Regularity Lemma and its Applications

Tibor Szabó

Extremal Combinatorics, FU Berlin, WiSe 2017–18

1 Introduction

The high level intuitive meaning of the Regularity Lemma of Szemerédi is that **every** graph is made up of three parts: a *structured* part, a *quasirandom* part, and an *error* part. The size of the structured part, the quality of quasirandomness, and the size of the error all depend on a parameter $\epsilon > 0$ that can be chosen arbitrary small. More concretely, the Regularity Lemma states that given any $\epsilon > 0$ **every** graph can be approximated by the disjoint union of constantly many "random-like" bipartite graphs. Crucially, the number of these bipartite graphs depends only on ϵ, and so is how closely they resemble a truly random graph. Furthermore all but a tiny fraction of all possible pairs are contained in one of the quasirandom bipartite graphs.

Our concept of quasirandomness of a bipartite graph will be formulated through the distribution of edges. It will require that, just like in a typical random graph, the edges are distributed evenly all over the graph, that there are no unnaturally dense parts and there are no unnaturally sparse parts either. That is, for any pair of large enough subsets of the two sides the fraction of pairs that are edges is roughly the same as the fraction of pairs that are edges in the whole bipartite graph.

Definition 1.1. *For a graph G and two disjoint subsets $A, B \subseteq V(G)$ os the vertices let $e(A,B) = |\{(a,b) \in A \times B : ab \in E(G)\}|$ denote the number of edges between A and B and let*

$$d(A,B) := \frac{e(A,B)}{|A| \cdot |B|}$$

denote the density *of the pair $\{A,B\}$. The pair $\{A,B\}$ is called* ϵ-regular *if for every $A' \subseteq A$ with $|A'| \geq \epsilon|A|$ and every $B' \subseteq B$ with $|B'| \geq \epsilon|B|$ we have*

$$|d(A',B') - d(A,B)| \leq \epsilon.$$

Remarks.
1. Note that $0 \leq d(A,B) \leq 1$ for every A, B.
2. The definition becomes trivial for $\epsilon = 1$ since then *every* pair is 1-regular.
3. The lower bound on the sizes of the subsets A' and B' is necessary in order to have a meaningful definition. Taking it to the extreme, if we were to require the density condition to hold for every pair of subsets $A' \subseteq A$ and $B' \subseteq B$, then only the complete bipartite graph or the empty bipartite graph would be ϵ-regular with some $\epsilon < \frac{1}{2}$. Indeed the density of a pair of 1-element subsets is 0 or 1, so one of them would be at least $\frac{1}{2}$-away from $d(A,B)$.
4. To see a concrete example of this definition: if a bipartite graph G on vertex set $A \cup B$ with $|A| = |B| = k$ is $\frac{1}{100}$-regular and has $0.37k^2$ edges, then for **every** subsets $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq \frac{|A|}{100}$ and $|B'| \geq \frac{|B|}{100}$ the number of edges between A' and B' is between $0.36|A'| \cdot |B'|$ and $0.38|A'| \cdot |B'|$.
5. The property of being ϵ-regular becomes stronger and stronger as ϵ decreases. Namely for any $\epsilon > \epsilon' > 0$, a graph G being ϵ'-regular implies that it is also ϵ-regular. This is immediate because checking ϵ-regularity requires checking a *weaker* inequality holds for *less* subsets than checking ϵ'-regularity.

6. It is justified to think of ϵ-regularity as a "quasirandom property" as in the next homework you will show that it holds with probability tending to 1 for the random bipartite graph with edge probability p.

HW *The random bipartite graph $G(A, B, p)$ on vertex set $A \cup B$ is the probability space where each pair $(a, b) \in A \times B$ is an edge with probability p, such that all these choices are mutually independent. For every $\epsilon > 0$, a graph sampled from $G(A, B, p)$ is ϵ-regular with probability tending to 1 (as the sizes $|A| = |B| = k$ tend to infinity).*
Hint: You should apply the union bound combined with a Chernoff bound.

