

# A TOWER LOWER BOUND FOR THE DEGREE RELAXATION OF THE REGULARITY LEMMA

Frederik Garbe<sup>\*1</sup> and Jan Hladký<sup>†2</sup>

<sup>1</sup>Heidelberg University, Faculty of Mathematics and Computer Science,  
Im Neuenheimer Feld 205, 69120 Heidelberg, Germany  
garbe@informatik.uni-heidelberg.de

<sup>2</sup>Institute of Computer Science of the Czech Academy of Sciences,  
Pod Vodárenskou věží 2, 182 07 Prague, Czechia  
hladky@cs.cas.cz

Submitted: Oct 15, 2024; Accepted: Jun 10, 2025; Published: Dec 20, 2025

© The authors. Released under the CC BY license (International 4.0).

**Abstract.** It is well-known that if  $(A, B)$  is an  $\frac{\varepsilon}{2}$ -regular pair (in the sense of Szemerédi) then there exist sets  $A' \subseteq A$  and  $B' \subseteq B$  with  $|A'| \leq \varepsilon|A|$  and  $|B'| \leq \varepsilon|B|$  so that the degrees of all vertices in  $A \setminus A'$  differ by at most  $\varepsilon|B|$  and the degrees of all vertices in  $B \setminus B'$  differ by at most  $\varepsilon|A|$ . We call such a property  $\varepsilon$ -*degeneracy*. This leads to the notion of an  $\varepsilon$ -*degeneracy partition* of a graph in the same way as the definition of  $\varepsilon$ -regular pairs leads to the notion of  $\varepsilon$ -regular partitions.

We show that there exist graphs in which any  $\varepsilon$ -degeneracy partition requires the number of clusters to be  $\text{tower}(\Theta(\varepsilon^{-1/3}))$ . That is, even though degeneracy is a substantial relaxation of regularity, in general one cannot improve much on the bounds that come with Szemerédi's regularity lemma.

**Keywords.** Szemerédi's regularity lemma, degree

**Mathematics Subject Classifications.** 05C35

## 1. Introduction

### 1.1. Regularity Lemma

A key concept of the famous Regularity Lemma of Szemerédi [Sze78] from 1976 is that of a regular pair. To this end, we recall that for two nonempty sets  $X$  and  $Y$  of vertices of a graph  $G$ , the *density of*  $(X, Y)$  is defined as  $d_G(X, Y) = \frac{e(X, Y)}{|X||Y|}$ . A pair  $(A, B)$  of disjoint nonempty sets of vertices is  $\varepsilon$ -*regular* if

$$|d_G(A, B) - d_G(A', B')| \leq \varepsilon \tag{1.1}$$

<sup>\*</sup>Supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 428212407.

<sup>†</sup>Supported by Czech Science Foundation Project 21-21762X. With institutional support RVO:67985807.

for every  $A' \subseteq A$  and  $B' \subseteq B$  with  $|A'| > \varepsilon|A|$  and  $|B'| > \varepsilon|B|$ .

Regular pairs are the building blocks of regular partitions. More precisely, if  $G$  is a graph, we say that a partition  $V(G) = V_1 \sqcup \dots \sqcup V_\ell$ , where the sets  $V_1, \dots, V_\ell$  are called *clusters*, is an  $\varepsilon$ -regular partition of  $G$  of complexity  $\ell$  if we have

$$(P1) \quad ||V_i| - |V_j|| \leq 1 \text{ for every } i, j \in [\ell].$$

$$(P2) \quad \text{For each } i \in [\ell] \text{ for at least } (1 - \varepsilon)\ell \text{ indices } j \in [\ell] \setminus \{i\} \text{ we have that } (V_i, V_j) \text{ is an } \varepsilon\text{-regular pair.}$$

We note that several slightly different, but essentially equivalent, definitions of regular partitions appear in the literature, each of which may have advantages in specific contexts. We discuss two such variants in Sections 1.4 and 1.5.

Observe that (P2) forces that

$$\ell \geq \varepsilon^{-1}, \tag{1.2}$$

which is often a convenient lower bound for applications of the regularity method.

The Regularity Lemma then says that for each  $\varepsilon > 0$  the number

$$R(\varepsilon) := \sup\{\min\{\ell : G \text{ has an } \varepsilon\text{-regular partition of complexity } \ell\} : G \text{ is a graph}\}$$

is finite. This statement, which has been central to extremal graph theory, information theory, and theoretical computer science, naturally, calls for studying the value  $R(\varepsilon)$ . Szemerédi's proof gives an upper bound  $R(\varepsilon) \leq \text{tow}(\Theta(\varepsilon^{-5}))$ . Here,  $\text{tow} : [0, \infty) \rightarrow \mathbb{N}$  is the *tower function*, defined as  $\text{tow}(x) = 1$  for  $x \in [0, 1)$  and  $\text{tow}(1 + x) = 2^{\text{tow}(x)}$  for  $x \geq 1$ .<sup>1</sup>

It was a major question to obtain any nontrivial lower bound on  $R(\varepsilon)$ . A breakthrough was achieved by Gowers [Gow97] who obtained a qualitatively matching lower bound  $R(\varepsilon) > \text{tow}(\Theta(\varepsilon^{-1/16}))$ . In fact, Gowers's paper has two parts. In the first one, he gives a bound  $R(\varepsilon) > \text{tow}(\Theta(\log(\frac{1}{\varepsilon})))$  using a fairly short argument. In the second part, which was called a *tour de force* by Bollobás in his official description of the work for Gowers's Fields medal [Bol98], he then proves the already mentioned bound  $R(\varepsilon) > \text{tow}(\Theta(\varepsilon^{-1/16}))$ . Subsequently, Gowers's construction was streamlined by Moshkovitz and Shapira [MS16] who based on the previously mentioned short argument proved the lower bound  $R(\varepsilon) > \text{tow}(\Theta(\varepsilon^{-1/2}))$ . The state of the art is [FL17] in which a tight bound  $R'(\varepsilon) = \text{tow}(\Theta(\varepsilon^{-2}))$ , where  $R'(\varepsilon)$  is a slight variant of  $R(\varepsilon)$  defined using the 'aggregate form' described in Section 1.4. Note that [Gow97] and [FK99] obtain lower-bounds in a finer setting in which the three different roles of  $\varepsilon$  — the tolerance of density deviation in (1.1), the lower bound on the size of the sets  $A'$  and  $B'$  in (1.1), and the proportion of regular pairs quantified in (P2) — are considered as separate parameters.

The importance of the Regularity Lemma (and its most important consequence, the Removal Lemma) initiated a whole line of research and led to many decomposition statements which are more or less related. One of the most prominent ones is the Weak Regularity Lemma of Frieze and Kannan [FK99]. The main upshot of the Weak Regularity Lemma is that the number of clusters in this type of regularization is guaranteed to be only  $2^{O(\varepsilon^{-2})}$ , and this bound is known to be almost tight, [CKKN18]. Quantitative aspects of other regularity lemmas are obtained and surveyed in [CF12]. Another recent example how to achieve a partition with a substantially smaller number of cluster, but for a weaker regularity-type property, can be found in [Csa21].

<sup>1</sup>This operation is sometimes also called 'tetration'.

## 1.2. Regularity versus degeneracy

The usage of the word *regular* is somewhat unfortunate as it traditionally refers to Szemerédi’s concept of partition’s and at the same time to another property of all the degrees being equal.<sup>2</sup> Throughout the paper, the term *regular pair* or *regular partition* refers to the former concept whereas the term *regular graph* refers to the later. In this paper we study the link between these two properties.

We call a pair  $(A, B)$  of disjoint vertex sets of a graph  $\varepsilon$ -*degeneracy* if there exist sets  $A^- \subseteq A$  and  $B^- \subseteq B$  with  $|A^-| \leq \varepsilon|A|$  and  $|B^-| \leq \varepsilon|B|$  such that

$$\left| \deg(a, B) - \frac{e(A, B)}{|A|} \right| \leq \varepsilon|B| \text{ and } \left| \deg(b, A) - \frac{e(A, B)}{|B|} \right| \leq \varepsilon|A|, \quad (1.3)$$

for every  $a \in A \setminus A^-$  and  $b \in B \setminus B^-$ .

It is well-known that an  $\varepsilon$ -regular pair is  $2\varepsilon$ -degeneracy. For completeness we recall a proof of this fact.

