TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 371, Number 10, 15 May 2019, Pages 6779–6814
https://doi.org/10.1090/tran/7414

A SPARSE REGULAR APPROXIMATION LEMMA

GUY MOSHKOVITZ AND ASAF SHAPIRA

Abstract. We introduce a new variant of Szemerédi's regularity lemma which we call the *sparse regular approximation lemma* (SRAL). The input to this lemma is a graph G of edge density p and parameters ϵ, δ, where we think of δ as a constant. The goal is to construct an ϵ-regular partition of G while having the freedom to add/remove up to $\delta|E(G)|$ edges. As we show here, this weaker variant of the regularity lemma already suffices for proving the graph removal lemma and the hypergraph regularity lemma, which are two of the main applications of the (standard) regularity lemma. This of course raises the following question: can one obtain quantitative bounds for SRAL that are significantly better than those associated with the regularity lemma?

Our first result answers the above question affirmatively by proving an upper bound for SRAL given by a tower of height $O(\log 1/p)$. This allows us to reprove Fox's upper bound for the graph removal lemma. Our second result is a matching lower bound for SRAL showing that a tower of height $\Omega(\log 1/p)$ is unavoidable. We in fact prove a more general multicolored lower bound which is essential for proving lower bounds for the hypergraph regularity lemma.

1. Introduction

Szemerédi's regularity lemma [22] asserts that every graph can be partitioned into a bounded number of vertex sets $Z_1, \dots, Z_k$ so that the bipartite graphs between almost all pairs (Z_i, Z_j) behave "randomly". More precisely, for every $\epsilon > 0$ there is a smallest integer $M = M(\epsilon)$ such that for every graph, and every vertex equipartition $\mathcal{P}_0$ of order at most $1/\epsilon$, there is an equipartition $\mathcal{Z}$ that refines $\mathcal{P}_0$, is ϵ-regular, and has order at most M.[1] The precise definitions of the above standard notions are given in Section 2. The regularity lemma has become one of the most widely used tools in extremal graph theory, as well as in many other fields. See [12] for a survey. Unfortunately, the proof in [22] gave $M(\epsilon) \le \text{twr}(\text{poly}(1/\epsilon))$, where $\text{twr}(x)$ is a tower of twos of height x. A celebrated result of Gowers [8] states that $M(\epsilon)$ indeed grows as $\text{twr}(\text{poly}(1/\epsilon))$. Hence, the applications of the lemma are all of asymptotic nature and supply very weak quantitative bounds.

It has long been observed that in some cases one does not need the full strength of Szemerédi's lemma. For example, when one is only interested in global counts such as the total number of triangles in a graph or the size of the largest cut, then far weaker notions of regularity suffice. Two examples are the so-called *weak*

Received by the editors November 3, 2016, and, in revised form, August 28, 2017.

2010 *Mathematics Subject Classification.* Primary 05C35, 05D99.

The second author was supported in part by ISF Grant 1028/16 and ERC-Starting Grant 633509.

[1] One can use an independent parameter for the upper bound on the size of $\mathcal{P}_0$ rather than the $1/\epsilon$ we have here. For the sake of simplicity we decided to drop this parameter as it never has any real effect on the quantitative bound.

regularity lemma of Frieze and Kannan [7] (see Section 2 for more details) and the *cylinder regularity lemma* of Duke, Lefmann and Rödl [4]. The main advantage of these relaxed regularity lemmas is that the bounds involved are far better than the $\mathrm{twr}(\mathrm{poly}(1/\epsilon))$ bounds that are usually obtained when applying the regularity lemma. For example, the above-mentioned variants of the regularity lemma have bounds that are only exponential in a power of $1/\epsilon$.

Our main objective in this paper is to introduce and study a new relaxed notion of regularity. As we will show, this relaxed version of the lemma will turn out to be strong enough to imply two of the most important applications of the regularity lemma, while at the same time be weak enough to have bounds better than $\mathrm{twr}(\mathrm{poly}(1/\epsilon))$. The idea in this relaxation of the regularity lemma is to allow the freedom to modify a small *percentage* of the graph's edges. We call this new variant the *sparse regular approximation lemma* or SRAL for short. The precise definition is the following.

Definition 1.1. For every $\epsilon, \delta, p > 0$ let $S = S(\epsilon, \delta, p)$ be the smallest integer such that if G is a graph of density at least p, and $\mathcal{P}_0$ is an equipartition of $V(G)$ of order at most $1/\epsilon$, then one can add/remove at most $\delta|E(G)|$ edges and thus turn G into a graph that has an ϵ-regular equipartition that refines $\mathcal{P}_0$ and has order at most S.

Let us make some simple observations regarding the above definition. We first note that trivially $S(\epsilon, \delta, p) \leq M(\epsilon)$ since one can just apply the usual regularity lemma without taking advantage of G's sparseness and of the freedom to modify G. In particular the function S is well defined.

It is natural to ask if one can take advantage of the sparseness of G even without using the freedom to modify its edges. As it turns out, this is not the case. It follows from the construction in [14] that for every p and $\epsilon = p^{12}$, there is a graph G of edge density p such that every ϵ-regular partition of G has order $\mathrm{twr}(\mathrm{poly}(1/\epsilon))$. In other words, even when $\epsilon = \mathrm{poly}(p)$, if one wants to beat the $\mathrm{twr}(\mathrm{poly}(1/\epsilon))$ bound that follows from simply applying the usual regularity lemma, then one has to modify G. Let us also observe that if we allow δ to depend on ϵ, say if $\delta = \epsilon^4$, then $S(\epsilon, \epsilon^4, p) \geq M(2\epsilon) \geq \mathrm{twr}(\mathrm{poly}(1/\epsilon))$. Indeed, this follows from the simple observation that an ϵ-regular bipartite graph remains 2ϵ-regular if only an ϵ^3-fraction of the possible edges are added/removed.[2] At the other extreme, we trivially have $S(\epsilon, 1, p) = 1/\epsilon$.

Hence, the main interest in SRAL is when $\delta < 1$ is constant. As we show below, even in this case SRAL has some unexpected applications. In fact, SRAL will be interesting even when $\epsilon = \mathrm{poly}(p)$, hence our main interest will be in bounding the function $S(\mathrm{poly}(p), \delta_0, p)$ for constant δ_0.

1.1. **Applications of SRAL.**

Let us now explain the main motivation for introducing the sparse regular approximation lemma (SRAL). One of the first, and most important, applications of the regularity lemma is the graph removal lemma of Ruzsa and Szemerédi [21], which states that for every fixed graph H there is a function $\mathrm{Rem}_H(\epsilon)$ such that if one must remove from an n-vertex graph G at least ϵn^2 edges in order to make it H-free, then G contains at least $n^h/\mathrm{Rem}_H(\epsilon)$

---

[2] Observe that we can combine the above two observations and get that for every p, $\epsilon = p^{12}$ and $\delta = \epsilon^4 = p^{48}$, there is a graph of density p such that even after adding/removing $\delta|E(G)|$ edges, every ϵ-regular partition of the resulting graph has order at least $\mathrm{twr}(\mathrm{poly}(1/\epsilon))$.

copies of H, where $h = |V(H)|$. The standard proof of the removal lemma, via the regularity lemma, establishes the bound $\mathrm{Rem}_H(\epsilon) \le M(\mathrm{poly}(\epsilon)) = \mathrm{twr}(\mathrm{poly}(1/\epsilon))$. Our first motivation for introducing SRAL is that one can in fact prove the removal lemma using SRAL. This is stated explicitly in the following theorem.

