof. Nitin Saxena
Department of Computer Science & Engineering
Indian Institute of Technology-Kanpur

Lecture - 12
Relation between Algebraic and Combinatorial Expanders

Last time we showed the equivalence of, we showed that an algebraic expander is also an edge expander with this rho factor in the expansion being $1 - \lambda/2$

(Refer Slide Time: 00:31)



So now we will show an almost converse statement. So you start with a rho edge expander. Now why is it an algebraic expander? In other words, why is the spectral gap large? So we started the proof. Assume that G is an n, d, rho edge expander and we will again use this kind of this Laplacian quadratic form, right? So you start with eigenvector of the second largest eigenvalue of the matrix A lambda 2, which is this equation.

So $\bar{u}\in\bar{1}\perp$, which is the vector having coordinates 1 /n in each place. Nonzero, since it is orthogonal to 1 there has to be positive negative coordinates and not all can be 0. So let us collect positives in v, negatives in w. So that both v and -w are non-negative coordinates only. And then based on that, so what you can say is without loss of generality $\bar{v}$ has $\leq$ n /2 nonzero entries because, if it is more than n / 2 then you can look at –u right?

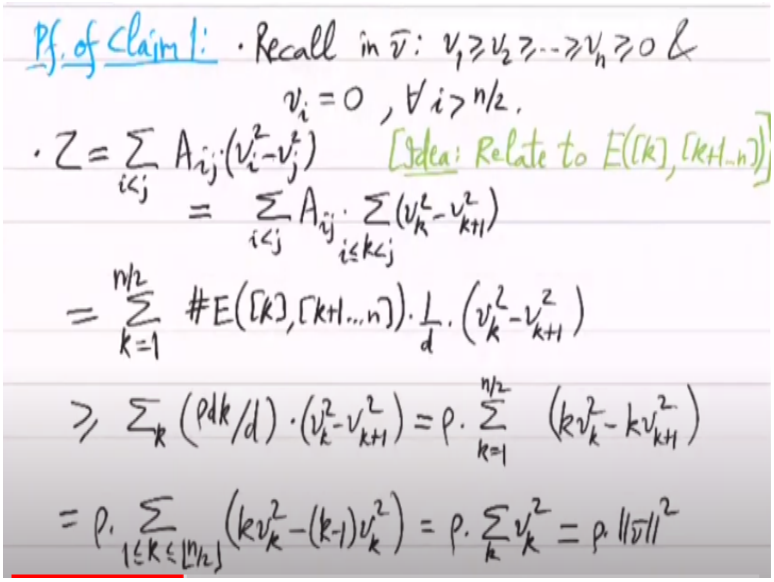
So if v has more than n /2 nonzero entries, then w will have less than n /2. So you can use w. So for that you can work with –u. So consider again this quadratic form $Z: = \sum_{i<j\epsilon[n]} A_{i,j}(v_i^2 - v_j^2)$ So that is the definition of Z. Notice that this is slightly different from what you used in the proof of theorem 1, right?

There we had $(x_i - x_j)^2$ and here we have difference of square. And using the edge expansion we want to get information about λ_2 or estimate λ_2. So we will show that and here also assume them to be ordered. So $v_1 \geq v_2 \geq \ldots\ldots\ldots \geq v_n \geq 0$, okay. So v_1 is the largest coordinate. And then you are looking at these successive differences or all possible pairs actually $v_i^2 - v_j^2$

But since we are looking at i <j, so this is each of these is non-negative, okay. So z is a non-negative number. So we will show that Z is large, namely $Z \geq \rho||\bar{v}||^2$. So here we will use the edge expansion and we will show that Z is small$Z \leq \sqrt{8(1-\lambda_2)}.||\bar{v}||^2$. So these two claims together will give you the information about λ_2, okay using given ρ.

So the first claim we will show using edge expansion because it is about ρ, so we will use edge expansion here. And the second claim we will here we will use some matrix analysis and then these two claims will give you the proof of theorem 2, will prove theorem 2, okay. So you will get the best possible information about how λ_2 and ρare related. So let us first use edge expansion and proof claim 1.