**Fact 1.1.** *Every  $\varepsilon$ -regular pair  $(A, B)$  is also  $2\varepsilon$ -degeneracy. In particular, for the sets  $A_1 := \{a \in A : \deg(a, B) < \frac{e(A, B)}{|A|} - \varepsilon|B|\}$  and  $A_2 := \{a \in A : \deg(a, B) > \frac{e(A, B)}{|A|} + \varepsilon|B|\}$  we have  $|A_1|, |A_2| \leq \varepsilon|A|$ . By symmetry, we have a similar property for  $B$ .*

*Proof.* We will only prove that  $|A_1| \leq \varepsilon|A|$ . The same argument can be used to bound  $|A_2|$ . If  $A_1 \neq \emptyset$  then

$$\frac{e(A, B)}{|A||B|} - \frac{e(A_1, B)}{|A_1||B|} = \frac{e(A, B)}{|A||B|} - \frac{\sum_{a \in A_1} \deg(a, B)}{|A_1||B|} > \varepsilon.$$

If we had  $|A_1| > \varepsilon|A|$  then we could use the pair  $(A_1, B)$  in (1.1) as a witness of  $\varepsilon$ -irregularity of  $(A, B)$ , which would be a contradiction.  $\square$

If we replace in (P2) the notion of  $\varepsilon$ -regular pairs by that of  $\varepsilon$ -degeneracy pairs then (P1) and (P2) lead to the notion of an  $\varepsilon$ -degeneracy partition of  $G$  of complexity  $\ell$ .

As we saw, regularity is at least as strong as degeneracy. But in fact, often it is much stronger. To see this, consider a pair  $(A, B)$  where  $A = A^1 \sqcup A^2$  and  $B = B^1 \sqcup B^2$  with  $|A^1| = |A^2|$  and  $|B^1| = |B^2|$  such that for  $i = 1, 2$ , the pair  $(A^i, B^i)$  is complete and the pair  $(A^i, B^{3-i})$  is edgeless. The pair  $(A, B)$  is 0-degeneracy while it is not even 0.499-regular. Indeed, this example can be used to show that key features from the theory of the Regularity Lemma, such as the Blow-up lemma or the Counting lemma, fail badly when relaxed to degeneracy pairs. This example also shows that the key property that a non-negligible subpair of a regular pair is again regular with a similar density, does not hold for degeneracy. Note that we show in Fact 2.6 that a one-sided counterpart of the statement holds, though.

There are a number of complex applications of regularity in which some steps actually utilize only the degeneracy property rather than the full regularity property. Take for example, ‘Łuczak’s

---

<sup>2</sup>Some authors in the 1990’s used the term *Uniformity lemma* instead of the Regularity Lemma, see e.g. [FR92, Gow97].

connected matching argument’, introduced in [Lu99] and used many times since to establish the existence of a long path in a graph  $G$  from the fact that its cluster graph  $\mathbf{G}$  contains a large matching  $\mathbf{M}$  in one connected component.<sup>3</sup> The argument has three parts. In the first part, connectivity of  $\mathbf{G}$  is used to find a system  $\mathcal{S}$  of vertex-disjoint paths in  $G$  between each two edges of  $\mathbf{M}$ . In the second part, each regular pair corresponding to an edge  $e \in \mathbf{M}$  is filled-up with a path  $P_e$  in  $G$ . In the third part, the paths of  $(P_e)_{e \in \mathbf{M}}$  are merged using  $\mathcal{S}$  into a single desired path. While for the second part, one needs the full strength of regular pairs, the system of paths  $\mathcal{S}$  in the first part can be planted using a single one-step extension strategy for which the degularity property is sufficient. There are examples where degularity is the only property used from an application of the Regularity Lemma. This point is made explicitly in [HH23], which we quickly and briefly recall. The concept studied in [HH23] is fractional isomorphism. This is a well-studied concept in graphs, and recently a graphon counterpart was introduced in [GR22]. The main result is that if  $U$  and  $W$  are two fractionally isomorphic graphons, then  $U$  can be approximated in the cut distance by a graph  $G$  and  $W$  can be approximated in the cut distance by a graph  $H$  in a way that  $G$  and  $H$  are fractionally isomorphic. The proof starts with a regularization of a ‘common equitable partition’ of  $U$  and  $W$ .<sup>4</sup> The regularization is one of the two main steps in building  $G$  and  $H$ . Since fractional isomorphism is a concept which is heavily based on degrees and much less on the actual structure, the regularization could be replaced by degularization. We believe that with minor changes, degularity could similarly replace regularity also in [BBCR10] and [HNT18].

### 1.3. Main result: Lower bound on the number of clusters in degular partitions

Fact 1.1 tells us that every  $\varepsilon$ -regular partition is also  $2\varepsilon$ -degarular, that is, for the number

$$D(\varepsilon) := \sup\{\min\{\ell : G \text{ has an } \varepsilon\text{-degarular partition of complexity } \ell\} : G \text{ is a graph}\}$$

we have  $D(\varepsilon) \leq R(\varepsilon/2)$ , yielding in particular  $D(\varepsilon) \leq \text{tow}(\Theta(\varepsilon^{-5}))$ . Given how much weaker the degularity property is compared to regularity — as we illustrated above and also in Section 1.6 — one could hope for a much better bound. Our main result shows that such a hope is in vain.

**Theorem 1.2.** *We have  $D(\varepsilon) \geq \text{tow}(\Theta(\varepsilon^{-1/3}))$ .*

#### About the proof

The proof of Theorem 1.2 builds heavily on the paper of Moshkovitz and Shapira [MS16] which gives the lower bound  $R(\varepsilon) \geq \text{tow}(\Theta(\varepsilon^{-1/2}))$ . The main idea of [MS16] (which is in turn taken from the seminal paper [Gow97]) for the construction of a graph with no  $\varepsilon$ -regular partitions of complexity less than  $\text{tow}(\Theta(\varepsilon^{-1/2}))$  is the following. We partition the vertex set into a constant number of blobs. Then in  $L = \Theta(\varepsilon^{-1/2})$  many rounds, we partition each current blob into a number of subblobs, the number of subblobs being exponential in the number of current blobs.

<sup>3</sup>‘Cluster graph’ on the vertex set  $\{V_1, \dots, V_\ell\}$  whose edges correspond to those pairs that are regular, and of sufficient edge density.

<sup>4</sup>The theory of regularizations of graphons is analogous to the theory of regularizations of graphs.

Overall, this results in  $\text{tow}(\Theta(\varepsilon^{-1/2}))$  blobs on the  $L$ -th level. The edges are then ingeniously defined based on this nested blob structure. The key inductive step in the proof of the lower bound is the following property (loosely):

- (♥) If an  $\varepsilon$ -regular partition (almost) refines blobs of some level  $\ell < L$  then it also (almost) refines blobs on level  $\ell + 1$ .

This key feature (♥) relies, besides the combinatorial definition of the edges, on a basic property of regular pairs, namely that a subpair of a regular pair is regular. As we already illustrated before, this property does not carry over to degular pairs and the same argument therefore does not work to show Theorem 1.2. Moreover in Section 1.7 we discuss that the construction from [MS16] has a  $\varepsilon$ -degener partition of relatively small complexity. Thus the main work in our proof of Theorem 1.2 is to circumvent this shortcoming using a one-sided version of the above subpair property which is still valid in the degular setting, namely Fact 2.6. This in turn leads to introducing additional orientations on the blobs of each level and further refined adjustments of the construction.

## Structure of the paper

The proof of Theorem 1.2 is given in Section 2. In the rest of this section, we discuss further aspects of our main result. In Sections 1.4 and 1.5 we discuss two alternative and looser definitions for degular partitions and argue that a lower bound similar to that in Theorem 1.2 holds for them as well. In Section 1.6 we show that regular graphs have degular partitions of small complexity. We use this to show that each graph is an induced graph of a graph twice as big, and the latter has a degular partition of small complexity. These features are in stark contrast to the setting of the Regularity Lemma. In Section 1.7 we discuss the gap between complexities of degular and regular partitions.

### 1.4. Two sparsity conditions for irregular pairs

The condition (P2) is sometimes called ‘the degree form of the Regularity Lemma’. There is also an ‘aggregate form’, which instead requires that at most  $\varepsilon\ell^2$  pairs  $(V_i, V_j)$  are  $\varepsilon$ -irregular. The degree form obviously implies the aggregate one. However, there is a well-known transformation of a  $\frac{\varepsilon^2}{8}$ -regular partition  $U_1 \sqcup \dots \sqcup U_h$  in the aggregate form into the degree form which goes as follows. Remove from the collection each cluster  $U_i$  which forms an  $\frac{\varepsilon^2}{8}$ -irregular pair with more than  $\frac{\varepsilon}{4}h$  other clusters. Simple counting gives that there are at most  $\frac{\varepsilon}{4}h$  such clusters. We distribute the contents of these problematic clusters evenly over the non-problematic ones. It is easy to check that the resulting partition is  $\varepsilon$ -regular in the degree form. The same works for degular partitions.