7. If you were wondering whether besides the above somewhat existential statement that "almost all bipartite graphs are ϵ-regular", are there any "concrete", "palpable" ϵ-regular graphs. Here is an example, the Payley biapartite graph. For each prime p we define a bipartite graph BP_p with vertex set $L \cup R$, where both L and R are copies of the p-element field $\mathbb{F}_p$. The vertex $a \in L$ is adjacent to the vertex $b \in R$ if $a - b$ is a *quadratic residue* (i.e., a square) in $\mathbb{F}_p$. Exactly $\frac{p-1}{2}$ of the non-zero elements of $\mathbb{F}_p$ are squares, hence BP_p has density $\frac{1}{2}$. In next semester's Constructive Combinatorics course we will show that BP_p is ϵ-regular for every $\epsilon > 0$ and large enough p.

Definition 1.2. *Let $G = (V, E)$ be a graph and $\epsilon > 0$ a real number. A partition $V = V_0 \cup V_1 \cup \cdots \cup V_k$ is called an* ϵ-regular partition *if*

- $|V_0| \leq \epsilon |V|$
- $|V_1| = |V_2| = \cdots = |V_k|$
- *at most $\epsilon \binom{k}{2}$ pairs $\{V_i, V_j\}$ are not ϵ-regular*

Again, smaller ϵ means a stronger property, and more control over the edges of G. Formally, for any $\epsilon > \epsilon' > 0$ an ϵ'-regular partition is also an ϵ-partition.

Theorem 1.3 (Regularity Lemma (Szemerédi, 1975))**.** *For every real $\epsilon > 0$ and positive integer $m \in \mathbb{N}$, there exists an integer $M = M(\epsilon, m)$ such that every graph $G = (V, E)$ with at least m vertices has an ϵ-regular partition $V = V_0 \cup V_1 \cup \cdots \cup V_k$ where $m \leq k \leq M$.*

Remarks.
1. The purpose of the Regularity Lemma is to gain some sort of control over "most edges" of G. In particular the lemma gives a lot of information about the placement of edges between the two sides of an ϵ-regular pair, but it gives essentially no information about the other edges. These include edges incident to the exceptional set V_0, edges inside the parts V_i, $1 \leq i \leq k$, and edges between pairs that are not ϵ-regular.
2. The real number $\epsilon > 0$ clearly represents some sort of upper bound on the error in this approximation of G, but what is the role of m? Without requiring at least m parts in the partition, the Lemma becomes trivially true: the partition $V = V_1$ is clearly ϵ-regular for every $\epsilon > 0$. This statement however is not very useful: it does not give any information about any set of edges of G. Choosing m a large constant in applications, say $m = \lceil \frac{1}{\epsilon} \rceil$, will ensure that only a small fraction of the edges are inside the parts V_i.
3. It is worthwhile to think over a proof of the Lemma for a graph with $k \leq M$ vertices. In this case we can choose our partition to contain only 1-element subsets. This is clearly an ϵ-regular partition for any $\epsilon > 0$, since any pair of 1-element subsets are trivially ϵ-regular for any $\epsilon > 0$.
4. An important issue about the Regularity Lemma is whether the various imperfections contained in its statement are indeed necessary.

- The presence of the exceptional vertex set V_0 is not crucial. One can state a lemma without it if one relaxes the condition of equipartition and allows that two parts differ in size by at most 1. In many aplication however it is easier to work with completely equal parts and rather allow the exceptional part.

- The presence of the exceptional set of $\epsilon\binom{k}{2}$ non ϵ-regular pairs is necessary. A construction showing this will be discussed in the Homework.