(Refer Slide Time: 06:31)



So remember that we have already sorted the coordinates of v bar and we have assumed that v bar has less than equal to n by 2 nonzero coordinates, right. So recall that in $\bar{v}$:$v_1 \geq v_2 \geq \ldots\ldots\ldots \geq v_n \geq 0$and $v_i = 0$ for all i$\geq n/2$. So only the first n /2 or less are nonzero, in fact positive, positive reals in order. So you can write $Z := \sum_{i<j} A_{i,j}(v_i^2 - v_j^2)$as, the nice thing about this expression is this will slowly connect to number of edges going out of the subset of vertices 1 to k.

So we want to relate this expression to the edges that go out of vertices 1 to k, k + 1 to n. And once you have that relationship then from there you can use the ρ edge expansion, right? So what we will do is we will rewrite this expression by looking at consecutive differences. So let us use telescopic sum $\sum_{i<j} A_{i,j} \sum_{i \leq k < j} v_k^2 - v_{k+1}^2$. And now flip the sum. There are two summations, so flip them. So bring k on the top.

So remember that we only need to go up to k = n /2 because the remaining v's are 0. And for fixed K, what are the ij's that are of interest, right? So the ij's of interest are the crossing edges because only then is $A_{i,j}$ nonzero. So you will get number of edges that cross divided by d

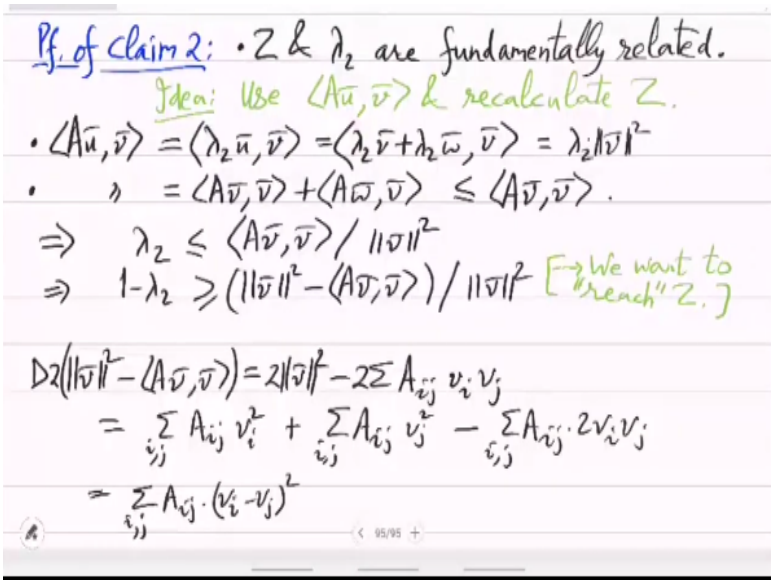
because A $_{i,j}$gives you 1 / d. And difference of square remains the same, right? $\sum_{k=1}^{n/2} \#E([k],[k+1 \ldots . n]).1/d.(v_k^2 - v_{k+1}^2)$That is what you get when you flip the sum.

So the second outer sum when it goes inside it gives you this number of crossing edges. Now you use the edge expansion. So this is $\geq \Sigma_k(\rho dk/d).(v_k^2 - v_{k+1}^2)$. So this will be then $\rho \sum_{k\epsilon[n/2]} (kv_k^2 - kv_{k+1}^2)$ right?

That is the simple expression. So what is this sum? So this sum is equal to we again reduce this to some telescopic sum. Let me not use this notation. So k is 1 to n/ 2, right? So you will get $\rho \sum_{1\leq k\leq[n/2]} (kv_k^2 - (k-1)v_k^2$. Can I do that? So you can do that because the second term, So you are just shifting it back.

