

YUFEI ZHAO: Last time, we considered the relationship between pseudo-random graphs and their eigenvalues. And the main message is that the smaller your second largest eigenvalue is, the more pseudo-random a graph is. In particular, we were looking at this class of graphs that are d -regular-- they are somewhat easier to think about.

And there is a limit to how small the second largest eigenvalue can be. And that was given by the Alon-Boppana bound. You should think of d here as a fixed number-- so here, d is a fixed constant. Then, as the number of vertices becomes large, the second largest eigenvalue of a d -regular graph cannot be less than this quantity over here. So this is the limit to how small this second largest eigenvalue can be.

And last time, we gave a proof of this bound by constructing an appropriate function that witnesses this λ_2 . We also gave a second proof which proves a slightly weaker result, which is that the second largest eigenvalue in absolute value is at least this quantity. So in spirit, it amounts to roughly the same result-- although technically, it's a little bit weaker. And that one we proved by counting walks.

And also at the end of last time, I remarked that this number here-- the fundamental significance of this number is that it is the spectral radius of the infinite d -regular tree. So that's why this number is here. Of course, we proved some lower bound. But you can always ask the question, is this the best possible lower bound? Maybe it's possible to prove a somewhat higher bound. And that turns out not to be the case.

So that's the first thing that we'll see today is some discussions-- I won't show you any proofs-- but some discussions on why this number is best possible. And this is a very interesting area of graph theory-- goes under the name of Ramanujan graphs. So I'll explain the history in a second, why they're called Ramanujan graphs. Ramanujan did not study these graphs, but they are called them for good reasons.

So by definition, a Ramanujan graph is a d -regular graph, such that, if you look at its eigenvalue of the adjacency matrix, as above, the second largest eigenvalue in absolute value is, at most, that bound up there-- $2\sqrt{d} - 1$. So it's the best possible constant you could put here so that there still exists infinitely many d -regular Ramanujan graphs for fixed d -- and the size of the graph going to infinity. And the last time, we also introduced some terminology.

Let me just repeat that here. So this is, in other words, an n -regular graph, with λ at most $2\sqrt{d} - 1$.

Now, it is not hard to obtain a single example of a Ramanujan graph. So I just want some graph such that-- or the top eigenvalue is d . I want the other ones to be small. So for example, if you get this link, it's d -regular.

Here, the top eigenvalue is d . And if it's not too hard to compute that all the other eigenvalues are equal to exactly minus 1. So this is an easy computation. But the point is that I want to construct graphs-- I want to understand whether they are graphs where d is fixed. So this is somehow not a good example.

What we really want is fixed d and n going to infinity. Large number of vertices. And the main open conjecture in this area is that for every d , there exists infinitely many Ramanujan d -regular graphs.

So let me tell you some partial results and also explain the history of why they're called Ramanujan graphs. So the first paper where this name appeared and coined this name-- and I'll explain the reason in a second-- is this important result of Lubotzky, Phillips, and Sarnak. From the late '80s. So their paper was titled Ramanujan graphs. So they proved that this conjecture is true.

So the conjecture is true for all d , such that $d - 1$ is a prime number. I should also remark that the same result was proved independently by Margulis at the same time. Their construction of this graph is a specific Cayley graph. So they gave an explicit construction of a Cayley graph, with the group being the projective special linear group-- $PSL(2, q)$.

So, some group-- and this group actually comes up a lot. It's a group with lots of nice pseudo-randomness properties. And to verify that, the corresponding graph has the desired eigenvalue properties, they had to invoke some deep results from number theory that were related to Ramanujan conjectures. So that's why they called these graphs Ramanujan graphs. And that name stuck.

So these papers, they proved that these graphs exist for some special values of d -- namely, when $d - 1$ is a prime. There was a later generalization in the '90s, by Morgenstern, generalizing such constructions showing that you can also take $d - 1$ to be a prime power. And really, that's pretty much it. For all the other values of d , it is open whether there

exists infinitely many d -regular Ramanujan graphs. In particular, for d equal to 7, it is still open.

Do there exist infinitely many semi-regular Ramanujan graphs? No. What about a random graph? If I take a random graph, what is the size of its second largest eigenvalue? And there is a difficult theorem of Friedman-- I say difficult, because the paper itself is more than 100 pages long-- that if you take a fixed d , then a random d -regular graph.

So what does this mean? So the easiest way to explain the random d -regular graph is that you look at the set of all possible d -regular graphs on a fixed number of vertices and you pick one uniformly at random. So random-- such graph is almost Ramanujan, in the following sense-- that the second largest eigenvalue in absolute value is, at most, $2\sqrt{d-1} + \epsilon$, where the ϵ goes to 0 as n goes to infinity.