### 1.5. Equipartitions versus partitions

There is an frequent alternative definition of regularity which does not require equipartitions. That is, the sets  $V_1, \dots, V_\ell$  in (P1) may be of arbitrary positive cardinalities. In such a setting (P2)

would be useless. Indeed we could take  $V_1, \dots, V_{\ell-1}$  singletons and  $V_\ell$  the rest. Then all pairs would be regular. Thus, (P2) is replaced with the condition

$$\sum_{i,j:(*)} |V_i||V_j| \geq (1 - \varepsilon)n^2.$$

where  $(*)$  runs over all  $(i, j) \in [\ell]^2$ ,  $i \neq j$  such that  $(V_i, V_j)$  is  $\varepsilon$ -regular.

One can get from an  $\frac{\varepsilon}{2}$ -regular partition  $U_1, \dots, U_\ell$  of clusters of arbitrary cardinalities to the equi-sized setting by breaking up each set  $U_i$  into random subsets of cardinality  $\frac{\varepsilon n}{2\ell}$  and distributing the leftovers randomly. Note that this way we would only get the aggregate form of the regularity and have to apply the transformation described in Section 1.4 to transform it to the degree form. The same can be done in the degular setting; we give the details in a related situation in Section 1.6.

## 1.6. Degular partitions of regular graphs

In this section, we show two things. Firstly, we show that each approximately regular graph has a degular partition of arbitrarily small complexity. Secondly, we show that each graph is an induced subgraph of a regular graph of twice the order of the former one, which has by the above a degular partition of arbitrarily small complexity. Both these features fail notably in the setting of regular partitions.

Suppose that  $G$  is a regular graph on  $n$  vertices. Suppose that  $\varepsilon > 0$  and  $L \in \mathbb{N}$  are given, with the condition that  $L = O_\varepsilon(\log n)$ . Basic concentration inequalities give that with positive probability a random partition of  $V(G)$  into  $L$  sets  $V_1, \dots, V_L$  of individual sizes  $\lfloor \frac{n}{L} \rfloor$  and  $\lceil \frac{n}{L} \rceil$  has the property that all pairs  $(V_i, V_j)$ ,  $\{i, j\} \in \binom{[L]}{2}$ , are  $\varepsilon$ -degular. That is, no non-trivial lower bound holds for the complexity of degular partitions. This can be extended to approximately regular graphs, that is, any graph in which the minimum and the maximum degrees differ by at most  $\frac{\varepsilon n}{2}$ . This is in contrast with the situation of regular partitions, see Section 1.7 for details.

This argument can be substantially extended as follows. Suppose that  $G$  is an arbitrary graph of order  $n$  on vertex set  $V$ . We construct a graph  $H$  of order  $2n$  on vertex set  $V \sqcup W$ . Firstly, we take  $H[V] := G$ . Secondly, step-by-step we take each  $v \in V$  and by introducing edges between  $v$  and  $W$  we make sure that  $\deg_H(v) = n - 1$ . The edges are introduced in such a way that in each step we have  $\max_{w \in W} \deg(w, V) - \min_{w \in W} \deg(w, V) \leq 1$ . In the last step, we construct a graph  $H[W]$  in such a way that for each  $w \in W$  we have  $\deg_H(w) = n - 1$ , or equivalently,  $\deg_{H[W]}(w) = n - 1 - \deg_H(w, V)$ . By the above, this is a requirement for a construction of a regular or almost regular graph, and such constructions are well-known to exist.<sup>5</sup> To summarize, we have constructed an almost regular graph  $H$  on  $2n$  vertices of which  $G$  is an induced subgraph. By the discussion above,  $H$  has a degular partition  $V(H) = V_1 \sqcup \dots \sqcup V_\ell$  of arbitrarily small complexity  $\ell$  (subject to a trivial condition (1.2)). Again, this is in a stark contrast with the situation of regular partitions. Indeed, if the above partition were  $\varepsilon$ -regular, then the restricted partition  $V = (V_1 \cap V) \sqcup \dots \sqcup (V_\ell \cap V)$  would be a  $2\varepsilon$ -regular partition of  $G$  (albeit with (P1) dropped; this is known not to be a real shortcoming).

<sup>5</sup>As follows from the Erdős–Gallai theorem on graphic sequences.

### 1.7. Graphs with complex regular partitions but simple degular partitions

As we show in this paper, even though degeneracy seems as substantial relaxation of regularity, in general, one cannot use the relaxation to improve much on the construction from Szemerédi’s regularity lemma. However, there are graphs in which the gap between the most economic  $\varepsilon$ -regular partition and the  $\varepsilon$ -degeneracy partition is huge. To see that, recall that Moshkovitz and Shapira [MS16] construct a weighted template graphs of an arbitrarily large order  $n$  which is approximately degenerate<sup>6</sup> but each  $\varepsilon$ -regular partition has complexity at least  $\text{tow}(\Theta(\varepsilon^{-1/2}))$ . On the other hand, by the discussion in Section 1.6, such graphs have degeneracy partitions of complexity  $O(\varepsilon^{-1})$ .

## 2. Proof of Theorem 1.2

In this section, we prove Theorem 1.2. More specifically, in Section 2.2, for given  $\varepsilon$  and  $n$  we construct an  $n$ -vertex *weighted graph*, i.e. a graph where each edge has an assigned  $[0, \infty)$ -weight,  $G(n, s, \delta)$ ; here we will choose  $\delta \sim \varepsilon^{1/3}$  and  $s \sim \delta^{-1}$ . In Section 2.4 we transform this weighted graph into an unweighted graph and show that each  $\varepsilon$ -degeneracy partition of this graph has complexity at least  $\text{tow}(\Theta(\varepsilon^{-1/3}))$ .

There are two auxiliary parts needed. In Section 2.1 we construct ‘separators’. These are certain systems of subsets of a finite auxiliary set, strengthening the notion of ‘balanced partitions’ used in previous work on lower bounds for the Regularity Lemma. Separators are a key building block whose iterative use yields the graph  $G(n, s, \delta)$  of our main construction in Section 2.2.

The second auxiliary part is Section 2.3. There, in Fact 2.6 we show that the well-known fact that a subpair of a regular pair is again regular has a one-sided counterpart in the degeneracy setting. Fact 2.6 is used in the main proof in Section 2.4.

### 2.1. Separators

The iterated key step in our lower-bound construction (as well as in the lower-bound constructions of [Gow97] and [MS16]) comprises of the following splitting procedure. Given a subset of the vertices of the host graph, say consisting of  $M$  subsets of equal-sized clusters  $\mathcal{C}$  of vertices, and an integer  $D$ , the clusters  $\mathcal{C}$  are partitioned in  $D$  different ways into two equal-sized collections: the ‘first’ collection and the ‘second’ collection. The key nontrivial property is that of ‘balancedness’: for each two clusters  $C_t, C_{t'} \in \mathcal{C}$ , there is a substantial number of partitions (out of the  $D$  many above) where  $C_t$  belongs into the first collection and  $C_{t'}$  belongs to the second collection, or vice versa. The precise definition from [MS16], in which we identify  $\mathcal{C} \doteq [M]$  goes as follows: A sequence of bipartitions  $(A_i, B_i)_{i=1}^D$  of  $[M]$  is *c-balanced* if

(BP1) for each  $i \in [D]$ ,  $A_i$  is a subset of  $[M]$  of cardinality  $\frac{M}{2}$ , and

(BP2) for each  $t, t' \in [M]$  distinct,  $|\{i \in [D] : |\{t, t'\} \cap A_i| = 1\}| \geq (\frac{1}{2} - c)D$ .

---

<sup>6</sup>More precisely, they construct a perfectly degenerate weighted graph (see Equation (1) in [MS16]) which is then, by the randomization procedure of Claim 3, transformed into an approximately regular graph. Here, by an ‘approximately regular graph’, we mean that all the degrees differ by at most  $\varepsilon n/2$ .

In our construction, we need to add an additional property. This leads us to the notion of separators.