So you are replacing k by k - 1, right? So you can see that on one side k= 0, this term will be 0, summand will be 0. And on the other side for n/2 + 1 this summand will be 0. So you can actually shift sum back. So you get this expression and this expression is just$\sum v_k^2$. And what is $\sum v_k^2$? This is $||\bar{v}||^2$, right? So that is the proof of claim 1 that Z is large. It is larger than $\rho||\bar{v}||^2$.

(Refer Slide Time: 13:27)



Now proof of claim 2. So claim 2 was now you have to show that Z is small in terms of λ_2, right? As a function of λ_2 it is small. So we have to now look at Z in a different way by matrix action, do some matrix analysis. So Z andλ_2 are fundamentally related, okay. And how? So the idea here will be use this inner product $< A\bar{u}, \bar{v} >$. Then look at the action of A on $\bar{u}$ because that is what will give you λ_2 , right?

So look at the action of A on $\bar{u}$ and take the inner product with the positive part v bar. So what does that give you? So use this and recalculate Z. So < A $\bar{u}$,$\bar{v}$ > = < λ_2 $\bar{u}$,$\bar{v}$>, which is equal to< $\lambda_2\bar{v} + \lambda_2$ $\bar{w}$,$\bar{v}$ > right? Express u as v + w. And now notice that v and w orthogonal because they have disjoint support, right?

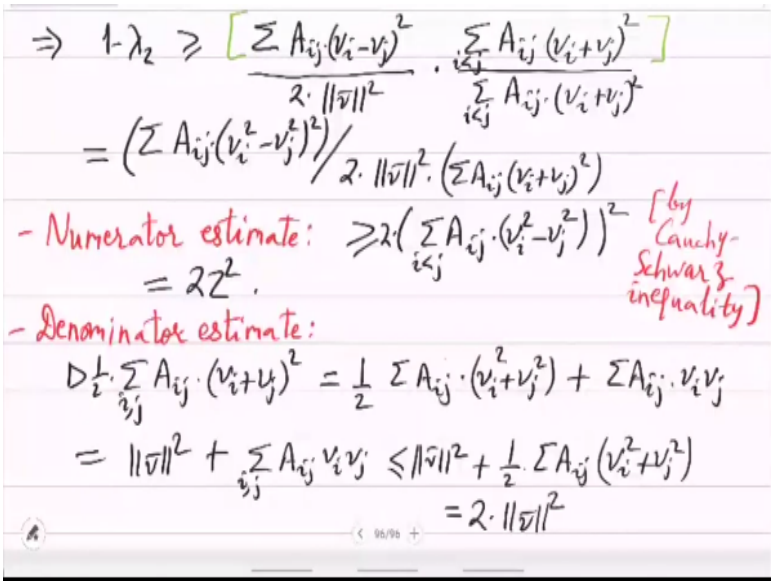
v is the positive places w is in negative places, so the inner product is 0. So you will get $\lambda_2||\bar{v}||^2$. Also this $< A\bar{u}, \bar{v} >$this you can write as$< A\bar{v}, \bar{v} >+ < A\bar{w}, \bar{v} >$.again because u is v + w. Now note that A has nonnegative entries, it is a matrix. v has nonnegative entries, it is a vector. But $\bar{w}$ has negative entries, right? So this sum will be less than equal to the first summand, right?

So we have these two things. So which means that $\lambda_2 \leq < A\bar{v}, \bar{v} > / ||\bar{v}||^2$, which means that 1 - $\lambda_2 \geq (||\bar{v}||^2 - < A\bar{v}, \bar{v} >) / ||\bar{v}||^2$. Remember that our goal is to reach Z right and Z had $v_i^2 - v_j^2$. So that is where we are trying to reach. So let us continue doing this. So we will do these calculations as follows.

This $(||\bar{v}||^2 - < A\bar{v}, \bar{v} >) = ||\bar{v}||^2 - \sum_{i,j} A\ v_i v_j$ And let us throw in a 2 here, double that$(||\bar{v}||^2 - < A\bar{v}, \bar{v} >) = 2||\bar{v}||^2 - 2 \sum_{i,j} A\ v_i v_j$.. So you have seen before that $||\bar{v}||^2$ can also be written as $\sum_{i,j} A_{i,j} v_i^2 + \sum_{i,j} A_{i,j} v_j^2 - \sum_{i,j} A\ 2 v_i v_j$.