So in other words, this constant cannot be improved, but this result doesn't tell you that any of these graphs are Ramanujan. Experimental evidence suggests that if you take, for fixed value of d -- let's say d equals to 7 or d equals to 3-- if you take a random d -regular graph, then a specific percentage of those graphs are Ramanujan. So the second largest eigenvalue has some empirical distribution, at least from computer experiments, where some specific fraction-- I don't remember exactly, but let's say 40% of three regular graphs-- is expected in the limit to be Ramanujan. So that appears to be quite difficult. We have no idea how to even approach such conjectures.

There were some exciting recent breakthroughs in the past few years concerning a variant, a somewhat weakening of this problem-- a bipartite analogue of Ramanujan graphs. Now, in a bipartite graph-- so all bipartite graphs have the property that its eigenvalues, its spectrum, is symmetric around 0. Its smallest eigenvalue is minus d .

So if you plot all the eigenvalues, it's symmetric around 0. This is not a hard fact to see-- I encourage you to think about it. And that's because, if you have an eigenvector-- so it lifts somewhere on the left and somewhere on the right-- I can form another eigenvector, which is obtained by flipping the signs on one part. If the first eigenvector has eigenvalue λ , then the second one has eigenvalue minus λ . So the eigenvalues come in symmetric pairs.

So by definition, a bipartite Ramanujan graph is one where it's a bipartite graph and I only require that the second largest eigenvalue is less than $2\sqrt{d-1}$. Everything's symmetric around the origin. So this is by definition.

If you start with a Ramanujan graph, I can use it to create a bipartite Ramanujan graph, because if I look at this 2 lift-- so where there's this construction-- this means if I start with some graph, G -- so for example, if G is this graph here, what I want to do is take two copies of this graph, think about having them on two sheets of paper, one on top of the other. And I draw all the edges criss-crossed. So that's G cross K_2 . This is G .

You should convince yourself that if G has eigenvalues λ then G cross K_2 has eigenvalues. The original spectrum, as well, it's symmetric-- so it's negation. So if G is Ramanujan, then G cross K_2 is a bipartite Ramanujan graph. So it's a weaker concept-- if you have Ramanujan graphs, then you have bipartite Ramanujan graphs-- but not in reverse.

But, still a problem of do there exist d -regular bipartite Ramanujan graphs is still interesting. It's a somewhat weaker problem, but it's still interesting. And there was a major breakthrough a few years ago by Marcus Spielman and Srivastava, showing that for all fixed d , there exist infinitely many d -regular bipartite Ramanujan graphs. And unlike the earlier work of Lubotzky, Phillips, and Sarnak-- which, the earlier work was an explicit construction of a Cayley graph-- this construction here is a probabilistic construction. It uses some very nice tools that they called interlacing families.

So it showed, probabilistically, using a very clever randomized construction, that these graphs exist. So it's not just take a usual d -regular random bipartite graph, but there's some clever constructions of randomness. And this is more or less the state of knowledge regarding the existence of Ramanujan graphs. Again, the big open problem is that there exists d -regular Ramanujan graphs. For every d , there are infinitely many such Ramanujan graphs. Yeah.

AUDIENCE: So in the conception of G cross K_2 , $\lambda + 1$ is equal to d ? Or is equal to original [INAUDIBLE].

YUFEI ZHAO: Right, so the question is, if you start with a d -regular graph and take this construction, the spectrum has $1 + d$ and it also has a minus d . If your graph is bipartite, its spectrum is symmetric around the origin. So you always have d and minus d .

So a bipartite graph can never be Ramanujan. But the definition of a bipartite Ramanujan graph is just that I only require that the remaining eigenvalues sit in that interval. I'm OK with having minus d here. So that's by definition of a bipartite Ramanujan graph. Any more questions?

All right, so combining a Alon-Boppana bound and both the existence of Ramanujan graphs and also Friedman's difficult result that random graph is almost Ramanujan, we see that this $2\sqrt{d-1}$ that number there is optimal. So that's the extent in which a d -regular graph can be pseudo-random. Now, the rest of this lecture, I want to move onto a somewhat different topic, but still concerning sparse pseudo-random graphs. Basically, I want to tell you what I did for my PhD thesis.

So, so far we've been talking about pseudo-random graphs, but let's combine it with the topic in the previous chapter-- namely, similarities regularity lemma. And we can ask, can we apply the regularity method to sparse graphs? So when we talk about similarities regularity, I kept emphasizing that it's really about dense graphs, because there are these error terms which are little and squared. And if your graph is already sparse, that error term eats up everything. So for sparse graphs, you need to be extra careful.

So I want to explore the idea of a sparse regularity. And here, sparse just means not dense. So sparse means ϵ density, little ϵ . So, the opposite of dense. We saw the triangle removal [INAUDIBLE].