5. The main point of the Regularity Lemma is the guarantee of the partition into *at most constant* number of parts, where this constant depends only on ϵ and m, but **not** on the number of vertices of G. So once you decided on the error you can tolerate in the approximation, the number of parts is bounded, no matter how large your input graph is. With this let us revisit the high level intuitive interpretation of the Regularity Lemma mentioned at the beginning, that **every** graph can be decomposed into three parts

- a *structured* part (of constant size)
- a *quasirandom* part
- an *error* part (of small size)

To make a more concrete sense of this heuristic, let us define the *regularity graph* $R = R(\mathcal{P})$ of an ϵ-regular partition $\mathcal{P} = \{V_0, V_1, \dots, V_k\}$ of the graph G on n vertices into k parts, with $m \leq k \leq M = M(\epsilon, m)$. The vertex set of R is just $[k]$, representing the indices of the non-exceptional parts of the partition $\mathcal{P}$. The edges of R correspond to the ϵ-regular pairs: $E(R) := \{ij : \{V_i, V_j\} \text{ is } \epsilon\text{-regular}\}$.

The *quasirandom part* of the heuristic "decomposition" of G consists of the (at most $\binom{k}{2}$) ϵ-regular graphs. The *structured part* of this decomposition is the regularity graph itself: R has constantly many vertices and describes the structure how the quasirandom pieces are put together. Finally, the error part is represented by the graph containing those edges we have essentially no information about. Let us count the edges in this "error part". The number of edges

- incident to V_0 is at most $|V_0|(n-1) \leq \epsilon n^2$;
- inside parts: $k\binom{|V_i|}{2} \leq k\binom{n/k}{2} \leq \frac{n^2}{2k} \leq \frac{n^2}{2m}$;
- between irregular pairs: $\epsilon\binom{k}{2}|V_i| \cdot |V_j| \leq \epsilon\frac{k^2}{2}\left(\frac{n}{k}\right)^2 = \frac{\epsilon}{2}n^2$;

All together these are at most $\left(\frac{3}{2}\epsilon + \frac{1}{2m}\right)n^2$ edges. We could force this error part to be as small a fraction of n^2 as we wish by choosing ϵ small enough and m large enough.

Before giving the proof of the Regularity Lemma we consider a couple of its most important applications. The first one is a relatively simple, yet extremely useful consequence, called the Triangle Removal Lemma. In order to demonstrate the power of the Regularity Lemma we will also discuss a couple of striking consequences of the Triangle Removal Lemma: one from number theory, and one from algorithms. Our second application of the Regularity Lemma will be slightly more technical: we derive a proof of the Erdős-Stone Theorem (which we already announced).

1.1 Application 1: Property Testing and the Triangle Removal Lemma

Suppose you are given a huge graph on n and you would need to decide whether it is triangle-free.[1] You have the possibility to query the status of any pair of vertices in the graph. To identify a graph as triangle-free might require for you to query the status of $\Theta(n^2)$ pairs, which potentially can take a lot of time if n is huge. Having no time for this, you are eager to gain at least some approximate information about the triangle-freeness of the graph, and gain that fast. You will be after identifying those graphs that are *far from being triangle-free* in the sense that one would need to delete *many* edges from the graph to make it triangle-free.

The Triangle Removal Lemma is the backbone of a randomized testing algorithm which does exactly this in constant(!) time. So the number of queries does *not* depend on the size of the graph. Here we first state and prove the lemma and formulate the algorithmic consequence as a homework exercise.

[1] Say, to decide whether the examined genome sequence contains an undesirable mutation.

Theorem 1.4 (Triangle Removal Lemma). *For every $\gamma > 0$ there exists $\delta = \delta(\gamma) > 0$ such that for every graph G at least one of the following is true.*

(A) there exists $F \subseteq E(G)$, $|F| \leq \gamma v(G)^2$, such that $G - F$ is triangle-free

(B) G contains at least $\delta v(G)^3$ triangles

Proof. Given an arbitrary $\gamma > 0$, we use the Regularity Lemma with $\epsilon = \frac{\gamma}{3}$ and $m = \lceil \frac{1}{\epsilon} \rceil = \left\lceil \frac{3}{\gamma} \right\rceil$ and receive $M = M(\epsilon, m)$. Then we define

$$\delta = (1 - 2\epsilon)\epsilon^3 \left(\frac{1-\epsilon}{M} \right)^3 .$$

Of course how the choice of these parameters came about will only be obvious from the proof. In fact when one does the proof the first time, one does not set these parameters, but one sees what is necessary for the various proof steps to go through.