So now you can see what we will do. We will write this as $(v_i - v_j)^2$ okay. So this is equal to $\sum_{i,j} A_{i,j} (v_i - v_j)^2$ right?. So let us continue this.

(Refer Slide Time: 19:38)



So this means what we have done is we have written $1-\lambda_2 \geq \frac{\sum_{i,j} A_{i,j}(v_i - v_j)^2}{2||\bar{v}||^2}$ right? There was this $||\bar{v}||^2$ in the denominator. But we are still we have still not reached $v_i^2 - v_j^2$, right? So this is still the Laplacian, it is not this modified Z that we have defined. So for that we have to go to $v_i^2 - v_j^2$.

So for to achieve to reach there you multiply this with $\sum A_{i,j}(v_i + v_j)^2$. You multiply with this, both numerator and denominator. So when you do this you will get

$1-\lambda_2 \geq \frac{\sum_{i,j} A_{i,j}(v_i - v_j)^2 \quad \sum A_{i,j}(v_i + v_j)^2}{2||\bar{v}||^2 \qquad\qquad \sum A_{i,j}(v_i + v_j)^2} \sum A_{i,j}(v_i^2 - v_j^2)^2 / 2||\bar{v}||^2 \sum A_{i,j}(v_i + v_j)^2$ That is the big expression you get.

So we have reached almost $v_i^2 - v_j^2$, but it is still not clear whether it is Z, this expression is Z. So let us estimate the two expressions separately. So first let us look at the numerator. So the numerator $\geq \sum_{i<j} A_{i,j}(v_i^2 - v_j^2)$. So the numerator is the product of these two things. So we are looking at this part. That is the numerator.

So if you use Cauchy-Schwarz inequality okay then the product of these two sums is at least the sum of individual products or I should say well I should, no not that but I should write $(\sum_{i<j} A_{i,j}(v_i^2 - v_j^2))^2$ and I should also specify what I am multiplying with. So here I actually multiplied only those summands for which i <j, okay.

And note that in this first $\sum_{i,j} A_{i,j}(v_i - v_j)^2$ if i and j are equal then it does not contribute. And when you flip i and j then you get the same contribution, right? So it is being counted twice. So

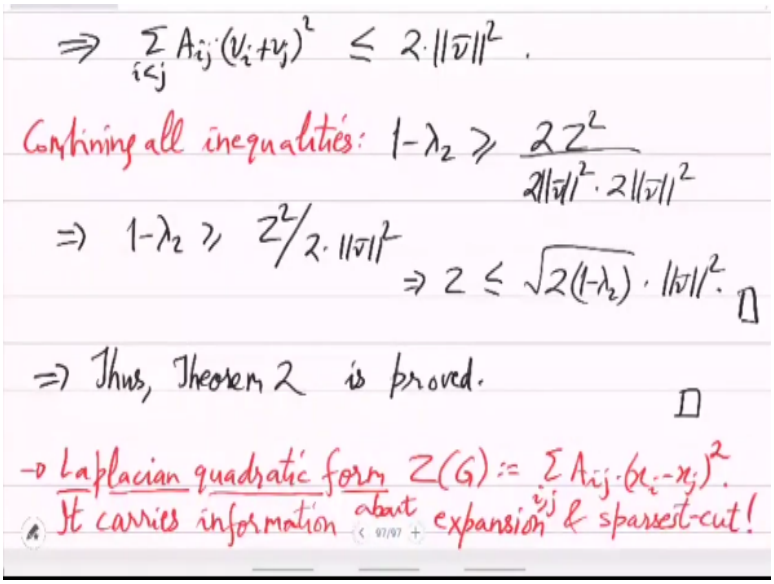
you get a factor of 2 here, okay. So the numerator estimate is giving you $2(\sum_{i<j} A_{i,j} (v_i^2 - v_j^2))^2$. So this is by Cauchy-Schwarz inequality, okay.