So, let me remind you the statement. It says that for every ϵ , there exists some δ , such that if G has a small number of triangles, then G can be made triangle-free by removing a small number of edges. I would like to state a sparse version of this theorem that works for graphs where I'm looking at sub constant ϵ densities.

So roughly, this is how it's going to go. I'm going to put in these extra p factors. And you should think of p as some quantity that goes to 0 with n .

So, think of p as the general scale. So that's the ϵ density scale we're thinking of. I would like to say that if G has less than that many triangles, then G can be made free by deleting a small number of edges. But what does small mean here? Small should be relative to the scale of ϵ densities you're looking at.

So in this case, we should add an extra factor of p over here. So that's the kind of statement I would like, but of course, this is too good to be true, because we haven't really modified anything. If you read the statement, it's just completely false. So I would like adding some conditions, some hypotheses, that would make such a statement true. And hypothesis is going to be roughly along those lines.

So I'm going to call this a meta-theorem, because I won't state the hypothesis precisely. But roughly, it will be along the lines that, if γ is some, say, sufficiently pseudo-random graph and vertices and x density, p , and G is a subgraph of γ , then I want to say that G has this triangle removal property, relatively inside γ . And this is true. Well, it is true if you're putting the appropriate, sufficiently pseudo-random condition. So I'm leaving here-- I'll tell you more later what this should be.

So this is a kind of statement that I would like. So a sparse extension of the triangle removal lemma says that, if you have a sufficiently pseudo-random host, or you think of this γ as a host graph, then inside that host, relative to the density of this host, everything should behave nicely, as you would expect in the dense case. The dense case is also a special case of the sparse case, because if we took γ to be the complete graph-- which is also pseudo-random, it's everything-- it's uniform-- then this is also true. And that's triangle removal among the dense case, but we want this sparse extension. Question.

AUDIENCE: Where does the c come into [INAUDIBLE]?

YUFEI ZHAO: Where does p come in to-- so p here is the edge density of γ .

AUDIENCE: [INAUDIBLE]

YUFEI ZHAO: Yeah. So again, I'm not really stating this so precisely, but you should think of p as something that could decay with n -- not too quickly, but decay at n to the minus some small constant. Yeah.

AUDIENCE: Delta here doesn't depend on γ ?

YUFEI ZHAO: Correct. So the question is, what does delta depend on? So here, delta depends on only epsilon. And in fact, what we would like-- and this will indeed be basically true-- is that delta is more or less the same delta from the original triangle removal lemma. Yeah.

AUDIENCE: If G is any graph, what's stopping you from making it a subgraph of some large [INAUDIBLE]?

YUFEI ZHAO: So the question is, if G is some arbitrary graph, what's to stop you from making it a subgraph of a large pseudo-random graph? And there's a great question. If I give you a graph, G , can you test whether G satisfies the hypothesis? Because the conclusion doesn't depend on γ . The conclusion is only on G , but the hypothesis requires us to γ .

And so my two answers to that is, one, you cannot always embed it in the gamma. I guess the easier answer is the conclusion is false with other hypothesis. So you cannot always embed it in such a gamma. But it is somewhat difficult to test.

I don't know a good way to test whether a given G lies in such a gamma. I will motivate this theorem in a second-- why we care about results of this form. Yes.

AUDIENCE:

Don't all sufficiently large pseudo-random graphs-- say, with respect to the number of vertices of G -- contain copies of every G ?

YUFEI ZHAO:

So the question is, if you start with a sufficiently large pseudo-random gamma, does it contain every copy of G ? And the answer is no, because G has the same number of vertices as gamma. A sufficiently pseudo-random-- again, I haven't told you what sufficiently pseudo-random even means yet. But you should think of it as controlling small patterns.

But here, G is a much larger graph. It's the same size, it's just maybe, let's say, half of the edges. So what you should think about is starting with gamma being, let's say, a random graph. And I delete adversarially, let's say, half of the edges of gamma. And you get G .

So let me go on. And please ask more questions. I won't really prove anything today, but it's really meant to give you an idea of what this line of work is about. And I also want to motivate it by explaining why we care about these kind of theorems.

So first observation is that it is not true with all the hypothesis-- hopefully all of you see this as obviously too good to be true. But will also see some specific examples. Here's a specific example. So this is not true without this gamma. So for example, you can have this graph, G . And we already saw this construction that came from Behren's construction, where you have n vertices and n to the 2 minus little 1 edges, where every edge belongs to exactly one triangle.

If you plug in this graph into this theorem, with all the yellow stuff-- if you add in this p -- you see it's false. You just cannot remove-- anyway. In what context can we expect such a sparse triangle removal lemma to be true?