Let G be an arbitrary graph and let us denote its number of vertices by $n = v(G)$. We use the Regularity Lemma with the above defined parameters ϵ and m to obtain an ϵ-regular partition $\mathcal{P} = \{V_0, V_1, \dots, V_k\}$, with $m \leq k \leq M$.

We create a somewhat sparsified version of the regularity graph $R(\mathcal{P})$ corresponding to the partition $\mathcal{P}$. For a density parameter d we define the regularity graph $R(\mathcal{P}, d) \subseteq R(\mathcal{P})$ on the vertex set $[k]$ as the subgraph of $R(\mathcal{P})$ where we keep only those edges $ij \in E(R(\mathcal{P}))$, where the corresponding ϵ-regular pair $\{V_i, V_j\}$ of the partition $\mathcal{P}$ has density $d(V_i, V_j) \geq d$. Note that the original regularity graph $R(\mathcal{P})$ is just $R(\mathcal{P}, 0)$.

For our particular purpose we choose this "minimum-density" $d = 2\epsilon$ and set $R = R(\mathcal{P}, 2\epsilon)$. We classify G according to whether R contains a triangle or not. We will show that if R contains a triangle then G itself contains many triangles, while if G is triangle-free then G can be made triangle-free with the removal of some γn^2 edges.

Case 1. R is triangle-free. We show that there is a set $F \subseteq E(G)$ of at most γn^2 edges of G, such that $G - F$ is triangle-free. Let F consists of all edges of G that are

- incident to the exceptional part V_0
- contained in one of the parts V_i for some i, $1 \leq i \leq k$
- go between a pair V_i, V_j that is not ϵ-regular
- go between a pair V_i, V_j that has density $d(V_i, V_j) < d = 2\epsilon$

We have already counted above that the number of edges of G in the first three parts is not more than $\left(\frac{3}{2}\epsilon + \frac{1}{2m}\right) n^2 \leq 2\epsilon n^2$. We still need to count the edges that are going in between parts with density $< d = 2\epsilon$. There are at most $\binom{k}{2}$ such pairs, and between each there are at most $d \cdot \frac{n}{k} \cdot \frac{n}{k}$ edges. All together this means at most $\frac{k^2}{2} d \frac{n^2}{k^2} = \epsilon n^2$ edges. Therefore F contains at most $3\epsilon n^2$ edges.

Let us see now why the removal of F makes G triangle-free.

Since in $G - F$ the exceptional set consists of isolated vertices any triangle would have to have its vertices in $\cup_{i=1}^{k} V_i$. Since there are no edge inside any V_i, any triangle would have to use vertices from three different parts. By the definition of our R and F, there is an edge in $G - F$ between two parts V_i and V_j if and only if $ij \in R$. Since R is triangle-free there are no three different parts V_i, V_j and V_k such that between all pairs there are edges in $G - F$. Consequently $G - F$ is triangle-free.

Case 2. R contains a triangle, say $ab, bc, ca \in E(R)$. We show that there are at least δn^3 triangles going across in $V_a \cup V_b \cup V_c$ (see Figure 1). This is an immediate consequence of the following lemma.

Lemma 1.5. *Let $A, B, C \subseteq V(G)$ are pairwise disjoint independent sets in some graph G. If the pairs $\{A, B\}, \{B, C\}, \{A, C\}$ are all ϵ-regular then the number of triangles in G is at least*

$$(1 - 2\epsilon)(d(A, B) - \epsilon)(d(A, C) - \epsilon)(d(B, C) - \epsilon)|A| \cdot |B| \cdot |C|.$$

Indeed, in our case a, b, c form a triangle in R, so each of the pairs of V_a, V_b, V_c is ϵ-regular and for their density we have $d(V_a, V_b), d(V_b, V_c), d(V_c, V_a) \geq 2\epsilon$. So the lemma is applicable and the number of triangles is at least

$$(1-2\epsilon)\epsilon^3|V_a||V_b||V_c| \geq (1-2\epsilon)\epsilon^3\left(\frac{(1-\epsilon)n}{k}\right)^3 \geq (1-2\epsilon)\epsilon^3\left(\frac{(1-\epsilon)n}{M}\right)^3 = \delta n^3$$

