YUFEI ZHAO: For the last few lectures, we've been talking about the extremal problem of forbidding a complete bipartite graph. So today I want to move beyond the complete bipartite graph and look at other sparser bipartite graphs. So we'll be looking at what happens to the extremal problem if you forbid a sparse bipartite graph.

Recall the Kovari-Sos-Turan theorem, which tells you that the extremal number for KSC is upper bounded by something on the order of m to the 2 minus 1 over m. So if I give you some bipartite H, we know that because it's a bipartite graph, it is always contained in some KST for some values of S and P.

So you already automatically have an upper bound, the extremal number for this H from the Kovari-Sos-Turan theorem. But as you might expect, we might be very wasteful in this step. And the question is, are there situations where we can do much better than what is just given by applying Kovari-Sos-Turan? And today, I want to show you several examples where you can significantly improve the bound given by Kovari-Sos-Turan for various sparser bipartite graphs.

The first result is the following theorem, which is initially due to Faraday. And then later, a different proof was given by Alon-Krivelevich-Sudov. And the latter proof is the one I want to present, because it presents an important and intricate probabilistic technique, which is the main reason for showing you this theorem. But let me tell you the theorem first.

Here, H will be a bipartite graph. And it has vertex bipartitions A and B, such that every vertex in A has degree at most R. The bipartition is A and B, and the degree in A is at most R for every vertex on the left side in the set A. And I want to understand, is there some upper bound on the extremal number that does better than Kovari-Sos-Turan theorem? And the theorem guarantees such a bound. So then there exists some constant depending on H, such that the extremal number is upper bounded by something to the order of n to the 2 minus 1 over R.

Compared with Kovari-Sos-Turan theorem, on one hand, if your H is the complete bipartite graph KST, then this is the same bound as Kovari-Sos-Turan theorem. On the other hand, you might have a lot more vertices in A and a lot more vertices in B. The hypothesis only requires what the degrees in A, max degrees, at most are. So it could be a much bigger graph. So if you apply Kovari-Sos-Turan, you would get a much worse bound compared to what this theorem guarantees.

This 1 of R is optimal. In this given statement, you cannot improve this 1 of R, because we know that from the KST example and the lower bounds I showed you last time, you cannot improve upon this 1 over R. Also, in this form, this theorem is best possible.

I want to show you a probabilistic technique for proving this theorem. And this is an important idea called dependent random choice. So let me first give you an informal interpretation of what's going on. So the idea is that if you have a graph G with many edges, then inside G, I can find a large subset of vertices U, such that all small subsets of vertices in U have many common neighbors. I won't tell you what the small and many are just yet, but we'll see it through the proof of the theorem, but that's the idea.

So I give you a graph. It's not too sparse. It's relatively dense. Then I should be able to find some subset that is fairly large so that, let's say, every pair of vertices in the subset has many common neighbors. So let me write down-- or at least attempt to write down-- the formal statement of dependent random choice. The statement of the theorem, as I will present, has a lot of parameters, but don't get-- I don't want you to be scared off by the parameters, so I won't even tell you what they are, but first tell you what the statement of the conclusion is. And then we'll derive what is the dependencies on the parameters, so the proof or the technique is much more important than the statement of the theorem itself. So I'll leave some space here.

The conclusion is that every graph with n vertices and at least alpha n squared over 2 edges contains a subset U of vertices. U is not too small, so the size of U is at least little u. And such that for every subset S of U with R elements, the set S has at least M common neighbors. So what's the idea here? I give you this graph, and I want you to produce the set U that has this property. How might you go about finding the set U?

Let me give you an analogy. Let's suppose you have the friendship graph on MIT campus. And I want to select a large set, let's say a hundred students, such that every pair of them or maybe even just most pairs of them have many common friends. Well, how might you go about doing that? Well, if you select a hundred students at random, you're unlikely to achieve that outcome. They're going to be pretty dispersed across campus.

But if you focus on some commonalities-- so, for example, you go to someone, some specific individual, and look at their circle of close friends, then it seems more likely you'll be able to identify a group of people who are very well-connected in that they, peer-wise, have lots of common friends. And that's the idea here. We're going to make the random choice by picking that core individual-- that's the random choice-- but then make the subsequent dependent random choice by looking at the group of friends from that specific random individual instead of choosing the hundred people uniformly at random, which is not going to work.