One setting for which it is true-- and this was a result that was proved about 10 years ago-- is that if your gamma is a truly random graph. So this is true for a random gamma if p is sufficiently large and roughly it's at least-- so there's some constant such that if p is at least c over [INAUDIBLE], then it is true. So this is the result of Conlon and Gowers. Yeah.

AUDIENCE: Is this random in the Erdos-Rényi sense?

YUFEI ZHAO: So this is random in the Erdos-Rényi sense. So, Erdos-Rényi random graph. But this is not the main motivating reason why I would like to talk about this technique. The main motivating example is the Green-Tao theorem. So I remind you that the Green-Tao theorem says that the primes contain arbitrarily long arithmetic progressions.

So the Green-Tao theorem is, in some sense, an extension of Szemerédi's theorem, but a sparse extension. Szemerédi's theorem tells you that if you have a positive density subset of the integers, then it contains long arithmetic progressions. But here, the primes-- we know from prime number theorem that the density of the primes up to n decays, like $1/\log n$. So it's a sparse set, but we would like to know that it has all of these patterns.

It turns out the primes are, in some sense, pseudo-random. But that's a difficult result to prove. And that was proved after Green-Tao proved their initial theorem-- so, by later works of Green and Tao and also Ziegler. But the original strategy-- and also the later strategy for the stronger result, as well.

But the strategy for the Green-Tao theorem is this-- you start with the primes and you embed the primes in a somewhat larger set. You start with the primes and you embed it in a somewhat larger set, which we'll call, informally, pseudoprimes. And these m , roughly speaking, numbers with no small prime divisors.

Because these numbers are somewhat smoother compared to the primes, they're easier to analyze by analytic number theory methods, especially coming from sieve theory. And it is easier, although still highly nontrivial, to show that these pseudoprimes are, in some sense, pseudo-random. And that's the kind of pseudo-random host that corresponds with the gamma over there.

So the Green-Tao strategy is to start with the primes, build a slightly larger set so that the primes sit inside the pseudoprimes in a relatively dense manner. So it has high relative density. And then, if you had this kind of strategy for a sparse triangle removal lemma-- but imagine you also had it for various other extensions of sparse hypergraph removal lemma, which allows you to prove Szemerédi's theorem. And now you can use it in that setting. Then you can prove Szemerédi's theorem in the primes.

That's the theorem and that's the approach. And that's one of the reasons, at least for me,

why something like a sparse triangle removal lemma plays a central role in these kind of problems. So I want to say more about how you might go about proving this type of result and also what pseudo-random graph means over here.

So, remember the strategy for proving the triangle removal lemma. And of course, all of you guys are working on this problem set and so the method of regularity hopefully should be very familiar to you by the end of this week. But let me remind you that there are three main steps, one being to partition your graph using the regularity lemma. The second one, to clean. And the third one, the count.

And I want to explain where the sparse regularity method fails. So you can try to do everything the same and then-- so what happens if you try to do all these things? So first, let's talk about sparse regularity lemma.

So let me remind you-- previously, we said that a pair of vertices is epsilon regular if, for every subset U of A and W of B -- neither too small. So if neither are too small, one has that the number of edges between U and W differs from what you would expect. So the x density between U and W is close to what you expect, which is the ordinal edge density between A and B . So they differ by no more than epsilon.

So this should be a familiar definition. What we would like is to modify it to work for the sparse setting. And for that, I'm going to add in an extra p factor. So I'm going to say epsilon, p regular. Well, this condition here-- now, oh, the densities are on the scale of p . Which goes to 0 as n goes to infinity.

So in what does the property compare them? I should add an extra factor of p to put everything on the right scale. Otherwise, this is too weak. And given this definition here, we can say that a partition of vertices is epsilon regular if all part but at most epsilon fraction pairs is epsilon regular-- so an equitable partition. And I would modify it to the sparse setting by changing the appropriate notion of regular to the sparse version, where I'm looking at scales of p .

I still require at most epsilon fraction-- that stays the same. That's not affected by the density scale. Previously, we had the irregularity lemma, which said that for every epsilon, there exists some M such that every graph has an epsilon regular partition into at most M parts. And the sparse version would say that if your graph has x density at most p -- and here, all of these constants are negotiable.

So when I say p , I really could mean 100 times p . You just change these constants. So if it's most p , then it has an ϵ , p regular partition into at most m parts. Here, m depends only on ϵ .

So previously, I wrote down the sparse triangle removal lemma. And I wrote down the statement and it was false-- with all the additional hypotheses, it was false. It turns out that this is actually true-- the version of the sparse regularity lemma, which sounds almost too good to be true, initially. We are adding in a whole lot of sparsity and sparsity seems to be more difficult to deal with. And the reason why I think sparsity is harder to deal with is that, in some sense, there are a lot more sparse graphs than there are dense graphs.