□

Proof of Lemma 1.5. We will classify triangles based on their vertex in C. For any $v \in C$, the number of triangles containing v is equal to the number of edges $e(N(v,A), N(v,B))$ between the neighborhood $N(v,A) = N(v) \cap A$ of v in A and the neighborhood $N(v,B) = N(v) \cap B$ of v in B. We can estimate this using ϵ-regularity provided the neighborhoods of v in A and B are both large enough (at least ϵ-fraction of A and B, respectively). We will now see that the ϵ-regularity of the pairs $\{A, C\}$ and $\{B, C\}$ implies that *most* vertices $v \in C$ have large neighborhoods both in A and B.

The ϵ-regularity condition regulates the number of edges between *large enough* subsets, and does not in particular require anything about the degrees of individual vertices. So for example the existence of a few isolated vertices is not impossible in a large ϵ-regular graph. However in the next lemma we will see that *most* vertices in an ϵ-regular graph must have approximately the right degree. This lemma is very useful in many application of the Regularity Lemma, we state it under a bit more general conditions.

Lemma 1.6 (Degree Lemma)**.** *Let $A, C \subseteq V(G)$ be an ϵ-regular pair of density at least d in some graph G and let $Y \subseteq A$ with $|Y| \geq \epsilon|A|$. Then*

$$|\{c \in C : |N(c,Y)| < (d-\epsilon)|Y|\}| < \epsilon|C|.$$

Proof. Let $X = \{c \in C : |N(c,Y)| < (d-\epsilon)|Y|\}$ the set of those vertices in C that have unnaturally low degree into Y. Since $d \leq d(A,C)$, we have $|N(c,Y)| < (d(A,C)-\epsilon)|Y|$ for all $c \in X$. Then for the number of edges between X and Y we have

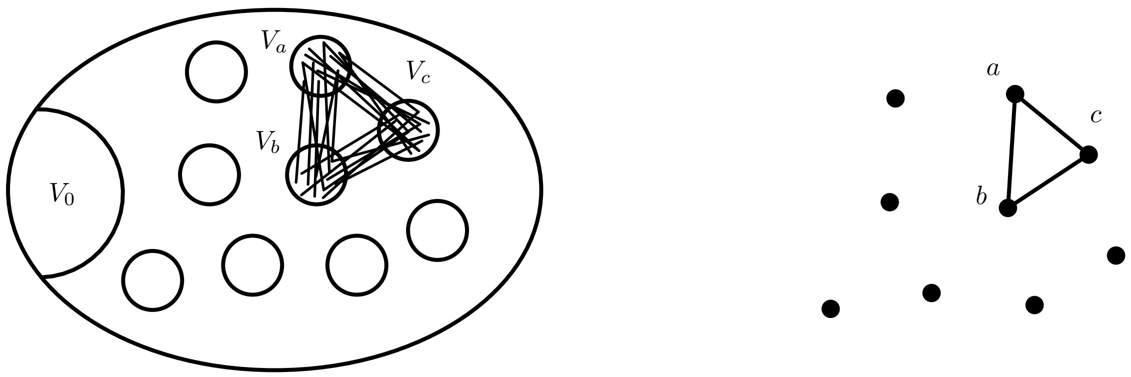
$$e(X,Y) = \sum_{c \in X} |N(c,Y)| < \sum_{c \in X} (d(A,C)-\epsilon)|Y| = (d(A,C)-\epsilon)|Y| \cdot |X|.$$

Consequently for the density we have $d(X,Y) < d(A,C) - \epsilon$. On the other hand if $|X| \geq \epsilon|C|$ then, since $|Y| \geq \epsilon|C|$, we could use the ϵ-regularity of $\{A, C\}$ to bound the density of the pair $d(X,Y)$ and in particular, $d(X,Y) \geq d(A,C) - \epsilon$. This is a contradiction, so $|X| < \epsilon|C|$. □