So let's execute that strategy on this graph. Let me take T to be the random set, so that's the core set. So let T be-- for convenience, I'm going to choose with repetition. Instead of just choosing one person, I'm going to use two T vertices, so a list of T vertices chosen uniformly at random from the vertex set. So G is going to be my graph. So the graph is G.

And what we are going to do is look at the set A, which is the set of common neighbors of T, the vertices that are adjacent to all of T. And that's the set I want to think about. So I want to basically argue that this set A has more or less the properties that we desire, maybe with some small number of blemishes, which we will fix by cleaning up.

First, I want to guarantee that A has a big size-- we are actually choosing a lot of vertices. So let's evaluate the size of A in expectation. By linearity of expectation, we need to compute the sum of the probabilities that individual vertices fall in this random A. For each particular vertex V, when is it in A? Well, it is in A if T is contained in the neighborhood of this V. So all of your children T fall into the neighborhood. Otherwise, V is not going to be contained in A.

So each individual probability can be computed easily by looking at the size, so the degree of V divided by the number of vertices raised to the power T-- T independent choices chosen with replacements. And by convexity applied to the final expression, we find that it is at least this quantity here where essentially we're taking the averages of the degrees. So this is by convexity. And finally, the graph has at least that many edges, so the final quantity there is at least N alpha to the T. Yes, question?

AUDIENCE: [INAUDIBLE]

YUFEI ZHAO: The question is, in our original list, are we allowed to have repeated vertices? Yes. So it's a list, and we're choosing every element independently at random, allowing repetition. So it's sometimes called choosing with replacement. You choose a T. You throw it back. Choose another one.

The property that we're looking for is that for every S subset, for every r-element subset of U, it has many common neighbors. So let's look at such an S. So for each S which is an r-element subset of the vertex set. So r-elements subset of V, what is the probability that this set F is contained in A? I'll give you this set S.

It is contained in A if-- well, let's think about how A is chosen. You want this to happen. An A is chosen as a common neighborhood of this T-- so assets containing A, if and only if, P is contained in the common neighborhood of f. So f is fixed for now. P is random.

So we draw the elements of T independently, uniformly, at random. Therefore, this probability is equal to the number of common neighbors of S as a fraction of the total number of vertices. This fraction raised to the power T. We want all S subsets of A-- all r-element subsets of this A-- to have at least m common neighbors. But maybe we cannot get that.

So let's figure out how many bad S's are there. How many S's do not satisfy this condition here? So let's call such a set S bad if it has fewer than m common neighbors. And from this equation, we see that for each fixed S that is in r-element subset of vertices, it is bad with probability-- it is bad if it has few common neighbors. Because if you have few common neighbors, then this probability is small. So it is bad with probability strictly less than m over n raised to the power T.

We chose this A in this dependent random way. And basically, we want all subsets. All r-element subsets have many common neighbors, but maybe we cannot get that at the first try. But we only have a small number of blemishes, so we can fix those blemishes by getting rid of the possible bad r subsets. And that's fine to do, as long as there are not too many of them.

And indeed, because S is bad with small probability, the expected number of bad r-element subsets of A is, at most-- well, I look over all possible r-element subsets of vertices. Each one of them is bad with this probability here. So we have that bound. And the point now is that if this number is significantly smaller than the expected size of A, then I can clean up all the bad subsets by plucking out one vertex from each bad subset.

So indeed, that is the case, because the expectation of the size of A minus-- so let me call this quantity here star-- so the size of A minus star by what we have shown is at least n alpha to the t minus the quantity just now. And I want this number to be somewhat large. And now let me put that as a hypothesis of the theory.

So dependent random choice-- let u, n, r, m, t be positive integers, alpha positive real, and suppose n alpha to the t minus n to the r. Cancel the n to the t is at least w. That's where that inequality comes from. So if this is at least w, then what we can do is delete one vertex from each bad subset S. And after deleting, A then becomes some smaller set A prime with at least u elements in expectation, but then there exists some new elements--

Let me put it this way. We know that this is true in expectation, thus there exists some T, such that that inequality is true without the expectation, such that there exists some T, such that-- so this quantity is at least that. Now, to delete a vertex from each bad subset, we obtain this A prime with at least u elements. And we have gotten rid of all the possible bad subsets-- so no bad r-element subsets. And that finishes off the proof of dependent random choice.

Just to recap, the idea is that if you have a dense enough graph, then the conclusion is that you can find a fairly large subset of vertices, so every pair, every r-element, have many common neighbors. And the way you do this, instead of choosing your set at random, which is not going to work, you choose a small set of anchors. T, you think of that as the anchors.