That is the lower bound on the numerator, which is equal to$2z^2$, right? So we have related this right hand side, at least the numerator we have related to Z nicely. Now let us look at the denominator. So the denominator we have to upper bound, right so that for the ratio we get a lower bound. So denominator is at most what? So I will first write $1/2\sum A_{i,j}(v_i + v_j)^2$ This expression is equal to$\frac{1}{2}\sum A_{i,j}(v_i^2 + v_j^2) + \sum A_{i,j} v_i v_j$.

And this I want to now relate with the denominator expression that you have. So this is again a bunch of calculation. So let us do that. So what is this first expression, $\frac{1}{2}\sum A_{i,j}(v_i^2 + v_j^2)$. Since we are going over all i j this will be equal to $||\bar{v}||^2$ as we have seen many times before. It will be half of two times $||\bar{v}||^2$.

$||\bar{v}||^2+\sum_{i,j} A_{i,j} v_i v_j \leq ||\bar{v}||^2 + \frac{1}{2}\sum A_{i,j}(v_i^2 + v_j^2)=2||\bar{v}||^2$. So we have estimated this thing, sigma over all i j. This is what we have estimated.

(Refer Slide Time: 26:54)



So from this you can deduce that $\sum_{i<j} A_{i,j}(v_i + v_j)^2 \leq 2||\bar{v}||^2$ because we have estimated half of that expression which is which also gives you an estimate on the denominator, this term. And now when you combine everything that we have learnt, you will get the following.

So combining all inequalities what you will get is $1 - \lambda_2 \geq \frac{2z^2}{2||\bar{v}||^2.2||\bar{v}||^2}$ So which implies that $1 - \lambda_2 \geq \frac{z^2}{2||\bar{v}||^2}$ which implies that $Z \leq \sqrt{2(1 - \lambda_2)}.||\bar{v}||^2$.

So there might be, I have to recheck this calculation. If there is a mistake then I will update in the PDF, okay. Let us move forward. So these are the basic inequalities and when you combine everything you will get the proof of claim 2. So because in the end I am getting something better than what claim 2 said. So this is why I have to recheck this but anyways you have seen the main points.

So we have shown both claim 1 and claim 2 and when you combine them together you will get the relationship between ρ and λ_2. So this finishes the proof of theorem 2 also. So thus theorem 2 is proved, okay. So the key things that we used in all these proofs to relate algebraic expansion

with combinatorial expansion was this Laplacian quadratic form Z(G): $= \sum_{i,j} A_{i,j} (x_i .... x_j)^2$ which is okay we use this, square of this.

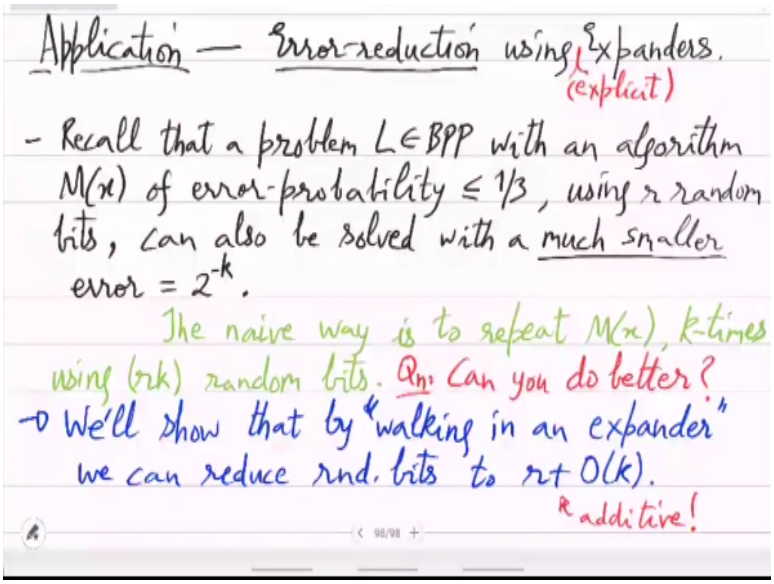
And that was certainly the case in theorem 1. In theorem 2 also although we started with $(x_i^2 - x_j^2)$ ultimately we did use this $(x_i - x_j)^2$ as well. When we did this Cauchy-Schwarz inequality, we actually got something like that, okay. So in both the proofs, this kind of object was critical. And the reason is, intuitive reason is that this quadratic form carries information about expansion and sparsest cut in graph, okay.