Using the degree lemma for the ϵ-regular pair $\{A, C\}$ with $Y = A$ and for the ϵ-regular pair $\{B, C\}$ with $Y = B$, we conclude that at least $(1-2\epsilon)|C|$ vertices in C are "normal" in the sense that they have at least $(d(A,C)-\epsilon)|A| \geq \epsilon|A|$ neighbors in A and at least $(d(B,C)-\epsilon)|B| \geq \epsilon|B|$ neighbors in B. Therefore by the ϵ-regularity of the pair $\{A, B\}$, the number of triangles is

$$\begin{aligned}
\sum_{v \in C} e(N(v,A), N(v,B)) &\geq \sum_{v \in C \text{ normal}} e(N(v,A), N(v,B)) \\
&\geq \sum_{v \in C \text{ normal}} (d(A,B)-\epsilon)|N(v,A)| \cdot |N(v,B))| \\
&\geq \sum_{v \in C \text{ normal}} (d(A,B)-\epsilon)(d(A,C)-\epsilon)|A| \cdot (d(B,C)-\epsilon)|B| \\
&\geq (1-2\epsilon)|C|(d(A,B)-\epsilon)(d(A,C)-\epsilon)|A| \cdot (d(B,C)-\epsilon)|B|.
\end{aligned}$$

□



(a) ϵ-regular partition

(b) Regularity Graph

Figure 1: Triangle removal lemma

1.1.1 Property Testing

In property testing one wants to have information about a huge input structure (say a graph) having a certain property (say triangle-freeness). Deciding whether the structure does have the property might require to query most of the input (say, looking at most pairs of vertices, and see whether they form an edge), which might be too large. Nevertheless, if we were content to only have information about the structure being *far* from having the property, then this can be decided much faster. Sometimes it is enough to query only a constantly many elements of the structure (constantly many edges of the graph), independent of its size!

A graph G is called *ϵ-far from being triangle-free*, if the removal of $\epsilon v(G)^2$ edges does not make G triangle-free.

Theorem 1.7. *For every $\epsilon > 0$ there exists an $f(\epsilon) > 0$ such that for every graph G there exists a randomized algorithm that queries at most $f(\epsilon)$ edges of G and then outputs* `accept` *or* `reject`*, such that*

- *if G is triangle-free then the algorithm* `accept`*s it with probability* 1*;*
- *if G is ϵ-far from being triangle-free, then the algorithm* `reject`*s it with probability* 0.999*.*

This statement is analogous to the (much simpler) phenomenon that exists in poll-making before elections. In order to know the approximate results of a binary election up to a certain error probability, it is enough to query say 1000 fully random people about their choices, where the size 1000 of the sample needed for the appropriate error-guarantees is irrespective of whether the election was in Steglitz or in the United States.[2]

1.2 Application 2: The Erdős-Stone Theorem

Let us revisit the statement of the more difficult direction of the Erdős-Stone Theorem, stating the upper bound on the Turán number of the Turán graph.

Theorem 1.8 (Erdős-Stone, 194?)**.** *For every $\gamma > 0$ and integers $r, s \in \mathbb{N}$ there exists an integer $N_0 > 0$ such that for every graph G on $n \geq N_0$ vertices with at least $\left(1 - \frac{1}{r-1}\right)\binom{n}{2} + \gamma n^2$ edges contains subgraph isomorphic to the Turán graph.*

Proof. Here is our plan:

[2]For the correctness of this statement we of course assumed that we are able to extract a truly independent sample of random people. This is much easier in theory than in practice. Nevertheless, the phenomenon that an effective sample size is not very sensitive to the size of the whole population does hold at large.