So let me pause for a second and explain that. It is not true that, in some sense, there are more sparse graphs. Because if you just count-- once you have sparser things, there are fewer of them. But I mean in terms of the actual complexity of the structures that can come up. When you have sparser objects, there's a lot more that can happen.

In dense objects, Szemerédi's regularity lemma tells us, in some sense, that the amount of complexity in the graph is bounded. But that's not the case for sparse graphs. In any case, we still have some kind of sparse regularity lemma. And this version here, as written, is literally true if you have the appropriate definitions-- and more or less, we have those definitions up there. But I want to say that it's misleading. This is true, but misleading.

And the reason why it is misleading is that, in a sparse graph, you can have lots of intricate structures that are hidden in your irregular parts. It could be that most edges are inside irregular pairs, which would make the irregularity lemma a somewhat useless statement, because when you do the cleaning step, you delete all of your edges. And you don't want that.

But in any case, it is true-- and I'll comment on the proof in a second. But the way I want you to think about the sparse regularity lemma is that it should work when-- so before jumping to that, a specific example where this happens is, for example, if your graph G is a click on a sublinear fraction of vertices. Somehow, you might care about that click. So that's a pretty important object in the graph. But when you do the sparse regularity partition, it could be that the entire click is hidden inside an irregular part. And you just don't see it-- that information gets lost.

The proper way to think about the sparse regularity lemma is to think about graphs, G , that satisfy some additional hypotheses. So in practice, G is assumed to satisfy some upper

regularity condition. And an example of such an hypothesis is something called no dense spots, meaning that it doesn't have a really dense component, like in the case of a click on a very small number of vertices.

So no dense spots-- one definition could be that there exists some η -- and here, just as in quasi-random graphs, I'm thinking of sequences going to 0. So there exists η sequence going to 0 and a constant, c , such that for all set x in the graph-- let's say X and Y . If X and Y have size at least η fraction of V , then the density between X and Y is bounded by at most a constant factor, compared to the overall density, p , that we're looking at. So in other words, no small piece of the graph has too many edges.

And with that notion of the no dense spots, we can now prove the sparse regularity lemma under that additional hypothesis. And basically, the proof is the same as the usual semi-regularity lemma proof that we saw a few weeks ago. So if you have proof of sparse regularity with no dense spots-- hypothesis.

OK, so I claim this as the same proof as Szemerédi's irregularity lemma. And the reason is that in the energy increment argument, you do everything the same. You do partitioning if it's not regular. You refine and you keep going.

In the energy increment argument, one key property we used was that the energy was bounded between 0 and 1. And every time, you went up by ϵ to the fifth. And now the energy increment argument-- that each step, the energy goes up by something which is like ϵ , let's say, to the fifth and p squared. The energy is some kind of mean square density, so this p squared should play a role.

So if you only knew that, then the number of iterations might depend on p -- it might depend on n . So, not a constant-- and that would be an issue. However, if you have no dense spots-- so, because no dense spots-- the final energy, I claim, is, at most, something like C squared p squared. Maybe some small error, because of all the ϵ s flowing around, but that's the final energy.

So you still have a bounded number of steps. So the bound only depends on ϵ . So the entire proof runs through just fine.

OK, so having the right hypothesis helps. But then I said, the more general version without the hypothesis is still true. So how come that is the case? Because if you do this proof, you run

into the issue-- you cannot control the number of iterations.

So here's a trick introduced by Alex Scott, who came up with that version there. So this is a nice trick, which is that, instead of using x^2 -- the function as energy-- let's consider a somewhat different function. So the function I want to use is f_e of x which is initially quadratic-- so, initially x^2 -- but up to a specific point. Let's say 2. And then after this point, I make it linear.

So that's the function I'm going to take. Now, this function has a couple of nice properties. One is that you also have this boosting, this energy increment step, because for all random variables, x -- so x is a non-negative random variable. So think of this as edge densities between parts on the refinement. If the mean of x is, at most, 1, then, if you look at this energy, it increases if x has a large variance.

Previously, when we used f_e as square, this was true. So this is true with equal to 1-- in fact, that's the definition of variance. But this inequality is also true for this function, f_e -- so that when you do the irregularity breaking, if you have irregular parts, then you have some variance in the edge densities.

So you would get an energy boost. But the other thing is that we are no longer worried about the final energy being much higher than the individual potential contributions. Because, if you end up having lots of high density pieces, they would contribute a lot.

So, in other words, the expectation for the second thing is that the expectation of f_e is upper-bounded by, let's say, 4 times the expectation of x . And so this inequality there would cap the number of steps you would have to do. You would never actually end up having too many iterations.

So this is a discussion of the sparse regularity lemma. And the main message here is that the regularity lemma itself is not so difficult-- that's largely the same as Szemerédi's regularity lemma. And so that's actually not the most difficult part of sparse triangle removal lemma. The difficulty lies in the other step in the regularity method-- namely, the counting step. And we already alluded to this in the past.