**Definition 2.1** (separators). Given two even positive integers  $D$  and  $M$ , we call a system  $(A_i, B_i)_{i=1}^D$  an  $(M, D)$ -separator if it is an 0.2-balanced system of bipartitions of  $[M]$ , and

$$(S1) \text{ for each } t \in [M], |\{i \in [D] : t \in A_i\}| = \frac{D}{2}.$$

The existence of a  $c$ -balanced system of bipartitions (for suitable parameters  $D$  and  $M$ ) was shown in [MS16] by a simple application of the probabilistic method. Using a somewhat more complex argument, we can prove the existence of separators.

**Lemma 2.2.** *For every even positive integer  $D$  and even positive integer  $M \leq 2^{\lceil D/9999 \rceil}$ , there exists an  $(M, D)$ -separator.*

*Proof.* We will only define the sets  $A_i$ , as the sets  $B_i$  are determined by  $B_i := [M] \setminus A_i$ . For  $D < 9999$  we have  $M = 2$ . We define  $A_1 = A_2 = \dots = A_{\frac{D}{2}} = \{1\}$  and  $A_{\frac{D}{2}+1} = A_{\frac{D}{2}+2} = \dots = A_D = \{2\}$ . Obviously, this satisfies all the required properties.

Suppose that  $D \geq 9999$ . The construction has two steps. We first construct sets  $(A_i)_{i=1}^{D^*}$ , where  $D^* := 2\lceil 0.45D \rceil$ , in a way that the following versions of the properties from the definition of balanced systems and separators hold.

$$(BP1^*) \text{ For each } i \in [D^*], A_i \text{ is a subset of } [M] \text{ of cardinality } \frac{M}{2}.$$

$$(BP2^*) \text{ For each } t, t' \in [M] \text{ distinct, } |\{i \in [D^*] : \{t, t'\} \cap A_i = 1\}| \geq 0.3D.$$

$$(S1^*) \text{ For each } t \in [M], |\{i \in [D^*] : t \in A_i\}| \in [0.41D, 0.49D].$$

The proof of this first step is by the probabilistic method. That is, we claim that if taking  $A_i$  for  $i = 1, \dots, D^*$  as a random subset of  $[M]$  of size  $\frac{M}{2}$ , the system  $(A_i)_{i=1}^{D^*}$  has the desired properties with positive probability. First, we calculate the probability of violating (S1\*) for a single  $t \in [M]$ . The random variable counting the number of indices  $i \in [D^*]$  containing  $t$  has mean  $\frac{D^*}{2} = 0.45D \pm 0.001D$ . Chernoff's bound gives that the probability of this random variable deviating more than  $0.039D$  from its mean is at most  $2 \exp(-2(0.039)^2 D)$ . Next, we calculate the probability of violating (BP2\*) for a single pair of distinct  $t, t' \in [M]$ . The random variable counting the number of indices  $i \in [D^*]$  containing exactly one of  $\{t, t'\}$  has mean  $\frac{D^*}{2} \geq 0.44D$ . Chernoff's bound gives that the probability that this random variable is at least  $0.14D$  below its expectation is at most  $\exp(-2(0.14)^2 D)$ . We use the union bound over the  $M$  many choices of  $t$  for which (S1\*) could fail, and also over the  $\binom{M}{2}$  many choices of  $t$  and  $t'$  for which (BP2\*) could fail. Since calculations give for  $D \geq 9999$  that

$$M \cdot 2 \exp(-2(0.039)^2 D) + \binom{M}{2} \cdot \exp(-2(0.14)^2 D) < 1,$$

we indeed conclude that the system  $(A_i)_{i=1}^{D^*}$  has the desired properties (BP1\*), (BP2\*) and (S1\*) with positive probability.

We now proceed with the second step of the construction, which is the addition of the sets  $(A_i)_{i=D^*+1}^D$ . Let  $R := D - D^*$ . Now, sequentially in steps  $j = 1, \dots, R$ , we add a set  $A_{D^*+j}$ . Before each step  $j$ , we require the following invariant:

For  $t \in [M]$  define  $\chi_j(t) := \frac{D}{2} - |\{i \in [D^* + j - 1] : t \in A_i\}|$ . We require  $\chi_j(t) \leq R - j + 1$ .

That is,  $\chi_j(t)$  counts how many times we still need to include the element  $t$  in  $A_{D^*+j}, A_{D^*+j+1}, \dots, A_D$  in order to satisfy (S1). Note that we always have

$$\sum_{t=1}^M \chi_j(t) = \frac{MD}{2} - (D^* + j - 1) \frac{M}{2} = (R - j + 1) \frac{M}{2}. \tag{2.1}$$

The condition  $\chi_j(t) \leq R - j + 1$  makes sure that we do not have to include  $t$  into more sets than the number of leftover steps. Also, note that we may not include an element  $t$  into  $A_{D^*+j}$  if we have  $\chi_j(t) = 0$ .

First note that this invariant is satisfied for  $j = 1$ . Indeed, by (S1\*),

$$\chi_1(t) \leq \frac{D}{2} - 0.41D \leq 0.1D - 0.001D \leq D - D^* = R - 1 + 1.$$

Now, in step  $j \geq 1$ , we define  $A_{D^*+j}$  as consisting of the  $\frac{M}{2}$  elements  $t \in [M]$  with the highest value of  $\chi_j(t)$  (breaking ties arbitrarily). We first verify that none of these selected elements has  $\chi_j(t) = 0$ . As none of the summands in (2.1) is bigger than  $R - j + 1$  by the invariant, at least  $\frac{M}{2}$  many are nonzero. Next, we verify the invariant for  $\chi_{j+1}$ . Suppose for a contradiction that there exists  $t$  such that  $\chi_{j+1}(t) \geq R - j + 1$ . We see that  $t$  was not selected for  $A_{D^*+j}$  (indeed, otherwise, using the invariant for  $\chi_j$ ,  $\chi_{j+1}(t) = \chi_j(t) - 1 \leq (R - j + 1) - 1$ ). From the general monotonicity  $\chi_j \geq \chi_{j+1}$ , we have  $\chi_j(t) \geq R - j + 1$ . But since  $t$  was not selected for  $A_{D^*+j}$ , we see that there were at least  $\frac{M}{2}$  additional elements with  $\chi_j$  at least as high. In particular,

$$\sum_{t=1}^M \chi_j(t) \geq \left(\frac{M}{2} + 1\right) (R - j + 1),$$

a contradiction to (2.1). □

The next lemma is Lemma 2.3 from [MS16], with changed constants. We omit the proof as it follows *mutatis mutandis*.

**Lemma 2.3.** *Let  $D, M \in \mathbb{N}$ , and  $\zeta \in (0, \frac{1}{8})$  be arbitrary. If  $(A_i, B_i)_{i=1}^D$  is a sequence of bipartitions of  $[M]$  that is  $1/5$ -balanced, then for every  $\lambda = (\lambda_1, \dots, \lambda_M)$  with  $\lambda_t \geq 0$ ,  $\|\lambda\|_1 = 1$  and  $\|\lambda\|_\infty \leq 1 - 8\zeta$ , at least  $D/15$  of the bipartitions  $(A_i, B_i)$  satisfy  $\min(\sum_{t \in A_i} \lambda_t, \sum_{t \in B_i} \lambda_t) \geq \zeta$ . In particular this holds for an  $(M, D)$ -separator.*

## 2.2. Construction

Given integers  $n, s$  and  $\delta > 0$  we construct a weighted graph  $G = G(n, s, \delta)$  on  $[n]$ . In fact, we construct  $s$  weighted graphs  $G_1, \dots, G_s$  and then obtain  $G$  by putting these graphs together.

Later we will, for a given  $\varepsilon > 0$ , choose  $n$  large in terms of  $\varepsilon$ ,  $\delta$  and  $s$ . Since we work with weighted graphs we extend the definition of the degree of a vertex  $\deg(v)$  and the edge density between two sets of vertices  $d(A, B)$  straightforwardly by summing over the corresponding edge weights.

Now, we define recursively  $m_0 = 2$  and for  $r \in [s]$  we set  $m_r = m_{r-1}M_r$ , where  $M_r = 2^{\lceil m_{r-1}/(4 \cdot 9999) \rceil}$ . We always assume that  $n$  has the right divisibility for what follows and consider a sequence of equipartitions  $(\mathcal{X}_r)_{r \in [s]}$  of  $[n]$  in which  $\mathcal{X}_0 = \{\{1, \dots, n/2\}, \{n/2 + 1, \dots, n\}\}$  and iteratively  $\mathcal{X}_r$  is a refinement of  $\mathcal{X}_{r-1}$  by partitioning each partite set of  $\mathcal{X}_{r-1}$  into  $M_r$  many consecutive intervals. Note that  $m_r = |\mathcal{X}_r|$ .