You choose a bunch of anchors. And then you look at their common neighborhoods and use that as a starting point. That will almost work. It might not work perfectly, but then you fix things up by removing the blemishes. So this is a very tricky probabilistic idea. It's also a very important one. It will allow us to prove the theorem over there about the extremal numbers of standard degree graphs. Question?

AUDIENCE: Is the definition of bad set for any subset of V or is it only for subsets of A?

YUFEI ZHAO: The question is, in the definition of bad, do I use this definition for all subsets of V, all r-element subsets, or just subsets of A? So I use it to mean all subsets of V, because A is random. A is random, so the definition of bad does not depend on the randomness. It only depends on the original graph.

AUDIENCE: How is the badness dependent on any probability?

YUFEI ZHAO: The badness does not-- so the question us, how does the badness depend on any probability? The badness does not depend on the probability, but A is random. So the number of bad r-element subsets of A is a random variable. So each individual-- so you start with a graph. Some r-element subsets of graphs. Some are not.

And now I choose this random A in this dependent random manner. And A might contain some bad subsets. I'm trying to calculate how many bad subsets does A have. Question?

AUDIENCE: [INAUDIBLE], because an S is neither bad or not bad. You said it's bad with probability.

YUFEI ZHAO: So your concern is each S is bad with probability-- ah. Sorry.

AUDIENCE: [INAUDIBLE]

YUFEI ZHAO: Fine. Yeah. Thank you. So each bad subset-- OK, so for each fixed bad subset, it is contained in A with probability. Thank you. I hope this makes it clear. So the probability of the-- the property of being bad is not random, but it containing A is random. Question?

AUDIENCE: [INAUDIBLE]

YUFEI ZHAO: So the question is, why is this true? So why are these two events the same? And it's kind of hard to steer the definition a bit. So T is chosen as the common neighborhoods-- sorry. You choose T at random. And you choose A to be the common neighborhood of T. And so how do you characterize subsets of A if every element is connected to all of T? You have to think about it. Any more questions? Yes?

AUDIENCE: [INAUDIBLE]

YUFEI ZHAO: We pick T at random. T is uniform. So the question is, how are we picking T? T is uniform at random. Oh, great. So the question is, how do we pick the little t in the theorem statement? It depends on the application.

And that's a little bit weird, because in the statement of the theorem, the little t shows up in the inequality, but not in the conclusion. So you think of t as an auxiliary parameter. So the little t comes up in the proof, but not really in the conclusion. Any more questions?

It's a tricky lemma. It's a tricky idea. So now let me use it to prove the statement over there. And here, it's not so hard. So it's mostly an application of this dependent random choice lemma. So for all H that's in the theorem-- so let me prove this lemma-- so for all H that's in the theorem, there exists a constant C, such that every graph with at least C n to the 2 minus 1 over r edges contains a vertex U with-- so vertex subset U with the size of U equal to B-- so B comes from the vertex by partition of H. It's a constant-- such that every r-element subset in B has-- so r-element subset in U has lots of common neighbors. That's at least another constant, which is the number of vertices in H-- that many common neighbors.

So you see, it's a direct corollary of the dependent random choice lemma by setting it in the right parameters and, indeed, by the dependent random choice lemma where we choose this T-- this auxiliary variable T in independent random choice lemma-- equal to r. So suffices to check that there exists of C such that-- plug it into that expression, that inequality, up there. So n2Cn to the minus 1 over r-- raised to the power r minus n choose r. And then this expression here.

So I'm just plugging in the various graph parameters into the dependent random choice statement. And I want to show that you can find a constant C such that this inequality is true. And indeed, these exponents, they cancel out, so the first term is simply 2C raised to the r. And the second one, because, again, you notice that the exponents work out just fine, so it is, at most, a constant. So you can choose C big enough so that this is true. So it's a direct verification of the hypothesis of the lemma. And now we're ready to prove the theorem over there. Yes, question?

AUDIENCE: How do you have size A plus B common neighbors? Isn't that the entire graph? Or can I just [INAUDIBLE]?

YUFEI ZHAO: The question is, how do you have size A plus B commons neighbors? So H is fixed. And A and B are constants, so A plus B is the number of vertices of H. That's a constant.

AUDIENCE: Oh, sorry. Oh, OK.