- Given the parameters $\gamma > 0$ and $r, s \in \mathbb{N}$, we will choose a small parameter $d > 0$, a tiny parameter $\epsilon > 0$, and a large parameter $m \in \mathbb{N}$. We apply the Regularity Lemma to obtain an integer $M = M(\epsilon, m)$, an ϵ-regular partition $\mathcal{P}$ of G, and construct the regularity graph $R(\mathcal{P}, d)$.
- When choosing the parameters d, ϵ, and m, we make sure that the number of edges of G that are *not* going between ϵ-regular pairs of density at least d is small, in fact less than $\frac{\gamma}{2}n^2$.
- We conclude that *most* edges of G go between parts V_i, V_j with $ij \in E(R)$. From this we will infer that R must also have many edges, in fact so many, that Turán's Theorem will imply the existence of a K_r in R.
- We consider r parts of the partition $\mathcal{P}$ in G, which correspond to a K_r in R. Using the pairwise ϵ-regularity of these r parts, we derive that they induce a copy of $T_{rs,r}$.

Nice plan, now let's get down to business. Given a real $\gamma > 0$, let us define parameters

$$d = \frac{\gamma}{6}, \qquad m = \left\lceil \frac{6}{\gamma} \right\rceil, \qquad \epsilon = \frac{1}{rs}\left(\frac{d}{2}\right)^{rs^2}, \qquad N_0 = \frac{s}{\epsilon(1-\epsilon)}M.$$

As we have shown in the proof of the Triangle Removal Lemma, the number of edges in G that do *not* go between parts V_i, V_j with $ij \in E(R)$ is at most $\left(\frac{3}{2}\epsilon + \frac{1}{2m} + \frac{d}{2}\right)n^2$. With our choice of parameters this is at most $\frac{\gamma}{2}n^2$ (note that our ϵ is much less than our d). Hence the number of edges of G that *do* go between parts V_i, V_j with $ij \in E(R)$ is, on the one hand, at least $\left(1 - \frac{1}{r-1}\right)\binom{n}{2} + \frac{\gamma}{2}n^2$. On the other hand, the number of these edges is trivially not more than the number of all pairs between V_i and V_j with $ij \in E(R)$, that is $|E(R)|\left(\frac{n}{k}\right)^2$. Consequently,

$$|E(R)| \geq \frac{k^2}{n^2}\left(\left(1 - \frac{1}{r-1}\right)\binom{n}{2} + \frac{\gamma}{2}n^2\right) \geq \left(1 - \frac{1}{r-1}\right)\binom{k}{2} + \frac{\gamma}{2}k^2,$$

where the last inequality is true since $n \geq N_0 \geq M \geq k$.

Hence for large k the number of edges in the regularity graph R is more than $e(T_{k,r-1}) \leq \left(1 - \frac{1}{r-1}\right)\binom{k}{2} + k$, which, by Turán's Theorem, is the maximum number of edges in K_r-free graph on k vertices. Our choice of $m = \left\lceil \frac{6}{\gamma} \right\rceil$ is large enough that for every $k \geq m$, the quadratic function $\frac{\gamma}{2}k^2$ is larger than the linear error term k in the expression for $e(T_{k,r-1})$. Therefore, by Turán's Theorem, R contains a K_r.

For ease of notation, let us assume that an r-clique in R can be found on the vertex set $[r]$. In G this means that the parts $V_1, \ldots, V_r$ are pairwise ϵ-regular with density at least d. We will find a copy of $T_{rs,r}$ in $\cup_{i=1}^r V_i$, by iteratively finding subsets $S_j \subseteq V_j$ of size s for $j = 1, \ldots \ell$, that are pairwise fully adjacent to each other, and have a large common neighborhood in each of the remaining parts V_i, $i = \ell + 1, \ldots r$. To this end the following generalization of the Degree Lemma will come in handy.

Lemma 1.9 (Common Neighborhood Lemma)**.** *Let $A, B \subseteq V(G)$ be an ϵ-regular pair of density at least d in some graph G. For every subset $Y \subseteq B$ with $(d-\epsilon)^{s-1}|Y| \geq \epsilon|A|$. Then*

$$|\{\vec{a} = (a_1, \ldots, a_s) \in A^s : |N(\vec{a}) \cap Y| < (d-\epsilon)^s|Y|\}| < s\epsilon|A|^s.$$

Remark This lemma states yet another quasirandom consequence of ϵ-regularity. The expected number of common neighbors of an s-element vertex set $\{a_1, \ldots, a_s\}$ in a truly random graph with edge-probability d is $d^s|Y|$ (since each vertex in Y is a common neighbor of all a_i with probability d^s). While we cannot expect *all* s-subsets having close to this many common neighbors, the Common Neighborhood Lemma states that *most* s-sets do.
2. The formulation of the lemma, talking about ordered s-tuples (and allowing repetitions), rather than subsets, is purely for technical convenience.