The point is that there is no counting lemma for sparse regular graphs. And we already saw an example where, if you start with a random graph which has a small number of triangles and I delete a small number of edges corresponding to those triangles-- one, I do not affect its

quasi-randomness. But two, there's no triangles anymore, so there's no triangle counting lemma. And that's a serious obstacle, because you need this counting step.

So what I would like to explain next is how you can salvage that and use this hypothesis here written in yellow to obtain a counting lemma so that you can complete this regularity method that would allow you to prove the sparse triangle removal lemma. And a similar kind of technique can allow you to do the Green-Tao theorem. So let's take a quick break.

OK, any questions so far. So let's talk about the counting lemma. So, the first case of the counting lemma we considered was the triangle counting lemma.

So remember what it says. If you have 3 vertex sets-- V_1, V_2, V_3 -- such that, between each pair, it is epsilon regular. And edge density-- that's for simplicity's sake-- they all have the same edge density.

Actually, they can be different. So d_{ij} -- so possibly different edge densities. But I have the set-up. And then the triangle counting lemma tells us that the number of triangles with one vertex in each part is basically what you would expect in the random case-- namely, multiplying these three edge densities together, plus a small error, and then multiplying the vertex sets' sizes together.

So what we would like is a statement that says that if you have epsilon p regular and x densities now at scale p, then we would want the same thing to be true. Here, I should add an extra p^3 , because that's the densities we're working with. And I want some error here-- OK, I can even let you take some other epsilon. But small changes are OK.

So that's the kind of statement we want-- and this is false. So this is completely false. And the example that I said earlier was one of these examples where you have a random graph. So this initial version is false, because if you take a G and p , with p somewhat less than $1/\sqrt{n}$, and then remove an edge from each triangle-- or just remove all the triangles-- then you have a graph which is still fairly pseudo-random, but it has no triangles. So you cannot have a counting lemma.

So there's another example which, in some sense, is even better than this random example. And it's a somewhat mysterious example due to a law that gives you a pseudo-random gamma. So it's, in some sense, an optimally pseudo-random gamma, such that it is d -regular with d on the order of $n^{3/2}$.

And it's an n^d lambda graph, where lambda is on the order of \sqrt{d} . Because here, d is not a constant. But even in this case, roughly speaking, this is as pseudo-random as you can expect. So the second eigenvalue is roughly square root of the degree. And yet, this graph is triangle free.

So you have some graph which, for all the other kinds of pseudo-randomness is very nice. So it has all the nice pseudo-randomness properties, yet it is still triangle free. It's sparse. So the triangle counting lemma is not true without additional hypotheses.

So I would like to add in some hypotheses to make it true. And I would like a theorem. So again, I'm going to put as a meta-theorem, which says that if you assume that G is a subgraph of a sufficiently pseudo-random Γ and Γ has edge density p , then the conclusion is true. And this is indeed the case.

And I would like to tell you what is the sufficiently pseudo-random-- what does that hypothesis mean? So that at least you have some complete theorem to take. There are several versions of this theorem, so let me give you one which I really like, because it has a fairly clean hypothesis. And the version is that the pseudo-randomness condition-- so here it is.

So, a sufficient pseudo-randomness hypothesis on Γ , which is that Γ has the correct number-- "correct" in quotes, because this is somewhat normative. So what I'm really saying is it has, compared to a random case, what you would expect. Densities of all subgraphs of $K_{2,2,2}$.

Having correct density of H means having H density. $1 + \epsilon$ times p , raised to the number of edges of H , which is what you would expect in a random case. So you should think of there, again, not being just one graph, but a sequence of graphs. You can also equivalently write it down in terms of δ s and ϵ s having error parameters. But I like to think of it having a sequence of graphs, just as in what we did for quasi-random graphs.

If your Γ has this pseudo-randomness condition, which is we're in this sparse setting. So if you try to compare this to what we did for quasi-random graphs, you might get confused. Because there, having the correct C_4 count already implies everything. This condition, it actually does already include having the correct C_4 count. So $K_{2,2,2}$ is this graph over here.

And I'm saying that if it has the correct density of H , whenever H is a subgraph, of $K_{2,2,2}$ --

then it has a correct density. So in particular, it already has a C_4 count, but I want more. And it turns out this is genuinely more, because in a sparse setting, having the correct C_4 count is not equivalent to other notions of pseudo-randomness. So this is a hypothesis. So if I start with a sequence of gammas, I have the correct counts of $K_{2,2,2}$ s as well as subgraphs of $K_{2,2,2}$. Then I claim that that pseudo-random host is good enough to have a counting lemma-- at least for triangles. Any questions?