YUFEI ZHAO: Yeah, I'm talking about common neighbors, not in H, but in the big graph G as n vertices. Now, I like questions. It's tricky. This is a tricky argument, so please do ask questions if you're confused. And there are times when I may not have explained it very well, so please do ask questions.

So let's prove the theorem. And now we're almost there. The idea is that we embed the vertices of B into the big graph G one by one. Sorry. First, embed B into the vertices of G using U from the lemma-- so the lemma that we just stated. And they claim that once you have done that-- so the vertices of B-- and now I need to embed the remaining vertices of A.

And I can do this one by one, because if I need to embed some vertex of A, has, at most, r neighbors to B. But I have embedded B in such a way that they have a lot of common neighbors inside G.

So I can always do it. So I can always embed the vertices of A one at a time in such a way that I even avoid collisions. I don't allow vertices to be embedded into the same place. This is all using that B, embedding a B, which is U, has many common neighbors in G. So once you put that in. And the rest, you just make one choice at a time.

And you need to embed a second vertex, or you can find somewhere in their common neighborhood that allows you to do it. So you embed the vertices one at a time, and then you finish embedding the whole graph. Any questions? It's a tricky argument. So let's take a break, and I want you to think about it.

Any questions? Yes.

AUDIENCE: [INAUDIBLE].

YUFEI ZHAO: So the question is, how do you embed A without having any collisions? So I put in the vertices of B so that every r of them have many neighbors. And now, I want to try to embed the vertices of A one by one in any order. Think about-- you pick the first vertex. Where can it go? It has-- let's say it's adjacent to the first three vertices of B. So, in the embedding, it has to go in the common neighborhood of those three vertices, which we know is large. So I put them anywhere. And I do the same for the second vertex, do the same for the fourth vertex. So I just keep on going.

Because the common neighborhood is large, it may be that some of the potential vertices I might embed is already used by the previous steps in the process. But because I always have at least A plus B common neighbors, I always have some possibilities that remain. Yes, question.

AUDIENCE: So [INAUDIBLE], like the last line of the proof to [INAUDIBLE] delete a vertex from each bad subset, and then to make that [INAUDIBLE].

YUFEI ZHAO: Ah. The question is, how does this bad subset deletion work? So you have this A which is fairly large. And you know that there exists some instance-- there's some incidence of this randomness that produces for you a situation where A has very few bad r subsets. So then I take A and I delete from A one vertex in each bad subset. I haven't changed the size of A very much. A is still quite large after this deletion, but now A has no bad subsets remaining, because I've gotten rid of one vertex from each one, from each bad subset.

So they're very similar to what we've seen before, the random process for creating an H-free graph. You generate a random H-free graph which has very few copies of H relative to the number of edges, and then you get rid of them by removing one edge from each copy of H.

So with the theorem that we saw in the first part of the lecture, we saw how to improve on the bound of Kovari-Sos-Turan in some circumstances, namely one where the graph that you're forbidding, this H, essentially has bounded degree. We stated something a bit stronger, namely has bounded degree from one side. And that's a pretty general result. And now I want to look at some more specific situations where you might be able to improve further.

So what are some nice bipartite graphs? One that comes up is kind of even cycles. So if you have C4, C6, and so on. And you see C4 is the same as K2,2, which we already saw before. But even C6, the techniques so far allow us to obtain, with the theorem that we just saw-- gives us a bound on C6 that's more or less the same as that of C4, namely n to the 3/2. So what's the truth for C6?

Well, it turns out that you can do much better. So this is the theorem of Bondy and Simonovits that, for all integers k at least 2, there exists some constant C such that the extremal number-- well, I'm going to use C too many times. So I'll just say that the extremal number of C2 sub k is, at most, on the order of n to the 1 plus 1 over k. So, in particular, for 6 cycles, the upper bound is 4/3 in the exponent. It's better than the 3/2.

So there's another class of graphs where there are some nice upper bounds. So you can ask, well, just like Kovari-Sos-Turan theorem for complete bipartite graphs, do we know matching lower bound constructions? And what is known is that it is tight only for a small number of cases. And the others, we do not know whether they are tight. So this Bondy-Simonovits theorem, it is tight for k being 2, 3, or 5, and open for others. So there are constructions for C4,3, C6,3 and C10,3, but not for C8,3. That's an open problem.