For every  $r \in [s-1]$  we construct a weighted graph  $G_{r+1}$ . A crucial homogeneity property in our construction will be that for every two distinct cells  $C_1, C_2 \in \mathcal{X}_{r+1}$ ,

$$\text{the edge weights of the bipartite graph } G_{r+1}[C_1, C_2] \text{ are either all equal to } \delta \text{ or } 0. \quad (2.2)$$

Write

$$\mathcal{X}_{r-1} = \{X_1, \dots, X_{m_{r-1}}\} \quad \text{and} \quad \mathcal{X}_r = \{X_{i,t}\}_{i \in [m_{r-1}], t \in [M_r]}, \quad (2.3)$$

where  $X_i$  is partitioned into  $\{X_{i,1}, \dots, X_{i,M_r}\}$ . For every  $1 \leq i < j \leq m_{r-1}$  choose a regular bipartite tournament on  $\{X_{i,t}\}_{t=1}^{M_r} \cup \{X_{j,t}\}_{t=1}^{M_r}$ . Such a bipartite tournament exists; for example we can take the directed edges

$$\{(X_{i,t}, X_{j,t'})\}_{t \in [M_r], t' \in J_t^+} \cup \{(X_{j,t}, X_{i,t'})\}_{t \in [M_r], t' \in J_t^{++}},$$

where  $J_t^+ \subseteq [M_r]$  is defined as the  $M_r/2$ -many elements (recall that  $M_r$  is even) that include and follow  $t$ , viewed cyclically in the set  $[M_r]$ ,

$$J_t^+ = \{t, t+1, \dots, t+M_r/2-1\} \pmod{M_r},$$

and similarly,  $J_t^{++}$  are the  $M_r/2$  many numbers that follow  $t$ ,

$$J_t^{++} = \{t+1, t+2, \dots, t+M_r/2\} \pmod{M_r}.$$

The above bipartite tournaments have disjoint edge sets. Their union is denoted by  $T_r$  and is itself a digraph on  $\mathcal{X}_r$ . Note that the indegree and the outdegree of every vertex in  $T_r$ , i.e. a set  $X_{i,t}$ , is equal to  $D_r = (m_{r-1} - 1) \cdot M_r/2 = (m_r - M_r)/2$ . Observe that

$$D_r \geq m_r/4 \quad (2.4)$$

and hence  $M_{r+1} \leq 2^{\lceil D_r/9999 \rceil}$ . Figure 2.1 is aimed to help with the steps that follow. Note that since (2.3) we mostly considered the layers  $\mathcal{X}_{r-1}$  and  $\mathcal{X}_r$ . Now the considered depth will move up by one and we will work with  $(\mathcal{X}_r, \mathcal{X}_{r+1})$ . Correspondingly, elements of  $\mathcal{X}_r$  will be denoted by a single index (e.g.,  $X_j$ ) and elements of  $\mathcal{X}_{r+1}$  will be denoted by pair of indices (e.g.,  $X_{j,t}$ ).

Lemma 2.2 tells us there exists an  $(M_{r+1}, D_r)$ -separator  $(A'_j, B'_j)_{j=1}^{D_r}$ . We define a weighted graph  $G_{r+1}$  on the vertex set  $[n]$  the following way. Fix  $j \in [m_r]$ . Let  $\{X_{i_1}, \dots, X_{i_{D_r}}\}$  be the

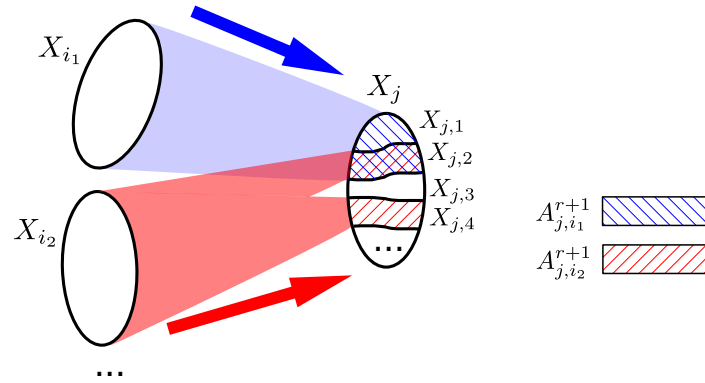


Figure 2.1: A part of the construction of the graph  $G_{r+1}$ . In this example,  $X_{i_1}$  and  $X_{i_2}$  are two inneighbours of  $X_j$  in  $T_r$ , and  $A'_{i_1} = \{1, 2, \beta, \mathbb{A}, \dots\}$  and  $A'_{i_2} = \{\lambda, 2, \beta, 4, \dots\}$ . The solid blue and red areas indicate positive weights inserted in (2.5).

inneighbourhood of  $X_j$  in  $T_r$ . Recall that  $X_j = \bigcup_{t \in [M_{r+1}]} X_{j,t}$ , where  $X_{j,t} \in \mathcal{X}_{r+1}$ . Then for every  $\ell \in \{1, \dots, D_r\}$  set  $A_{j,i_\ell}^{r+1} = \bigcup_{t \in A'_\ell} X_{j,t}$ . Finally, set

$$G_{r+1}(x, y) = \begin{cases} \delta, & \text{if } x \in X_i, y \in X_j \text{ for some } (X_i, X_j) \in T_r \text{ and } y \in A_{j,i}^{r+1}, \\ 0, & \text{otherwise.} \end{cases} \tag{2.5}$$

It is clear that (2.2) holds.

Additionally, for  $\mathcal{X}_0 = \{X_1, X_2\}$ , we define  $G_1$  to be the graph with

$$G_1(x, y) = \begin{cases} 0.1, & \text{if } x, y \in X_1, \\ 0.9, & \text{if } x, y \in X_2, \\ 0, & \text{otherwise.} \end{cases} \tag{2.6}$$

**Construction 2.4.** Given  $s \in \mathbb{N}$ ,  $\delta > 0$  and  $n \geq m_s$  with  $m_s \mid n$ , we set  $G = G(n, s, \delta) = \sum_{r=1}^s G_r$ .

Given a vertex  $x$  and a set of vertices  $X$  we write  $d(x, X)$  shorthand for  $d(\{x\}, X)$ .

**Observation 2.5.** For  $\delta > 0$ ,  $s \in \mathbb{N}$  and  $n \geq m_s$  with  $m_s \mid n$ , let  $G$  be given by Construction 2.4. Let  $r \in [s]$  and  $X \in \mathcal{X}_{r-1}$ . If  $x \in [n] \setminus X$ , then

$$d_{G_{r+1}}(x, X) + \dots + d_{G_s}(x, X) = \frac{1}{2}\delta(s - r).$$

*Proof.* Fix  $r' \in \{r + 1, \dots, s\}$  and fix the unique  $X_i \in \mathcal{X}_{r'-1}$  with  $x \in X_i$ . For  $X_j \in \mathcal{X}_{r'-1}$  we have

$$d_{G_{r'}}(x, X_j) = \begin{cases} 0, & \text{if } X_i = X_j, \\ \delta/2, & \text{if } (X_i, X_j) \in T_{r'-1}, \\ \delta, & \text{if } (X_j, X_i) \in T_{r'-1} \text{ and } x \in A'_{i,j}, \\ 0, & \text{if } (X_j, X_i) \in T_{r'-1} \text{ and } x \notin A'_{i,j}. \end{cases}$$

Since  $\mathcal{X}_{r'-1}$  is a refinement of  $\mathcal{X}_{r-1}$  there exists  $\mathcal{P} \subseteq \mathcal{X}_{r'-1}$  such that  $X = \bigcup_{X' \in \mathcal{P}} X'$ . Furthermore by the definition of  $T_{r'-1}$  we have  $|\mathcal{P} \cap \mathbf{N}_{T_{r'-1}}^{\text{in}}(X_i)| = |\mathcal{P} \cap \mathbf{N}_{T_{r'-1}}^{\text{out}}(X_i)| = |\mathcal{P}|/2$ . Furthermore, by the choice of separators in the construction and Definition 2.1 (S1) for exactly  $\frac{|\mathcal{P}|/2}{D_{r'-1}} \cdot \frac{D_{r'-1}}{2} = |\mathcal{P}|/4$  many  $X_j \in \mathcal{P} \cap \mathbf{N}_{T_{r'-1}}^{\text{in}}(X_i)$  we have  $x \in A_{i,j}^{r'}$  and for the same number of  $X_j \in \mathcal{P} \cap \mathbf{N}_{T_{r'-1}}^{\text{in}}(X_i)$  we have  $x \notin A_{i,j}^{r'}$ . Therefore

$$d_{G,r}(x, X) = \frac{1}{2} \cdot \delta/2 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot \delta = \delta/2$$

and the claim follows by summing over  $r' = r + 1, \dots, s$ .  $\square$

### 2.3. Density inheritance in degular pairs

A crucial property of an  $\varepsilon$ -degular pair  $(A, B)$  is that the edge-density between every subset  $X \subseteq A$  and the whole set of  $B$  is almost the same as the one between  $A$  and  $B$ . We summarize this straightforward observation in the following statement.