Now, you might want to ask for some intuitions about where this condition comes from. The proof itself takes a few pages. I won't try to do it here. I might try to give you some intuition how the proof might go and also what are the difficulties you might run into when you try to execute this proof.

But, at least how I think of it is that this $K_{2,2,2}$ condition plays a role similar to how previously, in dense quasi-random graphs, we had this somewhat magical looking C_4 condition, which can be viewed as a doubled version of an edge. So actually, the technical name is called a blow-up. It's a blow-up of an edge. Whereas the $K_{2,2,2}$ condition is a 2 blow-up of a triangle.

And this 2 blow-up hypothesis is some kind of a graph theoretic analogue of controlling second moment. Just as knowing the variance of a random variable-- knowing its second moment-- helps you to control the concentration of that random variable, showing that it's fairly concentrated. And it turns out that having this graphical second moment in this sense also allows you to control its properties so that you can have nice tools, like the counting lemma.

So let me explain some of the difficulties. If you try to run the original proof of the triangle removal lemma for the sparse setting, what happens? So if you start with a vertex-- so remember how the proof of triangle removal lemma went. You start with this set-up and you pick a typical vertex. This typical vertex has lots of neighbors to the left and lots of neighbors to the right. And here, a lot means roughly the edge density times the number of vertices-- and a lot of vertices over here.

And then you say that, because these are two fairly large vertex sets, there are lots of edges between them by the hypotheses on epsilon regularity, between the bottom two sets. But now, in the sparse setting, we have an additional factor of p . So these two sets are now quite small.

They're much smaller than what you can guarantee from the definition of epsilon, p regular. So you cannot conclude from them being epsilon regular that there are enough edges

between these two very small sets. So the strategy of proving the triangle removal lemma breaks down in the sparse setting.

In general-- not just for triangles, but for other H 's as well-- we also have this counting lemma. So, the sparse counting lemma. And also the triangle case, which I stated earlier. So this is drawing work due to David Conlon, Jacob Fox, and myself. Says that there is a counting lemma.

So let me be very informal. So, that there exists a sparse counting lemma for counting H , in this set-up as before. If Γ has a pseudo-random property of containing the correct density of all subgraphs of the 2 -blow-up of H .

Just as in the triangle, the 2 -blow-up is $K_{2,2,2}$. In general, the 2 -blow-up takes a graph, H , and then doubles every vertex and puts in four edges between each pair of vertices. So that's the 2 -blow-up of H .

If your Γ has pseudo-random properties concerning counting subgraphs of this 2 -blow-up, then you can obtain a counting lemma for H itself. Any questions? OK, so let's take this counting lemma for granted for now. How do we proceed to proving the sparse triangle removal lemma?

Well, I claim that actually it's the same proof where you run the usual simulated regularity proof of triangle removal lemma. But now, with all of these extra tools and these extra hypotheses, you then would obtain the sparse triangle removal lemma, which I stated earlier. And the hypothesis that I left out-- the sufficiently pseudo-random hypothesis on Γ -- is precisely this hypothesis over here, as required by the counting lemma.

And once you have that, then you can proceed to prove a relative version of Roth's theorem-- and also, by extension, two hyper-graphs-- also a relative version of Szemerédi's theorem. So, recall that the Roth's theorem tells you that if you have a sufficiently large-- so let me first write down Roth's theorem. And then I'll add in the extra relative things in yellow. So if I start with A , the subset of $\mathbb{Z}/N\mathbb{Z}$, such that A has size-- at least δn .

So then, Roth's theorem tells us that A contains at least one three-term arithmetic progression. But actually, you can boost that theorem. And you've seen some examples of this in homework. And also our proofs also do this exact same thing. If you look at any of the proofs that we've seen so far, it tells us that A not only contains one single $3A_p$, but it contains many $3A_p$'s, where C is some number that is positive.

So you can obtain this by the versions we've seen before, either by looking at a proof-- problem is in the the proof gift stack-- or by using the black box version of Roth's theorem. And then there's a super saturation argument, which is similar to things you've done in the homework. What we would like is a relative version. And a relative version will say that if you have a set, S , which is sufficiently pseudo-random.

And S has density, p . Here, [INAUDIBLE]. And now A is a subset of S . And A has size at least δ , that of S . Then, A contains still lots of $3A$'s, but I need to modify the quantity, because I am looking at density, p .

So this statement is also true if you're putting the appropriate hypothesis into sufficiently pseudo-random. And what should those hypotheses be? So think about the proof of Roth's theorem-- the one that we've done-- where you set up a graph.

So, you set up this graph. So, one way to do this is that you say that you put in edges between the three parts-- x , y , and z . So the vertex sets are all given by $z \pmod N$. And you put in an edge between x and y , if $2x$ plus y lies in S .