The proof of the Bondy-Simonovits theorem is slightly involved, but I want to show you a weaker result that already contains a lot of interesting ideas. So a weaker result is this. That for every integer k at least 2, there exists a constant C such that every n-vertex graph G with at least C-- so this-- the correct-- the same order of number of edges-- so we know for Bondy-Simonovits, it contains an even cycle of length exactly 2k. So we'll show something slightly weaker that contains an even cycle of length, at most, 2k. Which, in other words, says that the extremal number, if you-- so we haven't introduced this notation, but, hopefully, you can guess what it means.

Then, if you forbid all of these cycles, then it is this quantity there. So I'll show this weaker result. All right. Let's' do it. First, I want to show you a couple of easy preparatory lemmas. The first-- so every graph G contains a subgraph with min degree at least the half of-- at least half of the average degree of G.

So you have a graph G. It has large average degree. It has lots of edges. But maybe there are some small degree vertices. And it will be useful to know that the minimum degree is actually quite large as well. So it turns out, by passing to a subgraph, you can guarantee that. How do you think we might prove this? I give you a graph. I know it has lots of edges. But maybe some vertices have small degree. Yes.

AUDIENCE: [INAUDIBLE].

YUFEI ZHAO: So you suggest dependent random choice. That's a heavy hammer for such a simple statement. And so it turns out we can do something even simpler.

AUDIENCE: [INAUDIBLE].

YUFEI ZHAO: So we'll just throw out the min degree vertices. So throw out the small degree vertices. So if the average degree is 2t then the number of edges is number of vertices times t. And you see that removing a vertex of degree t, or at most t, cannot decrease average degree. So I have this process where I remove vertices that are less than half of average degree. And the average degree never goes down, so I keep on doing this. Average degree stays the same. Well, it can go up, but it never goes down.

And when I stop, I don't have any more small degree vertices to get rid of. Why does the process even terminate? Maybe we end up with the empty graph, which is not so useful. Yes.

AUDIENCE: The average degree is [INAUDIBLE].

YUFEI ZHAO: He said the average degree is non-decreasing. But how do I know the graph has at least some number of vertices when I stop? Yes.

AUDIENCE: [INAUDIBLE].

YUFEI ZHAO: So if you-- you must terminate. You must terminate because every graph, if you have way too small number of vertices-- because, for example, if-- we'll just notice that every graph with, at most, 2t vertices has average degree less than 2t. So you'll never get below 2t vertices. You'll run out of room if you go too far. Because that's the first preparation lemma.

The second one is that every graph G has a bipartite subgraph with at least half of the number of edges as the original graph. This is a very nice and quick exercise in the probabilistic method. So we can color every vertex black and white uniformly at random. And the expected number of black to white edges is exactly half of the original number of edges by linearity of expectations. So there's some instance with at least that many black-white edges, and that's a bipartite subgraph.

So now we can prove the theorem about even cycles, the weaker theorem at least. I start with a graph which has a lot of edges. But by these two lemmas, and changing constant somewhat, I can obtain a subgraph with-- that's bipartite and has min degree quite large. So I lose, at most, a constant factor, and I get basically within a factor of 2 of-- well, factor of 4 of the average degree. So let me call this quantity delta, the min degree in this subgraph, this bipartite subgraph.

Let's think about what happens to this graph if I start with an arbitrary vertex. So now I have a min degree condition. It's, really, all vertices are now kind of the same to me. So pick an arbitrary vertex, and look at its neighborhood. It has at least delta edges coming out. So let me call the first vertex level 0, and the second set level 1. It's bipartite, so there are no edges within level 1. Let's expand out even further.

Can there be some collisions where two of these edges go to the same vertex? Well, if there were, then I find a C4. So if I assume-- so let's assume that there's no C4, C6, and so on, C2k, for contradiction. So because there's no C4, all the endpoints of this path of length 2 are distinct in level 2.

And you keep going, or you keep expanding further. And so on. And all the way to level t. And all of these guys must be distinct as well. Because, otherwise, you'll find a cycle of length, at most, 2t. So when you do this expansion, each step, you get distinct vertices. And you also have no edges inside each part because you are looking in the bipartite setting.

So how many vertices do you get at the end? Well, I have a min degree condition. The min degree condition tells me that I expand by a factor of at least delta minus 1 each time. So the number of vertices in here is at least delta minus 1 raised to the power t-- so raised to the power k. Here's k. And expand all the way to the end.

But, you see, that number there is quite large. So, in particular, if t is large enough, then it's bigger than n. And that will be a contradiction, because you only had n vertices in the graph to begin with. Therefore, the assumption that there are no cycles, even cycles of length, at most, 2k, is incorrect. And that finishes the proof. Any questions? Yes.