Theorem 1. *For every $h \ge 3$ there are $\epsilon_0, \delta_0, C > 0$ such that if H is a graph on h vertices and $\epsilon \le \epsilon_0$, then*

$$\mathrm{Rem}_H(\epsilon) \le [S(\epsilon^C,\ \delta_0,\ \epsilon)]^C\ . \tag{1}$$

The proof of Theorem 1 is much more delicate than the usual proof of the removal lemma via the standard regularity lemma, mainly due to having to work with a modified version of the input graph. In particular, we will need to prove a counting lemma which is suitable for SRAL; see Lemma 4.1.

Our second motivation for studying SRAL is the hypergraph regularity lemma [9, 15, 19, 23]. For simplicity, we focus on the regularity lemma for 3-uniform hypergraphs (3-graphs for short), such as the one obtained by Frankl and Rödl [6], refraining from giving the exact definition of 3-graph regularity. Since all proofs of the regularity lemma for 3-graphs proceed by repeatedly applying the graph regularity lemma, they all produce partitions whose order is given by a wowzer-type bound, that is, an iterated-tower bound. It is a major open problem to decide if one can obtain tower-type bounds for the 3-graph regularity lemma, and more generally for the k-graph regularity lemma. One striking application for such a bound would be primitive recursive bounds for the multidimensional Szemerédi theorem [9], a result which currently has only Ackermann-type upper bounds. As in the case of the removal lemma, we can show that when proving the 3-graph regularity lemma, one can replace the application of the graph regularity lemma with an application of SRAL. In particular, we have the following, where $\mathrm{twr}_y(x)$ is a tower of x twos with y at the top.[3]

Proposition 1.2. *Suppose that for every $C > 0$ there is $c > 0$ such that*

$$S(p^C, \delta, p) \le \mathrm{twr}_{1/p}((1/\delta)^c)\ . \tag{2}$$

Then one can prove a tower-type upper bound for the 3-graph regularity lemma.

The proof of Proposition 1.2 proceeds by redoing the proof of the regularity lemma for 3-graphs [6], while observing that in the critical step when one applies the regularity lemma, it is in fact enough to use SRAL with no effect on the progress of the process of regularizing the hypergraph.

Summarizing, the above two theorems in particular imply that for a fixed δ, proving a bound on $S(\mathrm{poly}(\epsilon), \delta,\ \epsilon)$ which is significantly better than $\mathrm{twr}(\mathrm{poly}(1/\epsilon))$ would have the following immediate consequences. By Theorem 1, this would give an improvement over the standard bound $\mathrm{Rem}_H(\epsilon) \le \mathrm{twr}(\mathrm{poly}(1/\epsilon))$ for the graph removal lemma. By Proposition 1.2, if one can further prove an upper bound for $S(\mathrm{poly}(\epsilon), \delta,\ \epsilon)$ that is given by a bounded number of exponents, then one would significantly improve the bound on 3-graph regularity from wowzer-type to tower-type.

---

[3] So $\mathrm{twr}_y(2) = 2^{2^y}$ and $\mathrm{twr}_{\mathrm{twr}(x)}(x) = \mathrm{twr}(2x)$.

1.2. **The regular approximation lemma.**

Before describing our solution of the above problem, we first describe a related variant of the regularity lemma. As the name SRAL suggests, it is a variant of the so-called *regular approximation lemma* (RAL for short), a special case[4] which can be defined as follows.

Definition 1.3. For every $\epsilon, \delta > 0$ let $T = T(\epsilon, \delta)$ be the smallest integer such that if G is an n-vertex graph and $\mathcal{P}_0$ is an equipartition of $V(G)$ of order at most $1/\epsilon$, then one can add/remove at most δn^2 edges and thus turn G into a graph that has an ϵ-regular equipartition which refines $\mathcal{P}_0$ and has order at most T.

The RAL was introduced as part of the study of graph limits and of the hypergraph regularity lemma by Lovász and Szegedy [13] and Rödl and Schacht [17], respectively. Note that RAL differs from SRAL in that the number of edge modifications is a δ-fraction of n^2 rather than $|E(G)|$. Nonetheless, we still have the trivial relation

$$S(\epsilon, \delta, p) \leq T(\epsilon, \delta p) \ . \tag{3}$$

The upper bounds obtained in [13, 17], when specialized to Definition 1.3, are no better than the trivial $T(\epsilon, \delta) \leq M(\epsilon) = \mathrm{twr}(\mathrm{poly}(1/\epsilon))$ bound that follows from the regularity lemma. A considerably better bound was given by Conlon and Fox [2] who showed that $T(\epsilon, \delta) \leq \mathrm{twr}_{1/\epsilon}(\mathrm{poly}(1/\delta))$. Note that for a fixed δ, this is a fixed number of exponents, which is significantly better than the $\mathrm{twr}(\mathrm{poly}(1/\epsilon))$ bound given by the regularity lemma. Although this bound seems like the one we were aiming for in Proposition 1.2, observe that it only implies, via (3), that when δ is a fixed constant and $\epsilon = \mathrm{poly}(p)$ we have $S(\epsilon, \delta, p) \leq \mathrm{twr}_{1/\epsilon}(\mathrm{poly}(1/\delta p)) = \mathrm{twr}(\mathrm{poly}(1/\epsilon))$, which again does not improve over the regularity lemma.

1.3. **An upper bound for SRAL.**

Our first bound shows that one can improve upon the $\mathrm{twr}(\mathrm{poly}(1/\epsilon))$ bound of the regularity lemma, even when the number of modifications allowed is relative to the graph's density. In particular, we improve the bound given by the regularity lemma when $\epsilon = \mathrm{poly}(p)$, which is the setting of Theorem 1 and Proposition 1.2.

Theorem 2. *There is an absolute constant c such that*

$$S(\epsilon, \delta, p) \leq \mathrm{twr}_{1/\epsilon}(c \log(1/p)/\delta^2).$$

In particular, for every fixed $C, \delta_0 > 0$ we have

$$S(p^C, \delta_0, p) \leq \mathrm{twr}(O(\log(1/p))) \ .$$

Since we trivially have $T(\epsilon, \delta) \leq S(\epsilon, \delta, 1/2)$,[5] Theorem 2 immediately gives as a special case the bound $T(\epsilon, \delta) \leq \mathrm{twr}_{1/\epsilon}(\mathrm{poly}(1/\delta))$ for RAL, which was first proved in [2]. We note that our proof of Theorem 2 gives a much more general result—we can in fact guarantee that the partition is such that all pairs are ϵ-regular and

---

[4] The full-fledged RAL allows one to replace ϵ with an arbitrary function f, so that the equipartition $\mathcal{P}$ is such that all pairs are $f(|\mathcal{P}|)$-regular. See [2] for a detailed discussion. As we will mention later (see Section 3), the proof for SRAL that we give applies to this more general setting almost without any effect on the bounds. We opted to describe the simpler/weaker versions of RAL and SRAL since they suffice for the applications mentioned in Subsection 1.1 and, most importantly, since the lower bounds we will prove hold even in these simpler settings.