So not only is this useful in this mathematical analysis of the graph, telling you about what is the expansion for every subset of vertices and what is the spectral gap in the graph, but also in algorithms it is used for finding the sparsest cut, okay. So this is a very important object. And so you have to carefully look at these calculations again and get an idea of actually what we did. These were pretty intense calculations.

So next we will, what we will do is one application of expanders. So first in this business, the first thing you saw was this randomized algorithm to solve connectivity in a graph in logspace, right? But this was randomized logspace. Still it is a very non trivial algorithm the reason why it works. And from there we got the idea that maybe we can study graphs, which are highly connected.

So you can reach anywhere with equal probability in just log many steps. And we defined algebraic expansion and combinatorial expansion. Now at this point, we do not know whether these things exist. We have defined them, we do not know whether they exist. So we will solve that problem later. For now just take it on faith that they do exist. And so now how can you use these objects, okay. So there is this algorithmic application that we will see.

(Refer Slide Time: 33:37)



Which is called error reduction, using expanders. So this is something to do with randomized algorithm. So suppose you have a randomized algorithm for a problem $L \epsilon BPP$, makes an error probability $\leq 1/3$. So suppose you want to reduce this error further, okay exponentially below. So suppose random bits that it needs is r, can be solved with a much smaller error.

So this error probability of one third if there is an algorithm of this type, polynomial time algorithm, then there is also another polynomial time algorithm that can solve the same problem with error much smaller, something like 2^{-k}, right? So how is this thing done? This is just repeating the algorithm again and again with different random bits, right repeating this experiment k times.

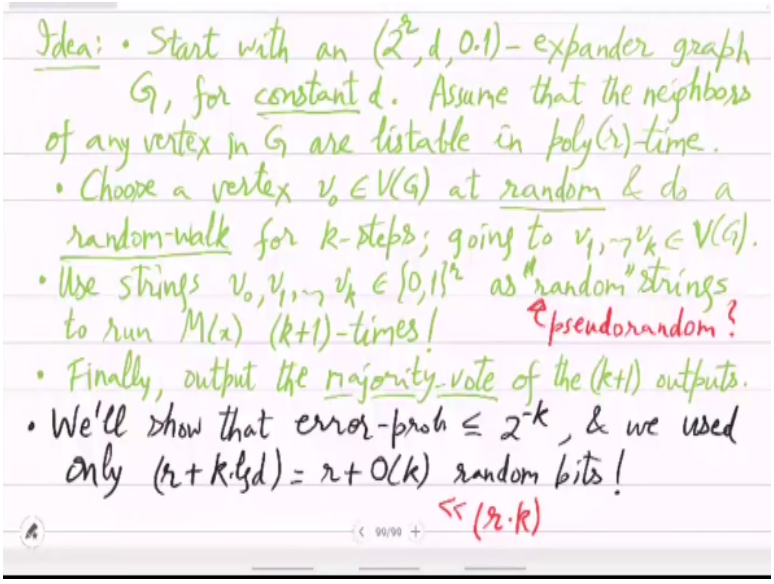
So the naive way is to repeat M (x) , k times using(rk) random bits. So if you repeat it k times random bits are also much more and you will get error probability exponentially smaller, time will be just k times. So the question we ask is can you save on these random bits? So can you do better?