**Fact 2.6.** *Let  $(A, B)$  be an  $\varepsilon$ -degular pair in a possibly weighted graph. For every  $X \subseteq A$  we have*

$$|d(X, B) - d(A, B)| \leq \left(1 + \frac{|A|}{|X|}\right) \varepsilon.$$

*Proof.* Let  $A^- \subseteq A$  be a set of size at most  $\varepsilon|A|$  as in (1.3). We have

$$\begin{aligned} d(X, B) &= \frac{1}{|X|} \left( \sum_{x \in X \setminus A^-} d(x, B) + \sum_{x \in X \cap A^-} d(x, B) \right) \\ &= \frac{1}{|X|} \left( \sum_{x \in X \setminus A^-} (d(A, B) \pm \varepsilon) + \sum_{x \in X \cap A^-} (d(A, B) \pm 1) \right) \\ &= d(A, B) \pm \left( \varepsilon + \varepsilon \cdot \frac{|A|}{|X|} \right). \end{aligned} \quad \square$$

### 2.4. Main proof

Below, we work with the natural extension of degeneracy and degular partitions to the setting of weighted graphs. We show the following lemma which is an adaptation of Lemma 2.4 from [MS16] to our altered construction. For two vertex sets  $A, B$  we write  $A \subseteq_\beta B$  if  $|A \cap B| \geq (1 - \beta)|A|$ . Furthermore, we say that a partition  $\mathcal{Z}$  of  $[n]$  is a  $\beta$ -refinement of a partition  $\mathcal{P}$  of  $[n]$  if for every  $Z \in \mathcal{Z}$  there exists a  $P \in \mathcal{P}$  such that  $Z \subseteq_\beta P$ .

**Lemma 2.7.** *Suppose that  $\varepsilon, \beta, \mu, \delta > 0$  are such that  $32\varepsilon/\mu, 4000\varepsilon, 1600\beta < \delta < 1/2$ , and  $1 \leq r \leq s$  and  $s\delta \leq 0.1$ . For  $n$  large enough let  $G(n, s, \delta)$  be given by Construction 2.4. If  $\mathcal{Z}$  is an  $\varepsilon$ -degular partition of  $G$  that  $\beta$ -refines  $\mathcal{X}_{r-1}$ , then it  $(\beta + 8\mu)$ -refines  $\mathcal{X}_r$ .*

*Proof.* Suppose that there is an  $\varepsilon$ -degener partition  $\mathcal{Z}$  of complexity  $\ell$  of the weighted graph  $G$  and for the sake of contradiction that there exists  $Z_0 \in \mathcal{Z}$  with  $Z_0 \subseteq_{\beta} X_{j^*}$  for some  $X_{j^*} \in \mathcal{X}_{r-1}$ , but  $Z_0 \not\subseteq_{\beta+8\mu} X_{j^*,t}$  for every  $t \in [M_r]$ . From now on we fix this  $Z_0$  and  $X_{j^*}$ .

At the end of the proof, we will show that there are more than  $\varepsilon\ell$  many sets  $Z \in \mathcal{Z}$  such that  $(Z, Z_0)$  is not  $\varepsilon$ -degener showing that  $\mathcal{Z}$  violates the degenerity counterpart of (P2) and is therefore not an  $\varepsilon$ -degener partition. We need quite some preparations to obtain this contradiction.

We say that a vertex  $v \in [n]$  is *useful* if for the unique  $X_i \in \mathcal{X}_{r-1}$  with  $X_i \ni v$  and for the unique  $Z \in \mathcal{Z}$  with  $Z \ni v$  we have  $Z \subseteq_{\beta} X_i$ . The following claim is crucial for the proof.

**Claim 1.** *There are at least  $m_{r-1}/120$  many sets  $X_i \in \mathcal{X}_{r-1}$  which are contributing, i.e.*

(i) *all but at most  $(1 - 120\beta)|X_i|$  many vertices of  $X_i$  are useful,*

(ii)  $(X_i, X_{j^*}) \in T_{r-1}$ ,

(iii)  $|Z_0 \cap A_{j^*,i}^r|, |Z_0 \cap (X_{j^*} \setminus A_{j^*,i}^r)| > \mu|Z_0|$ .

*Proof of Claim 1.* For each  $t \in [M_r]$  we set  $\lambda_t = |Z_0 \cap X_{j^*,t}|/|Z_0 \cap X_{j^*}| \geq 0$ . Let  $\lambda = (\lambda_1, \dots, \lambda_{M_r})$ . Note that  $\|\lambda\|_1 = 1$  and

$$\|\lambda\|_{\infty} < \frac{(1 - (\beta + 8\mu))|Z_0|}{(1 - \beta)|Z_0|} = 1 - 8 \cdot \frac{\mu}{1 - \beta},$$

as  $Z_0 \subseteq_{\beta} X_{j^*}$  and  $Z_0 \not\subseteq_{\beta+8\mu} X_{j^*,t}$  for every  $t \in [M_r]$ . Recall that  $|\mathbf{N}_{T_{r-1}}^{\text{in}}(X_{j^*})| = D_{r-1}$  and that the sets  $\{A_{j^*,i}^r\}_{X_i \in \mathbf{N}_{T_{r-1}}^{\text{in}}(X_{j^*})}$  are based on an  $(M_r, D_{r-1})$ -separator. Hence by Lemma 2.3 (with  $\zeta = \mu/(1 - \beta)$ ) we have at least  $D_{r-1}/15$  many sets  $X_i \in \mathbf{N}_{T_{r-1}}^{\text{in}}(X_{j^*})$  such that

$$|Z_0 \cap A_{j^*,i}^r|, |Z_0 \cap (X_{j^*} \setminus A_{j^*,i}^r)| \geq \frac{\mu}{1 - \beta} \cdot |Z_0 \cap X_{j^*}| \geq \mu \cdot |Z_0|,$$

as is needed for Item (iii).

Since  $\mathcal{Z}$  is a  $\beta$ -refinement of  $\mathcal{X}_{r-1}$ , there are in total at most  $\beta n$  many vertices which are not useful. Hence there are at most  $m_{r-1}/120$  sets  $X_i \in \mathcal{X}_{r-1}$  containing more than  $120\beta|X_i|$  non-useful vertices. Thus in total there are at least

$$D_{r-1}/15 - m_{r-1}/120 \stackrel{(2.4)}{\geq} m_{r-1}/60 - m_{r-1}/120 = m_{r-1}/120$$

many contributing sets. □

Now consider an arbitrary contributing set  $X_i \in \mathcal{X}_{r-1}$ . Set  $Z_A = Z_0 \cap A_{j^*,i}^r$  and  $Z_B = Z_0 \cap (X_{j^*} \setminus A_{j^*,i}^r)$ . First note that by (2.2) we have  $d_{G_{r'}}(Z_A, X_i) = d_{G_{r'}}(Z_B, X_i)$  for every  $r' < r$ , as  $Z_A$  and  $Z_B$  are always part of the same cell in  $\mathcal{X}_{r'}$ . Furthermore, we have for every  $x_a \in Z_A$  and  $x_b \in Z_B$  by (2.5) that  $d_{G_r}(x_a, X_i) = \delta$  and  $d_{G_r}(x_b, X_i) = 0$  and by

Observation 2.5 that  $\sum_{r'=r+1}^s d_{G_{r'}}(x_a, X_i) = \frac{1}{2}\delta(s-r) = \sum_{r'=r+1}^s d_{G_i}(x_b, X_i)$ . Together this leads to

$$\begin{aligned} & d_G(Z_A, X_i) - d_G(Z_B, X_i) \\ &= d_{G_r}(Z_A, X_i) - d_{G_r}(Z_B, X_i) + \sum_{i=r+1}^s (d_{G_i}(Z_A, X_i) - d_{G_i}(Z_B, X_i)) = \delta. \end{aligned} \quad (2.7)$$