[5] Indeed, we can either apply SRAL to G or to its complement.

that ϵ can be taken to be a function of the order of the partition.[6] For the precise statement see Theorem 9 in Section 3.

Our original proof of Theorem 2 applied a method similar to the one used by Scott [20] in his proof of a regularity lemma for sparse graphs. The idea is to build a sequence of partitions $\mathcal{P}_0, \mathcal{P}_1, \dots$ so that $|\mathcal{P}_{i+1}| \leq 2^{\mathrm{poly}(|\mathcal{P}_i|)}$, where in partition $\mathcal{P}_i$ all ϵ-irregular pairs have density at least $2^i p$ (thus, in particular, at most a 2^{-i}-fraction of the pairs are irregular). Since this process terminates after $\log(1/p)$ iterations, we get a bound similar to the one stated in Theorem 2. The main benefit of this proof is that it hints at how one should construct a lower bound for $S(\mathrm{poly}(p), \delta, p)$. See the discussion after Theorem 4.

The actual proof of Theorem 2 we give here uses a different approach which is shorter to prove. It is motivated by the one taken by Conlon and Fox [2], using an iterated version of the weak regularity lemma of Frieze and Kannan [7]. Our proof, however, differs in two important aspects. First, we use (and prove) a new variant of the weak regularity lemma which we need for our purposes. Second, we use the entropy potential function (first used by Fox [5]) together with Pinsker's inequality from information theory in order to control the ℓ_1-distance, *relative to the graph's density*, between partitions with similar entropy potentials. We believe this approach might be useful for studying other variants of the graph and hypergraph removal lemma.

An immediate application of Theorems 1 and 2 gives the following.

Corollary 3. *For every h-vertex graph H we have*

$$\mathrm{Rem}_H(\epsilon) \leq \mathrm{twr}(O(\log(1/\epsilon))) .$$

As is of course well known, the above bound for the removal lemma was first obtained by Fox [5], who was the first to improve upon the $\mathrm{twr}(\mathrm{poly}(1/\epsilon))$ bound that follows from applying the regularity lemma. Fox's breakthrough result relied on a purely ad hoc argument. This argument was simplified by Conlon and Fox [3], where a proof more in the spirit of the regularity lemma was given. We think it is important to see that the same bound can be derived in the framework of the regularity method, meaning by applying an appropriate regularity lemma (namely, Theorem 2) together with an appropriate counting lemma (see Lemma 4.1).

1.4. **A tight lower bound for SRAL.**

Recall that our second motivation for SRAL was the possibility of using it to improve the bounds for hypergraph regularity, as stated in Proposition 1.2. Theorem 2 does allow one to "improve" the bounds for hypergraph regularity by replacing an iterated version of the function $\mathrm{twr}(\mathrm{poly}(1/\epsilon))$ with an iterated version of the function $\mathrm{twr}(\log(1/\epsilon))$. However, the latter is still a wowzer-type function. This, and the possibility of obtaining even better bounds for the removal lemma (via Theorem 1), naturally raise the question if one can obtain even better bounds for SRAL, say, a $\mathrm{twr}_{1/\epsilon}(\mathrm{poly}(1/\delta))$ bound such as the one obtained by Conlon and Fox [2] for RAL. As our second result shows, such an improvement is impossible, even when $\epsilon = p^5$ and δ is a fixed constant.

Theorem 4. *There are fixed constants $\delta_0, c > 0$ such that*

$$S(p^5, \delta_0, p) \geq \mathrm{twr}(c \log(1/p)) . \tag{4}$$

---

[6] As we noted earlier, such stronger properties were also available for previous versions of RAL.

Furthermore, one can decompose the complete bipartite graph into $1/p$ *graphs of density* p *so that each of them witnesses* (4).

Theorems 2 and 4 give us the following tight bound for SRAL.

Corollary 5. *For every fixed* $\delta \leq \delta_0$ *and* $C \geq 5$ *we have*

$$S(p^C, \delta_0, p) = \mathrm{twr}(\Theta(\log(1/p))).$$

The proof of (4) is by far the most complicated part of this paper. While the construction has a (relatively) simple description, proving its correctness requires a very careful analysis, employing some ideas we used in [14], together with those of Gowers [8]. The main difficulty in proving (4) lies in handling an absolute constant[7] δ_0 (we obtain $\delta_0 = 10^{-10}$ but make no effort to optimize it), i.e., even when the graph is very sparse and one is allowed to change a constant fraction of its edges!

It is hard to give a short overview of the proof of Theorem 4 (nonetheless, we try to do so in Subsection 5.1). Let us thus only mention two interesting aspects of it. First, the graph we construct is designed to be "hard" for the proof of Theorem 2 based on the method of [20] (the one we do not describe in this paper). By this we mean that the idea is to show that in order to find an ϵ-regular partition of the graph (or even of a modified version of it), in a sense one cannot avoid executing the process of constructing the sequence of partitions $\mathcal{P}_i$ with the properties mentioned in the previous subsection. A second interesting aspect is that although we want to show that the graph has no small p^5-regular partition (even after modifying it), it *does* essentially have a $p^{\frac{8}{7}}$-regular partition of size 2, namely the graph itself is quite quasirandom. This property is key to the analysis of the construction.

1.5. **An approach for hypergraph regularity lower bounds.**

Returning to Proposition 1.2, inequality (4) implies that one cannot prove a tower-type upper bound for 3-graph regularity even if using SRAL instead of the regularity lemma. However, as we explain below, we believe that an even more important aspect of Theorem 4 is in being a major step towards showing that such an improvement is actually impossible.

All proofs of the 3-graph regularity lemma proceed by iterating the graph regularity lemma, and more generally, all proofs of the k-graph regularity iterate the $(k-1)$-graph regularity lemma. Yet, it seems that a lower bound proof for 3-graph regularity does not follow by iterating a lower bound for the graph regularity lemma. This can be explained by the fact that 3-graph regularity can already be proved by iterating SRAL (as mentioned in the discussion leading to Proposition 1.2), which implies that any proof of a wowzer-type lower bound for 3-graphs would have to give, at least implicitly, a tower-type bound for SRAL. It therefore seems to us that the correct approach for proving 3-graph lower bounds is by iterating the SRAL lower bound instead. More generally, we suggest that in order to prove lower bounds for the k-graph regularity lemma, one should "strengthen the induction hypothesis", that is, prove by induction a stronger statement—that k-graph SRAL requires a partition whose order is given by the kth level in the Ackermann hierarchy. One can thus view Theorem 4 as the induction basis in such a program.

---

[7] As we remarked earlier, it is easy to give tower-type lower bounds for $S(\epsilon, \delta, p)$ if one allows δ to depend on p.

We intend to return to this subject in the near future. We give more details regarding the relevance of Theorem 4 to lower bounds for hypergraph regularity in Subsection 5.7.

1.6. **Paper organization.** The rest of the paper is organized as follows. In Section 2 we define a variant of the notion of weak regularity, state the corresponding regularity lemma, and prove that a weak regular partition can be made regular by making an appropriate number of edge modifications. The upper bound for SRAL, stated in Theorem 2, is proved in Section 3 using an iterated weak regularity lemma together with a new *sparse* defect inequality. Our reduction of the removal lemma to SRAL, stated in Theorem 1, is proved in Section 4 using a variant of the well-known counting lemma suitable for applying it together with SRAL. Finally, the lower bound for SRAL, stated in Theorem 4, is proved in Section 5. Regarding Proposition 1.2, since Theorem 4 implies that the bound stipulated in (2) does not hold, and since proving Proposition 1.2 would require reproving the 3-graph regularity lemma in its entirety, we felt that including its proof would be redundant.