Can you manage this same error probability with much fewer random bits? So we will show that by expanders, walking in an expander we can manage this in r + k random bits. It can reduce random bits to r +O(k). So instead of multiplicative growth, you will only get additive growth,

okay. So that in practice could mean a lot of savings, okay. You have to toss fewer coins and still get a very good success rate, exponentially low error rate.

So what is the idea here? So in this idea, you will see why an expander graph can be so helpful and specially explicit expanders. So if there are explicit expander constructions, then you can achieve this. So the idea is as follows.

(Refer Slide Time: 39:20)



You pick a random vertex in your expander graph. So that will need let us say r bits and then you do a random walk k steps, the vertices which you reach you use them as random bits or random strings for your algorithm, the algorithm M(x). So start with a (2^r, d , 0.1) expander,graph G okay. We can think of d to be constant, say 3. And assume that the neighbors of any vertex.

So remember r was like the input size, right? So 2^r is a very large vertex set. So you cannot, you do not want to look at every vertex, you do not want to store this expander graph. You just want the ability to locally look at the neighbors wherever you are in this expander graph. So you want this a one-step connectivity information in this expander graph G. So suppose there is a fast algorithm for that.

So assume that the neighbors of any vertex in G are listable in poly (r) time, which is very fast in terms of the big expander graph exponentially large, it is very quickly you can find the neighbor

of any neighbor set of any vertex v in the graph. So suppose this is what I am calling explicit expander. So if you have such explicit expander construction, then you can actually do a random walk k steps, okay.

So you will get these k strings. And when corresponding to every vertex there is a string right, there is a string of r bits. So you will get these k strings each of r bit and use these k strings in your randomized algorithm M x as random strings. So we have to, obviously it is not clear what will be now the error probability of M, right? That seems to be a complicated thing. So next time we will do that analysis.

But the sketch of the algorithm is just this. So choose a vertex $v_o \in V(G)$ at random and do a random walk for k steps, okay going to vertices v_1 to v_k. So you pick a random vertex v_0 in your explicit expander G and then you do a random walk for k steps. So that is v_1 to v_k, okay. So now you have v_0 to v_k vertices. Remember each vertex you can think of as an r bit string. So use strings $v_0, v_1, v_{k \in \{0,1\}^r}$

. So binary r bit strings as random string. This is "random" because it is not that you picked v_0, you picked v_0 actually randomly but v_1 you really have not picked it randomly, right? It is just in this fixed explicit expander graph. You have actually taken a neighbor of v_0. So it is not really a random string.

But anyways you give this to your algorithm M(x) and v_2 and v_k so on as random strings to run M on the input x, (k + 1) times, okay. That is the thing. So you are trying to fool the algorithm M(x) by giving it a pseudorandom string. So will this work is the question? Or how well will this work?

What will be the error probability of this algorithm M repeated k + 1 times, right? So repetition of k + 1 times means that you run that many times and then take the majority vote of the answer. So finally, you take the majority vote. So finally output the majority vote of these k + 1 outputs. That is the modified algorithm M (x), okay.

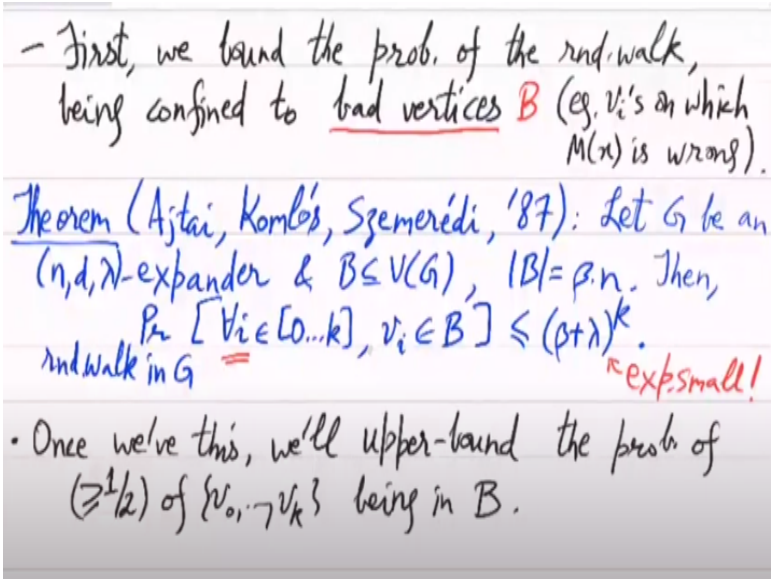
So according to where you reach in the expander, explicit expander graph, you use that as a random string, give it to M (x). M x will do a computation and output something. So you get (k + 1) outputs. Take the majority vote that is the final answer. So what we will show is the following. So we will show that error probability $\leq 2^{-k}$.