Let  $\mathcal{Z}_i = \{Z \in \mathcal{Z} \mid Z \subseteq_\beta X_i\}$ . By Claim 1(i) we have

$$|\mathcal{Z}_i| \geq \frac{|X_i|(1-120\beta)}{|Z_0|} = \frac{\frac{n}{m_{r-1}}(1-120\beta)}{\frac{n}{\ell}} = \frac{\ell(1-120\beta)}{m_{r-1}}. \quad (2.8)$$

Consider  $\tilde{X}_i = \bigcup_{Z \in \mathcal{Z}_i} Z$  and observe for any  $Z' \subseteq Z_0$  that (by using  $\beta < 1/3200$ )

$$|d_G(Z', X_i) - d_G(Z', \tilde{X}_i)| \leq \frac{|X_i \Delta \tilde{X}_i|}{|X_i \cap \tilde{X}_i|} \leq \frac{120\beta|X_i| + \beta|X_i|}{(1-120\beta)|X_i|} < 130\beta. \quad (2.9)$$

Define

$$\mathcal{Z}_i^\varepsilon = \{Z \in \mathcal{Z}_i \mid (Z, Z_0) \text{ is } \varepsilon\text{-degenerate}\}.$$

We have

$$\begin{aligned} \delta &\stackrel{(2.7)}{=} |d_G(Z_A, X_i) - d_G(Z_B, X_i)| \\ &\stackrel{(2.9)}{<} \left| d_G(Z_A, \tilde{X}_i) - d_G(Z_B, \tilde{X}_i) \right| + 260\beta \\ &\leq \frac{1}{|\mathcal{Z}_i|} \sum_{Z \in \mathcal{Z}_i} |d_G(Z_A, Z) - d_G(Z_B, Z)| + 260\beta \\ &\leq \frac{1}{|\mathcal{Z}_i|} \sum_{Z \in \mathcal{Z}_i^\varepsilon} (|d_G(Z_A, Z) - d_G(Z_0, Z)| + |d_G(Z_B, Z) - d_G(Z_0, Z)|) \\ &\quad + \frac{1}{|\mathcal{Z}_i|} \sum_{Z \in \mathcal{Z}_i \setminus \mathcal{Z}_i^\varepsilon} |d_G(Z_A, Z) - d_G(Z_B, Z)| + 260\beta. \end{aligned}$$

We can use Fact 2.6 on each summand with  $Z \in \mathcal{Z}_i^\varepsilon$ . To this end, we use Claim 1(iii) which tells us that both sets  $Z_A$  and  $Z_B$  span at least a  $\mu$ -fraction of  $Z_0$ . For each summand  $Z \in \mathcal{Z}_i \setminus \mathcal{Z}_i^\varepsilon$  we use that the densities are between 0 and  $0.9 + s\delta \leq 1$ . Therefore,

$$\delta \leq 2 \frac{|\mathcal{Z}_i^\varepsilon|}{|\mathcal{Z}_i|} (\varepsilon + \frac{\varepsilon}{\mu}) + (1 - \frac{|\mathcal{Z}_i^\varepsilon|}{|\mathcal{Z}_i|}) 1 + 260\beta.$$

Using the assumptions of the lemma, we get  $|\mathcal{Z}_i^\varepsilon| \leq (1 - \frac{1}{4}\delta)|\mathcal{Z}_i|$ . Hence, for each contributing  $X_i$ , we have at least

$$\frac{\delta}{4} |\mathcal{Z}_i| \stackrel{(2.8)}{\geq} \frac{\delta}{5} \ell / m_{r-1}$$

many pairs  $(Z, Z_0)$  with  $Z \in \mathcal{Z}_i$  that are not  $\varepsilon$ -degenerate. The families  $\mathcal{Z}_i$  are mutually disjoint for different contributing  $X_i$ 's. Therefore by Claim 1 we have in total at least  $m_{r-1}/120 \cdot \frac{\delta}{5} \ell / m_{r-1} = \delta\ell/600 > \varepsilon\ell$  many non- $\varepsilon$ -degenerate pairs  $(Z_0, Z)$  with  $Z \in \mathcal{Z}$ . Hence,  $\mathcal{Z}$  violates the degeneracy counterpart of (P2), a contradiction.  $\square$

We extend the definition of the tower function from base 2 to an arbitrary base  $a > 1$ ,  $\text{tow}_a : [0, \infty) \rightarrow (0, +\infty)$ ,  $\text{tow}_a(x) = 1$  for  $x \in [0, 1)$  and  $\text{tow}_a(1 + x) = a^{\text{tow}(x)}$  for  $x \geq 1$ . Note that for every  $a > 0$  we have  $\text{tow}_a(s) \geq \text{tow}(s - O(1))$  as  $s \rightarrow \infty$ .

**Theorem 2.8.** *There exists  $\varepsilon_0$  such that for every  $\varepsilon \in (0, \varepsilon_0)$  and every  $N \in \mathbb{N}$  there exists a weighted graph  $G$  of order  $n \geq N$  such that every non-trivial  $\varepsilon$ -degener partition  $\mathcal{Z}$  of  $G$  has complexity at least  $\text{tow}_a(s)$  where  $a = 2^{1/(4 \cdot 9999)}$  and  $s = \varepsilon^{-1/3} \cdot 10^{-6}$ .*

*Proof.* Set  $\delta = \varepsilon^{1/3}$ ,  $\mu = 33\varepsilon^{2/3}$  and  $s = \lfloor \varepsilon^{-1/3}/10^6 \rfloor$ . For  $\varepsilon$  small enough the following holds. Choose  $n \geq \max\{m_s, N\}$  with  $m_s \mid n$  and let  $G$  be the weighted graph given by Construction 2.4 with parameters  $\delta$ ,  $s$  and  $n$  and let  $\mathcal{Z}$  be a non-trivial  $\varepsilon$ -degener partition of  $G$  of complexity  $\ell$ .

**Claim 2.** *The partition  $\mathcal{Z}$  is an  $\varepsilon$ -refinement of  $\mathcal{X}_0 = \{X_1, X_2\}$ .*

*Proof of Claim 2.* Note that by construction and choice of  $\delta$  and  $s$  we have for every  $X \subseteq V(G)$  and  $x \in V(G)$  that

$$\text{deg}_G(x, X) = \sum_{r \in [s]} \text{deg}_{G_r}(x, X) = \text{deg}_{G_1}(x, X) \pm 0.1|X|. \tag{2.10}$$

Now suppose that there exists  $Z_0 \in \mathcal{Z}$  such that  $Z_0 \not\subseteq_\varepsilon X_i$  for  $i \in [2]$ . Then for every other  $Z \in \mathcal{Z} \setminus \{Z_0\}$  let  $A^- \subseteq Z_0$  with  $|A^-| \leq \varepsilon|Z_0|$  and  $B^- \subseteq Z$  with  $|B^-| \leq \varepsilon|Z|$  be such that (1.3) holds in the graph  $G$ . There exist  $x_1 \in (Z_0 \setminus A^-) \cap X_1$  and  $x_2 \in (Z_0 \setminus A^-) \cap X_2$ , and we fix them arbitrarily. For every  $Z \in \mathcal{Z}$  such that  $(Z_0, Z)$  is  $\varepsilon$ -degener in  $G$  we observe that

$$\begin{aligned} \boxed{\text{by (1.3)}} \quad & |\text{deg}_G(x_1, Z) - \text{deg}_G(x_2, Z)| \leq 2\varepsilon|Z|, \\ \boxed{\text{by (2.6) and (2.10)}} \quad & \text{deg}_G(x_1, Z) = 0.1|Z \cap X_1| \pm 0.1|Z|, \\ \boxed{\text{by (2.6) and (2.10)}} \quad & \text{deg}_G(x_2, Z) = 0.9|Z \cap X_2| \pm 0.1|Z|. \end{aligned}$$

Put together, we see that

$$0.1|Z \cap X_1| \geq 0.9|Z \cap X_2| - 0.3|Z| = 0.9(|Z| - |Z \cap X_1|) - 0.3|Z|,$$

which gives  $|Z \cap X_1| \geq 0.6|Z|$ . This leads to

$$\begin{aligned} |X_1| &= \sum_{Z \in \mathcal{Z}} |Z \cap X_1| \\ \boxed{\text{by (P1), (P2)}} \quad & \geq (1 - \varepsilon)\ell \cdot \frac{0.6n}{\ell} > n/2, \end{aligned}$$

contradicting that  $|X_1| = |X_2| = n/2$ . □

Note that  $s\delta \leq 0.1$ ,  $32\varepsilon/\mu < \varepsilon^{1/3} = \delta$ ,  $4000\varepsilon < \delta$  and for every  $r \in [s]$  and  $\beta = \varepsilon + 8\mu r$  we have  $1600\beta = 1600 \cdot (\varepsilon + 8\mu r) < 1600 \cdot 9\mu s \leq \frac{1600 \cdot 9 \cdot 33}{10^6} \varepsilon^{1/3} < \delta$ . Starting with Claim 2 and applying Lemma 2.7 repeatedly we get that  $\mathcal{Z}$  is an  $(\varepsilon + r \cdot 8\mu)$ -refinement of  $\mathcal{X}_r$  for every  $r \in [s]$ . Hence  $\mathcal{Z}$  is a  $(\varepsilon + s \cdot 8\mu) \leq 1/2$ -refinement of  $\mathcal{X}_s$  and therefore  $|\mathcal{Z}| \geq |\mathcal{X}_s|/2 = m_s/2 \geq \text{tow}_a(s)$ , as was needed. □

Finally, we can show Theorem 1.2 by turning the weighted graph from Construction 2.4 into an unweighted graph using a simple probabilistic argument.