2. From weak regularity to regularity

In this section we introduce a stronger notion of weak regularity and prove an upper bound on the number of edge modifications required to turn a weak regular bipartite graph into a regular graph.

2.1. **Preliminaries.** We use the following definitions in this section and throughout the paper. The density between two vertex subsets A, B in a graph G is $d_G(A,B) = e_G(A,B)/|A||B|$, where $e_G(A,B)$ is the number of ordered pairs $(u,v) \in A \times B$ with u connected to v. We say that the pair (A,B) is *ϵ-regular* if $|d_G(A,B) - d_G(A',B')| \le \epsilon$ for all $A' \subseteq A$ and $B' \subseteq B$ satisfying $|A'| \ge \epsilon|A|$ and $|B'| \ge \epsilon|B|$. A vertex *equipartition*[8] $\mathcal{Z} = \{Z_1, \dots, Z_k\}$ of G is *ϵ-regular* if all pairs (Z_i, Z_j) but at most ϵk^2 are ϵ-regular. The *order* of $\mathcal{Z}$ is k.

Suppose $G = (V,E)$. We say that $G' = (V,E')$ is *δ-close* to G if G' can be obtained from G by adding and/or removing at most $\delta|E|$ edges (i.e., $|E \triangle E'| \le \delta|E|$). The density of G is $d_G := 2|E|/|V|^2$. We sometimes write $e_G(x,A)$ for $e_G(\{x\},A)$. For partitions $\mathcal{P}, \mathcal{Q}$ we write $\mathcal{Q} \preceq \mathcal{P}$ if $\mathcal{Q}$ is a refinement of $\mathcal{P}$ (i.e., each part of $\mathcal{Q}$ is contained in a part of $\mathcal{P}$).

2.2. **Weak regularity.** The notion of weak regularity was introduced by Frieze and Kannan [7], and is crucial for the proof of Theorem 2. In our proof we will require a somewhat stronger notion than usual, as follows.

Definition 2.1. Given a graph $G = (V,E)$, a partition $\{V_1, \dots, V_k\}$ of V is *weak ϵ-regular* if for all disjoint sets $S, T \subseteq V$ with $|S|, |T| \ge \epsilon|V|$ we have, denoting $S_i = S \cap V_i$ and $T_i = T \cap V_i$,

$$\sum_{i,j=1}^{k} \frac{|S_i|\,|T_j|}{|S|\,|T|} \left| d(S_i,T_j) - d(V_i,V_j) \right| \le \epsilon \,.$$

For comparison, in the usual definition of a weak ϵ-regular partition we have

$$\left| d(S,T) - \sum_{i,j=1}^{k} \frac{|S_i|\,|T_j|}{|S|\,|T|} d(V_i,V_j) \right| = \left| \sum_{i,j=1}^{k} \frac{|S_i|\,|T_j|}{|S|\,|T|} \big(d(S_i,T_j) - d(V_i,V_j)\big) \right| \le \epsilon \,,$$

---

[8] $\mathcal{Z}$ is an equipartition (or simply *equitable*) if the sizes of all parts Z_i differ by at most 1.

that is, ϵ bounds the deviation of the average *difference* of $d(S_i, T_j)$ from its expected value. In contrast, ϵ in Definition 2.1 even bounds the average *deviation* of $d(S_i, T_j)$ from its expected value.

The weak regularity lemma asserts that every graph has a weak ϵ-regular partition whose order depends merely exponentially on $1/\epsilon$, as opposed to the tower-type dependence on $1/\epsilon$ in the usual regularity lemma.

Theorem 6. *Let $\epsilon > 0$. For every graph and initial vertex equipartition $\mathcal{P}_0$ there is a weak ϵ-regular equipartition (in the sense of Definition 2.1) refining $\mathcal{P}_0$ of order at most $|\mathcal{P}_0| \cdot 2^{\mathrm{poly}(1/\epsilon)}$.*

We note that we made no effort to optimize the bound in Theorem 6. The proof of this (stronger) weak regularity lemma is almost identical to the proof of the Frieze–Kannan weak regularity lemma, and for completeness we give the full proof in the appendix.

2.3. **Perturbation lemma.**

The main ingredient in the proof of Theorem 2 is a lemma showing that any weak regular partition can be made into a "genuine" regular partition by applying an appropriate perturbation. This is formally stated in the following lemma.

Lemma 2.2. *Let G be a bipartite graph of density d with vertex classes (A, B), and let $\mathcal{A} \cup \mathcal{B}$ be a weak ϵ-regular partition of G, where $\mathcal{A} = \{A_i\}_i$ and $\mathcal{B} = \{B_j\}_j$ partition A and B, respectively. That is, for every $S \subseteq A$, $T \subseteq B$ with $|S| \geq \epsilon |A|, |T| \geq \epsilon |B|$ we have, denoting $S_i = S \cap A_i$, $T_j = T \cap B_j$, and $d_{i,j} = d_G(A_i, B_j)$, that*

$$\sum_{i,j} \frac{|S_i|\,|T_j|}{|S|\,|T|} \left| d_G(S_i, T_j) - d_{i,j} \right| \leq \epsilon \,. \tag{5}$$

If $|A|, |B| \geq 8/\epsilon^4$, one can turn G into a 2ϵ-regular graph $\widetilde{G}$ by modifying at most Δ edges where

$$\Delta = \sum_{i,j} |d_{i,j} - d|\,|A_i||B_j| \,.$$

Proof. The idea is to add/remove edges between each pair (A_i, B_j) so as to equate their densities to d_G. We show that if this is done in a random manner, then, with high probability, the modified graph is 2ϵ-regular.[9] We henceforth assume $\epsilon \leq 1/2$, as otherwise there is nothing to prove.

Formally, we do the following for each pair (A_i, B_j). If $d_{i,j} = d$ we do nothing. If $d_{i,j} > d$ we remove each edge of G between A_i and B_j independently with probability $p_{i,j} := \frac{d_{i,j} - d}{d_{i,j}}$. If $d_{i,j} < d$ we add each nonedge of G between A_i and B_j independently with probability $p'_{i,j} := \frac{d - d_{i,j}}{1 - d_{i,j}}$. Let G' be the random graph obtained from G after applying the above procedure for all pairs (A_i, B_j). Clearly $\mathbb{E} d_{G'}(A_i, B_j) = d$ for every i, j, and so

$$\mathbb{E} d_{G'} = d. \tag{6}$$

Moreover, the number $|G' \triangle G|$ of edge modifications thus made satisfies

$$\mathbb{E}|G' \triangle G| = \sum_{i,j} |d_{i,j} - d|\,|A_i|\,|B_j| = \Delta \,.$$

---

[9] In fact, $(\epsilon + o(1))$-regular as $|V(G)| \to \infty$.