AUDIENCE: Are the ideas for the full Bondy-Simonovits theorem similar?

YUFEI ZHAO: So the question has to do with, in the original, in the full Bondy-Simonovits theorem, what do we need to do? Are they similar? So, certainly, you want to do something like this. But, then, you also want to think about, if you do have short cycles, how can you bypass them? So there's a more careful analysis of what happens with shorter cycles. And we will not get into that. Any more questions? All right. So the first thing that we did in today's lecture has to deal with what--

AUDIENCE: Quick question.

YUFEI ZHAO: Yes.

AUDIENCE: So why is-- so why are all vertices distinct for level k? Or are they just defined that way?

YUFEI ZHAO: So question is, why are the vertices distinct for level k? So at level k, if you have some collapse, then it came from two different paths, therefore forming a C2k.

AUDIENCE: OK.

YUFEI ZHAO: So now I want to revisit the first theorem that happened in today's lecture, namely if you have this H, bipartite A and B. So we saw the hypothesis was that if every vertex in A had bounded degree, degree, at most, r, then we got the upper bound on the extremal number. That was n to the 2 minus 1 over r. And, in particular, for r equals to 2, suppose I have that situation there.

Suppose that the degree for every vertex in A is, at most, 2. Then the first theorem guaranteed us that the extremal number is on the order, at most, n the 3/2, just like the extremal number for 4 cycles, for K2,2's. And, of course, this statement is tight, in the sense that if I change this 3/2 to any smaller number, then, well, just taking H to be K2,2 violates it. So I cannot replace in this generality 3/2 by any smaller number, because we know that the extremal number for K2,2 is this order.

But is that the only obstruction? So if H is not K2,2, well, you can make some sillier examples too by taking K2,2 and add some more edges. So lets forbid H from having a K2,2 subgraph. Can you do better now? So in this case, can you improve for the specific H that exponent? And we already saw one case where you can do this, namely in Bondy-Simonovits theorem for cycles. Or if you only applied this theorem here, you get 3/2, but Bondy-Simonovits tells you a much better exponent. So let's explore the situation.

And it turns out, in a very recent theorem that is only proved the last couple of years by David Conlon and Joonkyung Lee, they showed that for every H as above, there exists constants little c and big C such that the extremal number of H is upper bounded by something where I can decrease 3/2 to some even smaller number. So, somehow, this 3/2, now we understand is really because of the presence of K2,2. The graph has-- if H has no K2,2, then some smaller number suffices. And I want to use the rest of today to explain how to prove that theorem there. Yes, question.

AUDIENCE: The C is not independent of H [INAUDIBLE].

YUFEI ZHAO: That's right. So the question is, is C independent of H? So C depends on H. So C, they are dependent on H. Questions? Let me put this question in a slightly different formulation that is also equivalent. So in graph theory, there is a notion of a subdivision. And, in particular, a one subdivision of graph H is this operation where you start with a graph-- let's say this graph here-- and you add a vertex to the middle of every edge of this graph.

So, initially, it's 4 vertices. Now you add a new vertex to every edge. So you subdivide every edge into a path of two edges. That's called a subdivision. So for today's lecture, let me denote subdivisions by a prime. And, in particular, if this is K4, then I will denote this graph here K4 prime. So, for example, K3 prime, that's a triangle subdivided-- well, that's a C6, [INAUDIBLE].

So observe that every H that comes up in this theorem here is a subgraph of some subdivision, some one subdivision of a clique. Because the vertices on the left in A are degree 2, you think of them as midpoints of edges. But because you are K2,2 free, you-- if you collapse those path of lengths 2 to single edges, you do not end up with parallel edges. So it is the subgraph of this one subdivision of some graph, which, then, you can complete to a clique.

So this theorem here is equivalent, at least qualitatively, to the statement that for every t, there exists some constants, again depending on t, such that the extremal number of the one subdivision of a clique is bounded by something where we can improve upon the exponent in the first theorem in today's lecture. So that theorem there. So these two theorems are equivalent because of this remark. Any questions so far about the statements? Yes.

AUDIENCE: In the remark, how do you deal with, like, [INAUDIBLE].

YUFEI ZHAO: Question, in the remark, how do you deal with vertices with degree less than 2? Complete it to a vertex degree of-- vertex of degree 2. Add another edge.

AUDIENCE: [INAUDIBLE].