*Theorem 2.8 implies Theorem 1.2.* Given  $\varepsilon' > 0$  sufficiently small. We set  $\varepsilon = 4\varepsilon'$  and choose a constant  $\zeta$  sufficiently small with respect to  $\varepsilon$  and  $n$  sufficiently large, i.e.

$$\frac{1}{n} \ll \zeta \ll \varepsilon .$$

We define  $s$  and  $a$  as in Theorem 2.8. In particular we may assume that  $20\zeta^{-2} \log(n) < \varepsilon n/\ell$  for every  $\ell \leq \text{tow}_a(s)$ . The following claim taken from [MS16] is well-known and follows immediately from Chernoff's inequality.

**Claim 3** (Claim 2.1 from [MS16]). *Let  $\zeta > 0$ . Suppose  $G$  is a weighted complete graph on  $n$  vertices with weights in  $[0, 1]$ , and  $G'$  is a random graph, where each edge  $(x, y)$  is chosen independently to be included in  $G'$  with probability  $d_G(x, y)$ . Then with probability at least  $1/2$  we have  $|d_{G'}(A, B) - d_G(A, B)| \leq \zeta$  for all sets  $A, B$  of size at least  $20\zeta^{-2} \log(n)$ .*

Let  $G$  be given by Theorem 2.8 such that every non-trivial  $\varepsilon$ -degenerate partition of  $G$  has complexity at least  $\ell > \text{tow}_a(s)$ . Let  $G'$  be a graph, where each edge  $(x, y)$  is chosen independently to be included in  $G'$  with probability  $d_G(x, y)$ , and where we fix an outcome of this random selection such that the assertion of Claim 3 holds. Now to show Theorem 1.2 it suffices to show that every  $\varepsilon'$ -degenerate partition of  $G'$  of complexity  $\ell \leq \text{tow}_a(s)$  is also an  $\varepsilon$ -degenerate partition of  $G$  and therefore implying that its complexity is at least  $\text{tow}_a(s)$ . For this suppose that  $(A, B)$  is a  $\varepsilon'$ -degenerate pair in  $G'$  and such that  $|A| = |B| = n/\ell$ . We show that  $(A, B)$  is also an  $4\varepsilon'$ -degenerate pair in  $G$ .

Let  $A^* := \{a \in A : |\deg_{G'}(a, B) - \frac{e_{G'}(A, B)}{|A|}| > 4\varepsilon'|B|\}$  and  $B^*$  be defined similarly. We need to prove that  $|A^*| \leq 4\varepsilon'|A|$  and  $|B^*| \leq 4\varepsilon'|B|$ . We will show this argument only for  $A^*$ , the argument for  $B^*$  follows *mutatis mutandis*. Suppose for contradiction that  $|A^*| > 4\varepsilon'|A|$ .

Let  $A^- \subseteq A$  with  $|A^-| \leq \varepsilon'|A|$  be such that  $\deg_G(a, B) = \frac{e_G(A, B)}{|A|} \pm \varepsilon'|B|$  for every  $a \in A \setminus A^-$ . Without loss of generality we may then assume that there is a set  $A' \subseteq A^*$  of size at least  $2\varepsilon'|A|$  and such that  $\deg_{G'}(a, B) > \frac{e_{G'}(A, B)}{|A|} + 4\varepsilon'|B|$  for every  $a \in A'$ . Let  $\tilde{A} = A' \setminus A^-$ . We have  $|\tilde{A}| > \varepsilon|A|$  and

$$\begin{aligned} d_G(\tilde{A}, B) &= \\ \boxed{|\tilde{A}| > \varepsilon' n/\ell > 20\zeta^{-2} \log(n), \text{ Claim 3 applies}} &= d_{G'}(\tilde{A}, B) \pm \zeta \\ \boxed{(A, B) \text{ is } \varepsilon'\text{-degenerate in } G'} &= d_{G'}(A, B) \pm (\varepsilon' + \zeta) \\ \boxed{\text{Claim 3}} &= d_G(A, B) \pm (\varepsilon' + 2\zeta) < d_G(A, B) + 4\varepsilon' \\ \boxed{\text{definition of } \tilde{A}} &< d_G(\tilde{A}, B), \end{aligned}$$

a contradiction. □

## Acknowledgements

We thank Timothy Gowers for discussions and an anonymous referee for their remarks.

## References

- [BBCR10] Béla Bollobás, Christian Borgs, Jennifer Chayes, and Oliver Riordan. Percolation on dense graph sequences. *Ann. Probab.*, 38(1):150–183, 2010. doi:10.1214/09-AOP478.
- [Bol98] Béla Bollobás. The work of William Timothy Gowers. In *Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998)*, pages 109–118, 1998.
- [CF12] David Conlon and Jacob Fox. Bounds for graph regularity and removal lemmas. *Geom. Funct. Anal.*, 22(5):1191–1256, 2012. doi:10.1007/s00039-012-0171-x.
- [CKKN18] Jacob W. Cooper, Tomáš Kaiser, Daniel Král', and Jonathan A. Noel. Weak regularity and finitely forcible graph limits. *Trans. Amer. Math. Soc.*, 370(6):3833–3864, 2018. doi:10.1090/tran/7066.
- [Csa21] Béla Csaba. A new graph decomposition method for bipartite graphs. 2021. arXiv:2109.12429.
- [FK99] Alan Frieze and Ravi Kannan. Quick approximation to matrices and applications. *Combinatorica*, 19(2):175–220, 1999. doi:10.1007/s004930050052.
- [FL17] Jacob Fox and László Miklós Lovász. A tight lower bound for Szemerédi's regularity lemma. *Combinatorica*, 37(5):911–951, 2017. doi:10.1007/s00493-016-3274-4.
- [FR92] P. Frankl and V. Rödl. The uniformity lemma for hypergraphs. *Graphs Combin.*, 8(4):309–312, 1992. doi:10.1007/BF02351586.
- [Gow97] W. T. Gowers. Lower bounds of tower type for Szemerédi's uniformity lemma. *Geom. Funct. Anal.*, 7(2):322–337, 1997. doi:10.1007/PL00001621.
- [GR22] Jan Grebík and Israel Rocha. Fractional isomorphism of graphons. *Combinatorica*, 42(3):365–404, 2022. doi:10.1007/s00493-021-4336-9.
- [HH23] Jan Hladký and Eng Keat Hng. Approximating fractionally isomorphic graphons. *European J. Combin.*, 113:Paper No. 103751, 2023. doi:10.1016/j.ejc.2023.103751.
- [HNT18] Jan Hladký, Asaf Nachmias, and Tuan Tran. The local limit of the uniform spanning tree on dense graphs. *J. Stat. Phys.*, 173(3-4):502–545, 2018. doi:10.1007/s10955-017-1933-5.
- [Łu99] Tomasz Łuczak.  $R(C_n, C_n, C_n) \leq (4 + o(1))n$ . *J. Combin. Theory Ser. B*, 75(2):174–187, 1999. doi:10.1006/jctb.1998.1874.
- [MS16] Guy Moshkovitz and Asaf Shapira. A short proof of Gowers' lower bound for the regularity lemma. *Combinatorica*, 36(2):187–194, 2016. doi:10.1007/s00493-014-3166-4.
- [Sze78] Endre Szemerédi. Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976)*, volume 260 of *Colloq. Internat. CNRS*, pages 399–401. CNRS, Paris, 1978.