Since the random variable $|G'\triangle G|$ is a sum of (at most) $|A|\,|B|$ mutually independent indicator random variables, we have by Chernoff's inequality that

$$\mathbb{P}\big[|G'\triangle G|-\Delta>\epsilon^3|A||B|\big]<\exp\big(-2(\epsilon^3|A||B|)^2/\,|A|\,|B|\,\big)=\exp(-2\epsilon^6\,|A|\,|B|)\leq 1/6, \tag{7}$$

where the last inequality follows from the lemma's assumption that $|A|,|B|\geq 1/\epsilon^4$. Furthermore, for $S\subseteq A$, $T\subseteq B$ with $|S|\geq 2\epsilon|A|$ and $|T|\geq 2\epsilon\,|B|$, the random variable $e_{G'}(S,T)$ is a sum of (at most) $|S|\,|T|$ mutually independent indicator random variables, so by Chernoff's inequality,

$$\begin{aligned}\mathbb{P}\big[|e_{G'}(S,T)-\mathbb{E}e_{G'}(S,T)|>(\epsilon/4)|S||T|\big]&<2\exp(-2(\epsilon/4)^2\,|S|\,|T|)\\&\leq 2\exp(-\epsilon^4\,|A|\,|B|\,/2)\ .\end{aligned} \tag{8}$$

Note that the same bounds applies to $e_{G'}(A,B)$, that is,

$$\mathbb{P}\big[|e_{G'}(A,B)-\mathbb{E}e_{G'}(A,B)|>(\epsilon/4)|A||B|\big]<2\exp(-\epsilon^4\,|A|\,|B|\,/2)\leq 2\cdot 1/6\ . \tag{9}$$

(The last inequality above may be deduced from the last inequality in (7) as $\epsilon^4/2\geq 2\epsilon^6$.) Applying the union bound on (8), we get

$$\begin{aligned}\mathbb{P}\big[\exists S,T:\,|d_{G'}(S,T)-\mathbb{E}d_{G'}(S,T)|>\epsilon/4\big]&<2^{|A|+|B|}\cdot 2\cdot 2^{-\epsilon^4|A||B|/2}\\&\leq 2\cdot 2^{|A|(2-\epsilon^4|B|/2)}\leq 2\cdot 2^{-2|A|}\leq 1/2\end{aligned} \tag{10}$$

with S,T as above (i.e., $|S|\geq 2\epsilon|A|,|T|\geq 2\epsilon|B|$), where in the first inequality we assumed $|A|\geq|B|$ without loss of generality, and in the second inequality we used the assumption that $|A|\,,|B|\geq 8/\epsilon^4$.

Henceforth, let $S\subseteq A$, $T\subseteq B$ satisfy $|S|\geq 2\epsilon\,|A|\,,\,|T|\geq 2\epsilon\,|B|$. The crux of the proof is the claim that

$$|\mathbb{E}d_{G'}(S,T)-d|\leq\epsilon\ . \tag{11}$$

For this we will first need to prove that for any $X\subseteq A_i$, $Y\subseteq B_j$ we have

$$|\mathbb{E}d_{G'}(X,Y)-d|\leq|d_G(X,Y)-d_{i,j}|\ . \tag{12}$$

Recalling the construction of G' at the beginning of the proof, we need to consider three cases. First, if $d_{i,j}=d$, then (12) is trivial. Second, if $d_{i,j}>d$, then, setting $q_{i,j}:=1-p_{i,j}=\frac{d}{d_{i,j}}$, we have

$$|\mathbb{E}d_{G'}(X,Y)-d|=|q_{i,j}d_G(X,Y)-d|=q_{i,j}\,|d_G(X,Y)-d_{i,j}|\leq|d_G(X,Y)-d_{i,j}|\ .$$

Finally, if $d_{i,j}<d$ then, setting $q'_{i,j}:=1-p'_{i,j}=\frac{1-d}{1-d_{i,j}}$, we have

$$\begin{aligned}|\mathbb{E}d_{G'}(X,Y)-d|&=\big|d_G(X,Y)+p'_{i,j}(1-d_G(X,Y))-d\big|=\big|d_G(X,Y)q'_{i,j}+p'_{i,j}-d\big|\\&=\big|d_G(X,Y)q'_{i,j}-q'_{i,j}+(1-d)\big|=q'_{i,j}\,|d_G(X,Y)-d_{i,j}|\\&\leq|d_G(X,Y)-d_{i,j}|\ .\end{aligned}$$

Having established (12), we now prove (11). Denoting $S_i=S\cap A_i$ and $T_j=T\cap B_j$, we indeed have

$$\begin{aligned}|\mathbb{E}d_{G'}(S,T)-d|&=\left|\sum_{i,j}\frac{|S_i|\,|T_j|}{|S|\,|T|}\big(\mathbb{E}d_{G'}(S_i,T_j)-d\big)\right|\leq\sum_{i,j}\frac{|S_i|\,|T_j|}{|S|\,|T|}|\mathbb{E}d_{G'}(S_i,T_j)-d|\\&\leq\sum_{i,j}\frac{|S_i|\,|T_j|}{|S|\,|T|}|d_G(S_i,T_j)-d_{i,j}|\leq\epsilon\ ,\end{aligned}$$

where the second inequality follows from (12) with $X = S_i$ and $Y = T_j$, and the last inequality follows from the lemma's assumption that $\mathcal{A} \cup \mathcal{B}$ is a weak ϵ-regular partition of G, that is, (5).

We deduce from (7), (9), and (10) that there exists a graph, which we also denote by G' with a slight abuse of notation, that satisfies:

- $|G' \triangle G| \le \Delta + \epsilon^3 |A||B|$,
- $|d_{G'} - \mathbb{E} d_{G'}| \le \epsilon/4$,
- $|d_{G'}(S,T) - \mathbb{E} d_{G'}(S,T)| \le \epsilon/4$ for every S, T as above.

Note that G' is $3\epsilon/2$-regular, since for every S, T as above we have

$$\begin{aligned}\big|d_{G'}(S,T) - d_{G'}\big| &\le |d_{G'}(S,T) - \mathbb{E} d_{G'}(S,T)| + \big|\mathbb{E} d_{G'}(S,T) - d\big| + |d - d_{G'}| \\ &\le \epsilon/4 + \epsilon + \epsilon/4 = 3\epsilon/2 \ ,\end{aligned} \tag{13}$$

where to bound the first summand we used the third property of G', to bound the second summand we used (11), and to bound the third summand we used the second property of G' together with (6).

Finally, let $\widetilde{G}$ be obtained from G' by undoing some of the edge modifications, arbitrarily chosen, so that $|\widetilde{G} \triangle G| \le \Delta$. It remains to show that $\widetilde{G}$ is 2ϵ-regular. Indeed, for every S, T as above,

$$\begin{aligned}|d_{\widetilde{G}}(S,T) - d_{\widetilde{G}}| &\le |d_{\widetilde{G}}(S,T) - d_{G'}(S,T)| + |d_{G'}(S,T) - d_{G'}| + |d_{G'} - d_{\widetilde{G}}| \\ &\le 2\frac{\epsilon^3|A||B|}{|S||T|} + 3\epsilon/2 \le 2\epsilon,\end{aligned}$$

where to bound the first and third summands we used the first property of G' and to bound the second summand we used (13). ☐

3. Upper bound for SRAL

In this section we prove Theorem 2. The proof combines the perturbation lemma from Section 2 with iterated weak regularity using an entropy potential function which we introduce here.