And of course, we used only$(r + klogd) = r + O(k)$ So you are getting 2 $^{-k}$ error for random bits only r + k. The trivial would have given you r times k.

That is the improvement. So this is≪ (r . k), okay. So that is the big improvement, quadratic to linear. So the question is how do you show this? You do not know anything about the algorithm M, right? M is an arbitrary randomized polynomial time algorithm. And the explicit expander is also does not have any extra property. So the algorithm and the expander are unrelated.

So you have to now analyze the algorithm, this unknown algorithm using this unknown expander, right? So how will you do this? So this is a very interesting exercise. We will finish this proof next time.

(Refer Slide Time: 49:14)



So first we bound the probability of the walk, of the random walk being confined to bad vertices. So what are bad vertices? So when you are taking a random walk of some number of steps, we

want to understand, in this expander graph we want to understand what is the probability of not visiting certain vertices, okay. So these are the vertices which we are calling bad.

So in our application these will be the strings on which M fails. So example v_i's on which M (x) fails, okay. So these vertices we do not want to touch or we do not want to reach. So what is the chance of avoiding all these vertices, bad vertices? Let us first calculate that. So that is a famous theorem by Ajtai, Komlos and Szemeredi. So what this theorem will say is for an expander (n, d, λ)expander and bad vertices $B \leq V(G).|B| = \beta n$

Now of course, if the bad vertices are a lot in the graph, if B is a very big set, then you then the probability of avoiding them will be very low, right? So suppose that it is not it is just some fraction β, you can think of β is 10%. So say only 10% vertices are bad. So we do not want to reach them, okay. What will be the probability of that? Then the probability $Pr\ [\forall i \epsilon [0 \ldots k], v_i \epsilon B] \leq (\beta + \lambda)^k$.

So probability that every vertex is bad, right? So notice this for all. I did not emphasize it before, this is important. So we are saying by this idea of confined to B we actually meant that the random walk is this. So what is the chance of that happening. So you would expect it to be quite small, the chance to be quite small if the as long as the bad vertices are few. And that is what this statement is saying, okay.

So that it is exponentially small. So never able to leave this bad vertex set. That is a very small probability. It is exponentially small in k and depends on β and l. So basically this fraction of the bad vertices should be small, beta should be small and the second biggest eigenvalue of the graph lambda that also should be small. If the both are small, then you have this result.

So this is the theorem we will prove and once we have shown this then we will have to analyze or estimate the probability of the random walk reaching more than half vertices being bad, okay. So once we have this we will also estimate the probability of $\geq (1/2)$of $\{v_0 \ldots . v_k\}$being in B. And not just estimate, we will show it to be small, okay.

So once we have this, once we show that the probability of more than half of your vertices visited being bad is small, once we have shown that probability to be small then we are done, right? So that will mean that no matter what algorithm M was, since or one half or you can maybe even say one third.

So basically what it would mean is that in this random walk that we are taking, usually less than one third of the visited vertices will be in B, will be bad okay, the two third will be good. So two third of the answers will be right. So actually half would also do, half would also do. Yeah so less than half will be, usually less than half will be bad. So at least half are good. So on which M(x) will give the correct answer.

And so when you take the majority vote you will get the correct answer with high probability, okay. This is how we will approach this. So this theorem of Ajtai and others we will do next time by matrix analysis which will be an interesting exercise. We will build on what we have done before.