YUFEI ZHAO: Add another edge to a new vertex.

AUDIENCE: OK, sure.

YUFEI ZHAO: Any more questions? All right. So the proof I want to show you is due to Oliver Janzer, and for this clique subdivision theorem. And this proof produces that C sub t equals to 1 over 4t minus 6. So if you plug in t equals to 3, you find that the exponent here is actually right for the 6 cycle. So it actually agrees with what we know.

So I want to show you some of the main ideas from this proof. Just like the proof of-- that we saw for the cycles, even cycles theorem, it will be helpful to start with some preparation. You start with a graph that, even though it has a lot of edges, may have lots of vertices with high degree, lots of vertices with low degree. It's nice to clean it up somewhat. And so let me state a preparatory lemma, which we will not prove, but it's of a similar nature to this very easy lemma that we saw earlier but with a bit more work.

AUDIENCE: [INAUDIBLE].

YUFEI ZHAO: Yes. Originally, I put a C over here, but now it's OK. So there exists a Ct. So the preparation is that we're going to pass through a large, almost regular subgraph. The lemma-- so don't worry too much about the details, and I'll tell you what the idea is. So for every alpha, there exists constants beta and K such that for every C and n sufficiently large, every n-vertex graph G with lots of edges-- so C n to the 1 plus alpha edges-- has a subgraph G prime such that-- so I want some properties.

First, G prime has lots of vertices. So n to the beta, so it's still lots of vertices. You do some polynomial in n. And, 2, it has still lots of edges relative to the number of vertices it has. So, basically, changing the constants, if I start with n to the 1 plus alpha, I still have, roughly, number of vertices to the 1 plus alpha, number of edges.

It is almost regular, in the sense that the max degree of G does not differ from its minimum degree by more than a constant factor. So you do have vertices that are too small degree. You don't have vertices that are too large degree. And, finally, G prime is bipartite, and the two parts of this bipartition have sizes differing by a factor, at most, 2. So if you like, think of G as a regular bipartite graph. But this is the preparation lemma. We'll just make our life a bit easier. So, from now on, let's treat G as a constant in our asymptotic notation to simplify the notation.

So you have this graph G. It's a bipartite graph. And for a pair of vertices on one side A-- so there are no edges in A. But for a pair of vertices, I say that this u, v is light. So it's not an edge, but I talk about these pairs. I say that it is light if the number of common neighbors between-- of u and v is at least 1 and less than t choose 2. So it has some common neighbors, but not too many.

And then we say that this pair is heavy if the number of common neighbors is at least t choose 2. So if a pair u, v has some common neighbors, then it's either light or heavy. I claim that if G is a Kt prime-- so this is the one subdivision of Kt-- free bipartite graph with vertex bipartition A union B-- so U. Not A, but U union B. U will eventually be a subset of A. Such that all the vertices on the left of U have to be at least delta, and U is not too small. So it's at least 4Bt over delta. Don't worry about it for now.

So think of delta-- this is a min degree. So it is somewhat smaller than the average-- I mean, it's basically the average degree of your graph. And B, think of as n. It's more or less the whole set of vertices. Then the conclusion is that there exists a u in a lot of light pairs in the set U. There exists a vertex in many light pairs.

It's important that we assume that this graph G if Kt prime free. Because, otherwise, you could imagine a situation where, essentially, you have a complete bipartite graph and every pair of vertices is heavy. So you don't have any light pairs at all. So having Kt prime free somehow allows us to find light pairs. So let's see the proof of this lemma. So you combine some nice ideas that we've seen earlier in the course, namely double counting, and also uses Turan's theorem.

So, first, let's do a double counting argument similar to the proof of the Kovari-Sos-Turan theorem, where let's count the number of K1,2's like that. So the number of K1,2's like that between U and B. I claim one way to count this is to look through all the vertices on the right side, look at how many neighbors it has, and sum up the degrees choose 2. So skipping some-- I mean, I can tell you what comes out.

So, by convexity, we find that it is at least this quantity here. And then, assuming the minimum degree condition, we find that this quantity is quite large. So this is a calculation very similar to what we did for the proof of Kovari-Sos-Turan theorem.

The low degree vertices in B do not contribute much to the sum. So this sum is large, and it sums over all vertices in B, but the low degree vertices in B, they contribute very little. Because if we sum over all the vertices of B with degree less than 2t, then, for each summand, it's, at most, 2 t squared, which, again, by the assumption of delta, is less than half of the total sum. So the low degree vertices of B do not contribute very much to the sum.