3.1. **Entropy defect.** Let the function $H : \mathbb{R}^+ \to \mathbb{R}$ be given by

$$H(x) = x \ln x \ ,$$

where henceforth $0 \ln 0 = 0$. Note that H is a convex function. We will use H to define a potential function for vertex partitions. Crucially, we will need a "uniform" version of a defect inequality for H, which quantifies how convex H is in the following sense. The precise statement is the following.

Lemma 3.1. *Let $d_1, \dots, d_N, p_1, \dots, p_N \ge 0$ satisfy $\sum_{i=1}^N p_i = 1$ and $d := \sum_{i=1}^N p_i d_i \neq 0$. Then*

$$\sum_{i=1}^N p_i H(d_i) - H(d) \ge \frac{1}{2} d \bigg(\sum_{i=1}^N p_i \Big| \frac{d_i}{d} - 1 \Big| \bigg)^2 .$$

For the proof of Lemma 3.1 we will use Pinsker's inequality from information theory ([16], see also Lemma 6.2 in [11]), which lower bounds the Kullback–Leibler divergence of one probability distribution from another in terms of the total variation distance between the two distributions.

Theorem 7 (Pinsker's inequality)**.** *Let* $P = (p_1, \dots, p_N)$*,* $Q = (q_1, \dots, q_N)$ *satisfy* $p_i > 0, q_i \geq 0$ *and* $\sum_{i=1}^N p_i = 1$*,* $\sum_{i=1}^N q_i = 1$*. Then* $D_{KL}(Q \| P) \geq 2\delta(Q,P)^2$*, that is,*

$$\sum_{i=1}^N q_i \ln(q_i/p_i) \geq \frac{1}{2}\Big(\sum_{i=1}^N |q_i - p_i|\Big)^2 .$$

Proof of Lemma 3.1*.* Write $q_i = p_i d_i/d$ and note that

$$q_i \geq 0 \quad \text{and} \quad \sum_{i=1}^N q_i = \sum_{i=1}^N p_i d_i/d = 1 . \tag{14}$$

Assume without loss of generality that $p_i \neq 0$ for every i. By the definition of H we have

$$\begin{aligned}\sum_{i=1}^N p_i H(d_i) = \sum_{i=1}^N p_i d_i \ln(q_i d/p_i) &= \sum_{i=1}^N p_i d_i \ln(q_i/p_i) + \sum_{i=1}^N p_i d_i \ln d \\ &= d\sum_{i=1}^N q_i \ln(q_i/p_i) + H(d) .\end{aligned}$$

Since $(q_1, \dots, q_N)$ is a probability distribution by (14), we may apply Theorem 7 and deduce

$$\sum_{i=1}^N p_i H(d_i) - H(d) = d\sum_{i=1}^N q_i \ln(q_i/p_i) \geq d\cdot\frac{1}{2}\Big(\sum_{i=1}^N |q_i - p_i|\Big)^2 = d\cdot\frac{1}{2}\Big(\sum_{i=1}^N p_i|d_i/d - 1|\Big)^2,$$

as needed. ☐

3.2. **Potential function.** For the rest of this subsection let G be an n-vertex graph. We define the "potential" of a partition $\mathcal{P}$ of $V(G)$ by

$$\mathcal{H}(\mathcal{P}) = \sum_{V,V' \in \mathcal{P}} \frac{|V||V'|}{n^2} H(d(V,V')) , \tag{15}$$

where we recall that $H(x) = x \ln x$. Note that the summation in (15) is regarding ordered pairs (V, V'). It will be convenient to generalize the above definition. Henceforth, let $\mathcal{P}$ be a partition of $A \subseteq V(G)$, and let $\mathcal{P}'$ be a partition of $A' \subseteq V(G)$. We more generally define

$$\mathcal{H}(\mathcal{P}, \mathcal{P}') = \sum_{\substack{V \in \mathcal{P} \\ V' \in \mathcal{P}'}} \frac{|V||V'|}{|A||A'|} H(d(V,V')) ,$$

and in particular $\mathcal{H}(\mathcal{P}) = \mathcal{H}(\mathcal{P}, \mathcal{P})$ if $\mathcal{P}$ is a partition of $V(G)$.

Lemma 3.1 immediately implies the following bound on $\mathcal{H}(\mathcal{P}, \mathcal{P}') - \mathcal{H}(\{A\}, \{A'\})$, where we recall that $\mathcal{P}$ is a partition of A and $\mathcal{P}'$ is a partition of A'.

Corollary 8. *If* $d(A, A') \neq 0$*,*

$$\mathcal{H}(\mathcal{P}, \mathcal{P}') - \mathcal{H}(\{A\}, \{A'\}) \geq \frac{1}{2} d(A, A')\Big(\sum_{\substack{V \in \mathcal{P} \\ V' \in \mathcal{P}'}} \frac{|V||V'|}{|A||A'|}\left|\frac{d(V,V')}{d(A,A')} - 1\right|\Big)^2 .$$

Proof. The proof follows from Lemma 3.1 by setting $p_{(V,V')} = |V||V'|/|A||A'|$ and $d_{(V,V')} = d(V,V')$ for each $(V,V') \in \mathcal{P} \times \mathcal{P}'$, using the fact that

$$\sum_{\substack{V\in\mathcal{P}\\ V'\in\mathcal{P}'}} p_{(V,V')} d_{(V,V')} = \sum_{\substack{V\in\mathcal{P}\\ V'\in\mathcal{P}'}} e(V,V')/|A||A'| = d(A,A')\ .$$

□

Throughout the rest of the paper we will use the following notation: if $\mathcal{Q}$ is a refinement of $\mathcal{P}$ and $V \in \mathcal{P}$, then $\mathcal{Q}|_V$ will denote the partition of V that $\mathcal{Q}$ induces. We have the following properties.

Claim 3.2. If $\mathcal{Q}$ refines $\mathcal{P}$ and $\mathcal{Q}'$ refines $\mathcal{P}'$, then:

(i) $\mathcal{H}(\mathcal{Q},\mathcal{Q}') = \sum_{V\in\mathcal{P},V'\in\mathcal{P}'} \frac{|V||V'|}{|A||A'|}\mathcal{H}(\mathcal{Q}|_V, \mathcal{Q}'|_{V'})$.

(ii) $\mathcal{H}(\mathcal{Q},\mathcal{Q}') \geq \mathcal{H}(\mathcal{P},\mathcal{P}')$.