Claim *Let $\ell \in [r]$. Suppose we are given subsets $S_i \subseteq V_i$ of size s for $i = 1, \ldots, \ell - 1$, such that for their common neighborhood $A_j := N(\cup_{i=1}^{\ell-1} S_i) \cap V_j$ in the parts V_j, for $j = \ell, \ldots, r$ we have $|A_j| \geq (d - \epsilon)^{(\ell-1)s}|V_j|$. Then there exists a subset $S_\ell \subseteq V_\ell$ of size $|S_\ell| = s$ such that $|N(\cup_{i=1}^{\ell} S_i) \cap V_j| \geq (d - \epsilon)^{\ell s}|V_j|$ for every $j = \ell + 1, \ldots, r$.*

By repeated application of the Claim we obtain the required sets $S_1 \subseteq V_1, \ldots, S_r \subseteq V_r$, of size s each, which induce a copy $T_{rs,r}$.

For a proof of the Claim we use the Common Neighborhood Lemma for each $j = \ell + 1, \ldots, r$ with the ϵ-regular pair $\{V_\ell, V_j\}$ of density at least d, and with the subsets A_j as Y. For the size of Y we check that

$$(d - \epsilon)^{s-1}|Y| \geq (d - \epsilon)^{s-1}(d - \epsilon)^{(\ell-1)s}|V_j| \geq \epsilon|V_j|$$

holds, where the first inequality uses our assumption on $|A_j|$ and the second one follows from our choice of parameters for ϵ:

$$(d - \epsilon)^{s\ell-1} \geq (d - \epsilon)^{rs^2} \geq \left(\frac{d}{2}\right)^{rs^2} \geq sr\epsilon > \epsilon.$$

Therefore we can apply the Common Neighborhood Lemma.

From the lemma we obtain that for each $j = \ell + 1, \ldots, r$, the number of vectors $\vec{a} \in V_\ell^s$ for which the common neighborhood $N(\vec{a}) \cap A_j$ has size less than $(d - \epsilon)^s|A_j|$ is at most $s\epsilon|V_\ell|^s$. So for all but $(r - \ell)s\epsilon|V_\ell|^s$ s-tuples, the size of its common neighborhood in *each* of the A_j, $j = \ell + 1, \ldots, r$, is *at least* $(d - \epsilon)^s|A_j|$.

We want to find at least *one* such s-tuple in A_ℓ^s, which has pairwise different coordinates. The number of vectors with at least one pair of repeated coordinates is at most $\binom{s}{2}|V_\ell|^{s-1} \leq s\epsilon|V_\ell|^s$, since $|V_\ell| \geq \frac{1-\epsilon}{M}n \geq \frac{s-1}{2\epsilon}$ by our choice of N_0.

Since $|A_\ell^s| \geq \left((d - \epsilon)^{(\ell-1)s}|V_\ell|\right)^s \geq (d - \epsilon)^{rs^2}|V_\ell|^s > (r - \ell + 1)s\epsilon|V_\ell|^s$ by our choice of ϵ, we can conclude the existence of an s-set $S_\ell \subseteq A_\ell$, such that the neighborhood of S_ℓ inside the further $A_j = N(\cup_{i=1}^{\ell-1} S_i) \cap V_j$ is large enough:

$$|N(\cup_{i=1}^{\ell} S_i) \cap V_j| = |N(S_\ell) \cap A_j| \geq (d - \epsilon)^s|A_j| \geq (d - \epsilon)^{\ell s}|V_j|$$

for every $j = \ell + 1, \ldots, r$.

□