So let's look at the higher degree vertices. For the higher degree vertices, they contribute a substantial chunk. And the most important thing here is that, among these vertices, there are no t mutually heavy-- so not among these vertices, but in U. If you look at U, there are no t mutually heavy vertices in U.

If you have t mutually heavy vertices in U, then what happens? So if you have t mutually heavy vertices. If you had, let's say, three vertices in U, they're mutually heavy, each one of them, because they're heavy, I can find many common neighbors. So I can build this path of length 2. I can build another path of length 2. And I don't run out of vertices because they're all heavy. They all have at least t choose 2 common neighbors. So I can build the subdivision of Kt.

So there are no mutually heavy vertices in U. So where have we seen this before? So you have-- think about the neighborhood of a vertex in B. Because it's inside a neighborhood, all the pairs are either heavy or light, and there are no t mutually heavy vertices. So, then, Turan's theorem tells us that there must be many light vertices. That there are many light vertices in this neighborhood. So the number of light pairs in the neighborhood of this v, if it has at least t neighbors-- or else you run out of room-- is at least-- so if you think about what Turan's theorem says, that the number of known edges is at least this quantity here, which is at least the degree of v squared-- degree of v.

So Turan's theorem tells us that there cannot be so many heavy pairs inside a neighborhood of a vertex in B, so there must be many light pairs. And now we sum over all vertices in B. We obtain that U has a lot of light pairs. We might have overcounted and a little bit, because each light pair is overcounted only by a bounded number of times because it's light. So it's overcounted by less than K choose 2 times. So that's just a constant factor, and we're OK with that.

So that's the conclusion for now. This lemma tells us that you have lots of light pairs in U. And what we're going to do is to keep on shrinking this U. So U is going to be a subset of A. Initially, let's let U be the entire set A. It tells us that there's one vertex in A with lots of light neighbors. Take that vertex, choose its neighborhood. Apply the lemma again. Find another vertex with lots of internal light neighbors. Keep on going. And then we build a large light clique. So that's the idea.

So we'll find that-- so we'll see that if this delta is bigger than, basically, the quantity claimed in the theorem-- so t minus 2 over 2t minus 3-- and sufficiently large C, then there exists the sequence U1, U2, U3, so on, all the way to Ut, and the sequence of vertices v1, v2, vt, such that, initially, I take A. And the idea is, initially, I take v1 to be whatever comes out of that lemma.

And I want the property that all of the vi, vj's are light. And 2 is that no three of these v's have a common neighbor. And, I think, once you have these two properties, then you can find your clique subdivision. You find these t light vertices. So if you have v1, v2, v3, v4, you have these light vertices, and I can build a clique subdivision from these light vertices. Because they're light, they have at least one common neighbor for each pair. I just keep building them. So I build these common neighbors.

Well, you should be somewhat worried that I end up using the same vertex twice. But, of course, that should not be a worry if I guarantee that no three of them have a common neighbor. They cannot collapse. And you cannot have two vertices end up being the same. Otherwise, I would violate property b. So these two properties alone allow you to build a K_t subdivision.

But how do we find this sequence? Well, so we build it iteratively using that lemma. You start with one vertex guaranteed by that lemma. You look at its neighborhood. You pick another vertex guaranteed by the lemma. You look at its neighborhood. And so on. And you build this sequence of light clique. Yes.

AUDIENCE: [INAUDIBLE].

YUFEI ZHAO: Ah. The light neighborhood, yes. So we're not in the graph any more. So everything's inside A. Everything's inside A. So let me finish off the list of properties, and we're almost there. Third property I want is that, when I do this operation, I do not reduce my space of possibilities by too much. Namely, that the size of U does not go down too substantially. And that's guaranteed by the lemma.

And, 4 is that-- basically this picture over here-- v_i is light to all of U_{i+1}. So I claim that you can find the sequence satisfying these properties. And the reason is that you are repeatedly applying the lemma. So repeatedly apply the lemma.

The lemma doesn't address this part about triple vertices having common neighbors. But I claim that's actually not so hard to deal with. Because if you think about how many vertices, how many possibilities, this restriction eliminates, b only eliminates at each step t choose 2, because this is coming from the light restriction, times, at most-- so another t choose 2. And this comes-- the first one comes from pairs of v_1 through v_t. So eliminates, at most, this many-- times the max degree.