Proof. For the first item, we have

$$\begin{aligned}\mathcal{H}(\mathcal{Q},\mathcal{Q}') &= \sum_{\substack{V\in\mathcal{P}\\ V'\in\mathcal{P}'}} \sum_{\substack{U\in\mathcal{Q}|_V\\ U'\in\mathcal{Q}'|_{V'}}} \frac{|U|\,|U'|}{|A||A'|} H(d(U,U')) \\ &= \sum_{\substack{V\in\mathcal{P}\\ V'\in\mathcal{P}'}} \frac{|V||V'|}{|A||A'|} \sum_{\substack{U\in\mathcal{Q}|_V\\ U'\in\mathcal{Q}'|_{V'}}} \frac{|U|\,|U'|}{|V||V'|} H(d(U,U')) \\ &= \sum_{V\in\mathcal{P},V'\in\mathcal{P}'} \frac{|V||V'|}{|A||A'|}\mathcal{H}(\mathcal{Q}|_V, \mathcal{Q}'|_{V'})\ .\end{aligned}$$

As for the second item, if follows from the first item that

$$\begin{aligned}\mathcal{H}(\mathcal{Q},\mathcal{Q}') - \mathcal{H}(\mathcal{P},\mathcal{P}') &= \sum_{\substack{V\in\mathcal{P}\\ V'\in\mathcal{P}'}} \frac{|V||V'|}{|A||A'|}\Big(\mathcal{H}(\mathcal{Q}|_V, \mathcal{Q}'|_{V'}) - \mathcal{H}(\{V\},\{V'\})\Big) \\ &= \sum_{\substack{V\in\mathcal{P}\\ V'\in\mathcal{P}'}} \frac{|V||V'|}{|A||A'|}\Bigg(\sum_{\substack{U\in\mathcal{Q}|_V\\ U'\in\mathcal{Q}'|_{V'}}} \frac{|U|\,|U'|}{|V||V'|} H(d(U,U')) - H(d(V,V'))\Bigg) \\ &\geq 0\ ,\end{aligned}$$

where the inequality is due to the fact that each inner sum is nonnegative by Corollary 8 (in fact, Jensen's inequality suffices). □

The following claim gives lower and upper bounds for the potential function, where we recall that $\mathcal{P}$ is a partition of $V(G)$.

Claim 3.3. $d_G \ln(d_G) \leq \mathcal{H}(\mathcal{P}) \leq 0$.

Proof. The upper bound follows immediately from the fact that $H(x) \leq 0$ for every $0 \leq x \leq 1$. The lower bound $\mathcal{H}(\mathcal{P}) \geq H(d_G)$ follows from Jensen's inequality; indeed,

$$\begin{aligned}\mathcal{H}(\mathcal{P}) &= \sum_{V,V'\in\mathcal{P}} \frac{|V||V'|}{n^2} H(d(V,V')) \geq H\Big(\sum_{V,V'\in\mathcal{P}} \frac{|V||V'|}{n^2} d(V,V')\Big) \\ &= H\Big(\frac{2|E(G)|}{n^2}\Big) = H(d_G)\ .\end{aligned}$$

□

If $\mathcal{Q}$ refines $\mathcal{P}$ we write

$$\ell_1(\mathcal{Q},\mathcal{P}) = \frac{1}{2}\sum_{i,j=1}^{k} \sum_{\substack{U\in\mathcal{Q}|_{V_i}\\ U'\in\mathcal{Q}|_{V_j}}} |U||U'||d(U,U') - d(V_i,V_j)| \,.$$

We deduce the following relation between the ℓ_1-distance and the "entropy-distance" of partitions.

Lemma 3.4. *Suppose G has density p and $\mathcal{Q} \preceq \mathcal{P}$. If $\ell_1(\mathcal{Q},\mathcal{P}) \geq xpn^2$, then $\mathcal{H}(\mathcal{Q}) - \mathcal{H}(\mathcal{P}) \geq 2x^2p$.*

Proof. We have

$$\begin{aligned}
\frac{\mathcal{H}(\mathcal{Q}) - \mathcal{H}(\mathcal{P})}{p} &= \sum_{i,j} \frac{|V_i|\,|V_j|}{pn^2}\Big(\mathcal{H}(\mathcal{Q}|_{V_i},\mathcal{Q}|_{V_j}) - \mathcal{H}(\{V_i\},\{V_j\})\Big)\\
&\geq \frac{1}{2}\sum_{i,j}\frac{|V_i|\,|V_j|}{pn^2}\cdot d(V_i,V_j)\Bigg(\sum_{\substack{U\in\mathcal{Q}|_{V_i}\\ U'\in\mathcal{Q}|_{V_j}}} \frac{|U|\,|U'|}{|V_i|\,|V_j|}\left|\frac{d(U,U')}{d(V_i,V_j)} - 1\right|\Bigg)^2\\
&\geq \frac{1}{2}\Bigg(\sum_{i,j}\frac{|V_i|\,|V_j|}{pn^2}\cdot d(V_i,V_j)\sum_{\substack{U\in\mathcal{Q}|_{V_i}\\ U'\in\mathcal{Q}|_{V_j}}} \frac{|U|\,|U'|}{|V_i|\,|V_j|}\left|\frac{d(U,U')}{d(V_i,V_j)} - 1\right|\Bigg)^2\\
&= \frac{1}{2}\Bigg(\sum_{i,j}\frac{1}{pn^2}\sum_{\substack{U\in\mathcal{Q}|_{V_i}\\ U'\in\mathcal{Q}|_{V_j}}} |U||U'||d(U,U') - d(V_i,V_j)|\Bigg)^2\\
&= 2\bigg(\frac{\ell_1(\mathcal{Q},\mathcal{P})}{pn^2}\bigg)^2 ,
\end{aligned}$$

where all summations are over the ordered pairs (i,j) satisfying $d(V_i,V_j) > 0$. In the first line we used the first item of Claim 3.2. In the second we used Corollary 8, and in the third we used Jensen's inequality together with the fact that $\sum_{i,j}|V_i||V_j|d(V_i,V_j)/p|V|^2 = 2|E|/p|V|^2 = 1$. This completes the proof. ☐

3.3. **The iterative argument.** Here we show how to find a vertex partition $\mathcal{P}$ that has a refinement which is both weak ϵ-regular with ϵ that decreases with $|\mathcal{P}|$ and, simultaneously, close to $\mathcal{P}$ in terms of the entropy potential. The proof follows by iteratively finding better and better weak regular partitions (in the sense of Definition 2.1), similarly to the argument that Tao [24] used in order to provide an alternative proof for Szemerédi's regularity lemma (this type of argument has its roots in [1]). An analogous lemma, using the notion of Frieze–Kannan regularity, also appeared in [3].

Lemma 3.5. *Let $\alpha > 0$, let $s \in \mathbb{N}$, and let $g : \mathbb{N} \to (0,1)$ be a decreasing function. For every graph of density p and vertex equipartition $\mathcal{P}_0$ of order s, there are equipartitions $\mathcal{Q} \preceq \mathcal{P}$ refining $\mathcal{P}_0$ that satisfy:*

- *$\mathcal{Q}$ is weak $g(|\mathcal{P}|)$-regular.*
- *$\mathcal{H}(\mathcal{Q}) - \mathcal{H}(\mathcal{P}) < \alpha \cdot p\ln(1/p)$.*
- *$|\mathcal{P}| \leq E^{(\lfloor 1/\alpha\rfloor)}(s)$, where $E(x) = x\cdot 2^{\mathrm{poly}(1/g(x))}$.*

Proof. We construct $r+1$ equitable refinements $\mathcal{P}_0 \succeq \mathcal{P}_1 \succeq \cdots \succeq \mathcal{P}_r \succeq \mathcal{P}_{r+1}$ by letting $\mathcal{P}_i$ ($i \geq 1$) be the weak ϵ-regular refinement of $\mathcal{P}_{i-1}$ obtained by applying the weak regularity lemma in Theorem 6 with $\epsilon = g(|\mathcal{P}_{i-1}|)$. We stop once the potential difference between $\mathcal{P}_i$ and $\mathcal{P}_{i-1}$ drops below $\alpha p \ln(1/p)$. That is, r is chosen so that

$$\forall i \leq r: \ \mathcal{H}(\mathcal{P}_i) - \mathcal{H}(\mathcal{P}_{i-1}) \geq \alpha p \ln(1/p) \quad \text{and} \quad \mathcal{H}(\mathcal{P}_{r+1}) - \mathcal{H}(\mathcal{P}_r) < \alpha p \ln(1/p)\ .$$

We will show that the equipartitions $\mathcal{P} := \mathcal{P}_r$ and $\mathcal{Q} := \mathcal{P}_{r+1}$ satisfy the requirements in the statement. Note that, by construction, $\mathcal{Q}$ is weak $g(|\mathcal{P}|)$-regular and $\mathcal{H}(\mathcal{Q}) - \mathcal{H}(\mathcal{P}) < \alpha p \ln(1/p)$. Thus, it remains to bound $|\mathcal{P}| = |\mathcal{P}_r|$.