Pulling the edge between x and z -- if x minus z lies in S and a third edge between y and z , if y minus $2z$ lies in S . So this is a graph that we constructed in the proof of Roth's theorem. And when you construct this graph, either for S or for A -- as we did before-- then we see that the triangles in this graph correspond precisely to the $3A$'s in the set.

So, looking at the triangle counting lemma and triangle removal lemma-- the sparse versions-- then you can read out what type of pseudo-randomness conditions you would like on S -- so, from this graph. So, we would like a condition, which says that this graph here-- which we'll call γ sub S -- to have the earlier pseudo-randomness hypotheses. And you can spell this out. And let's do that.

Let's actually spell this out. So what does this mean? What I mean is S , being a subset of $\mathbb{Z} \pmod N$ -- we say that it satisfies what's called a 3-linear forms condition. If, for uniformly chosen random $x_0, x_1, y_0, y_1, z_0, z_1$ elements of $\mathbb{Z} \pmod N$.

Think about this $K_{2,2,2}$. So draw a $K_{2,2,2}$ up there. So what are the edges corresponding to the $K_{2,2,2}$? So they correspond to the following expressions-- x_0 minus $2z_0$ -- x_1 minus $2z_0$ -- x_0 minus $2z_1$ -- x_1 minus $2z_1$.

So those are the edges corresponding to the bottom. Draw C_4 across the bottom two vertex sets. But then there are two more columns. And I'll just write some examples, but you can fill in the rest.

OK, so there are at least 12 expressions. And what we would like is that, for random, the probability that all of these numbers are contained in S is within $1 \pm \epsilon$ factor of the expectation, if S were a random set. In other words, in this case, it's p raised to 12-- random set of density, p . And furthermore, the same holds if any subset of these 12 expressions are erased.

Now, I want you to use your imagination and think about what the theorem would look like for not $3A_p$'s, but for $4A_p$'s-- and also for kA_p 's in general. So there is a relative Szemerédi theorem, which tells you that if you start with S -- so here, we fix K . If you start with this S , that satisfies the k -linear forms condition. And A is a subset of S that is fairly large.

Then A has kA_p . So I'm being slightly sloppy here, but that's the spirit of the theorem-- that you have this Szemerédi theorem inside a sparse pseudo-random set, as long as the pseudo-random set satisfies this k -linear forms condition. And that k -linear forms condition is an extension of this 3-linear forms condition, where you write down the proof that we saw for Szemerédi's theorem, using hyper-graphs. Write down the corresponding linear forms-- you expand them out and then you write down this statement.

So this is basically what I did for my PhD thesis. So we can ask, well, what did Green and Tao do? So they had the original theorem back in 2006. So their theorem, which also was a relative Szemerédi theorem, has some additional, more technical hypotheses known as correlation conditions, which I won't get into.

But at the end of the day, they constructed these pseudoprimes. And then they verified that those pseudoprimes satisfied these required pseudo-randomness hypotheses-- that those pseudoprimes satisfied these linear forms conditions, as well as their now-extraneous additional pseudo-randomness hypotheses. And then combining this combinatorial theorem with that number adiabatic result. You put them together, you obtain the Green-Tao theorem, which tells you not just that the primes contain arbitrarily long arithmetic progressions, but any positive density subset of the primes also contains arbitrarily long arithmetic progressions.

All of these theorems-- now, if you pass down to a relatively dense subset, it still remains true. Any questions? So this is the general method.

So the general method is you have the sparse regularity method. And provided that you have a good counting lemma, you can transfer the entire method to the sparse setting. But getting the counting lemma is often quite difficult. And there are still interesting open problems-- in particular, what kind of pseudo-randomness hypotheses do we really need? Another thing is that you don't actually have to go through regularity yourself.

So there is an additional method-- which, unfortunately I don't have time to discuss-- called transference, where the story I've told you is that you look at the proof of Roth's theorem, the proof of Szemerédi's theorem. And you transfer the methods of those proofs to the sparse setting. And you can do that. But it turns out, you can do something even better-- is that you can transfer the results. And this is what happens in Green-Tao.

If you look at Szemerédi's theorem as a black-box theorem and you're happy with its statement, you can use these methods to transfer that result as a black box without knowing its proof to the sparse pseudo-random setting. And that sounds almost too good to be true, but it's worth seeing how it goes. And if you want to learn more about this subject, there's a survey by Colin Fox and by myself called "Green-Tao theorem, an exposition."

So, where you'll find a self-contained complete proof of the Green-Tao theorem, except no modulo-- the proof of Szemerédi's theorem, which we've called as a black box. But you'll see how the transference method works there. And it involves many of the things that we've discussed so far in this course, including discussions of the regularity method, the counting lemma. And it will contain a proof of this sparse triangle counting lemma. OK, good. We stop here.