First, we claim that $r \leq 1/\alpha$. This follows by bounding the difference $\mathcal{H}(\mathcal{P}_r) - \mathcal{H}(\mathcal{P}_0)$; indeed,

$$r\alpha p \ln(1/p) \leq \sum_{i=1}^{r} \Big(\mathcal{H}(\mathcal{P}_i) - \mathcal{H}(\mathcal{P}_{i-1})\Big) = \mathcal{H}(\mathcal{P}_r) - \mathcal{H}(\mathcal{P}_0) \leq p \ln(1/p)\ ,$$

where the lower bound follows by construction and the upper bound follows from Claim 3.3. Next, recall that by Theorem 6 we have $|\mathcal{P}_i| \leq E(|\mathcal{P}_{i-1}|)$ with $E(x) = x \cdot 2^{\mathrm{poly}(1/g(x))}$. We claim that $|\mathcal{P}_i| \leq E^{(i)}(s)$ for every $0 \leq i \leq r$, which we prove by induction on i. For the base case $i=0$ we trivially have $|\mathcal{P}_0| = E^{(0)}(s)$, and for the induction step we have

$$|\mathcal{P}_{i+1}| \leq E(|\mathcal{P}_i|) \leq E(E^{(i)}(s)) = E^{(i+1)}(s)\ ,$$

where the last inequality follows from the induction hypothesis and the fact that E is an increasing function, which proves our claim. As $r \leq \lfloor 1/\alpha \rfloor$ we conclude that $|\mathcal{P}_r| \leq E^{(r)}(s) \leq E^{(\lfloor 1/\alpha \rfloor)}(s)$, where the last inequality follows from fact that E is increasing and $E(x) \geq x$. This completes the proof. ☐

3.4. **Proof of Theorem 2.** Here we combine the results from this and the previous section in order to prove Theorem 2. In fact, we will prove the following stronger result. We say that a vertex partition of order k is *f-regular*, where $f : \mathbb{N} \to (0,1)$, if *all* distinct pairs are $f(k)$-regular.

Theorem 9. *Let $\delta > 0$, let $s \in \mathbb{N}$, and let $f : \mathbb{N} \to (0,1)$ be a decreasing function. For any graph G of density p and any initial vertex equipartition $\mathcal{P}_0$ of order s, one can add/remove at most $\delta|E(G)|$ edges to obtain a graph that has an f-regular equipartition refining $\mathcal{P}_0$ of order at most $F^{(h)}(s)$, where $F(x) = 2^{x/f(x)}$ and $h = O(\log \frac{1}{p}/\delta^2)$.*

Theorem 2 indeed follows from Theorem 9 by taking $s = 1/\epsilon$ and $f(x) = \epsilon$, in which case the resulting equipartition is ϵ-regular (as the fraction of irregular pairs is at most $1/s = \epsilon$) and has order at most $\mathrm{twr}_{1/\epsilon}(O(\log \frac{1}{p}/\delta^2))$, as required. Before proving Theorem 9 we first isolate a simple observation. We will slightly simplify the proofs by assuming, as we may, that in an equipartition all parts are of exactly the same size.

Claim 3.6. For a graph $G = (V, E)$, let $\mathcal{P} = \{V_1, \ldots, V_k\}$ be an equipartition of V and let $\mathcal{Q}$ be a weak ϵ-regular partition that refines $\mathcal{P}$. For every induced bipartite graph $G[V_a, V_b]$ with $a \neq b$, the partition $\mathcal{Q}|_{V_a} \cup \mathcal{Q}|_{V_b}$ is weak ϵk-regular (in the sense of Lemma 2.2).

Proof. Let $S \subseteq V_a$, $T \subseteq V_b$ with $|S| \geq \epsilon k\,|V_a|$, and let $|T| \geq \epsilon k\,|V_b|$. As $\mathcal{P}$ is an equipartition, this means that $|S|\,,|T| \geq \epsilon\,|V|$. Thus, by the weak ϵ-regularity of $\mathcal{Q}$ (recall Definition 2.1),

$$\sum_{i,j=1}^{k} \sum_{\substack{U\in\mathcal{Q}|_{V_i}\\ U'\in\mathcal{Q}|_{V_j}}} \frac{|S\cap U|\,|T\cap U'|}{|S|\,|T|}\,|d(S\cap U, T\cap U') - d(U,U')| \leq \epsilon\ .$$

Since $|S\cap V_i| = 0$ for any $i \neq a$ and $|T\cap V_j| = 0$ for any $j \neq b$, the above reduces to

$$\sum_{\substack{U\in\mathcal{Q}|_{V_a}\\ U'\in\mathcal{Q}|_{V_b}}} \frac{|S\cap U|\,|T\cap U'|}{|S|\,|T|}\,|d(S\cap U, T\cap U') - d(U,U')| \leq \epsilon \leq \epsilon k\ ,$$

which is what we needed to prove. ☐

Proof of Theorem 9. Let $G = (V,E)$ be a graph of density p, and let $\mathcal{P}_0$ be the given equipartition of order s. We apply Lemma 3.5 on G and $\mathcal{P}_0$ with parameters

$$\alpha = \delta^2/2\ln(1/p), \quad g(x) = f(x)/2x\ ,$$

where we note, since f is decreasing, that g is decreasing as well, as required by Lemma 3.5. Let $\mathcal{Q} \preceq \mathcal{P} \preceq \mathcal{P}_0$ be the obtained equipartitions, and put $k = |\mathcal{P}|$. This means that:

(i) $\mathcal{Q}$ is weak $g(k)$-regular,
(ii) $\mathcal{H}(\mathcal{Q}) - \mathcal{H}(\mathcal{P}) \leq p\delta^2/2$, and
(iii) $|\mathcal{P}| \leq E^{(\lfloor 2\ln(1/p)/\delta^2\rfloor)}(s)$, where $E(x) = 2^{\mathrm{poly}(x/f(x))}$.

We will show that one can modify at most $\delta\,|E|$ edges of G so as to make $\mathcal{P}$ an f-regular partition, provided $|V|$ is sufficiently large, which would complete the proof by item (iii).

Write $\mathcal{P} = \{V_1,\ldots,V_k\}$. For each $i<j$ we apply Lemma 2.2 on the induced bipartite subgraph $G[V_i,V_j]$ with the partition $\mathcal{Q}|_{V_i}\cup\mathcal{Q}|_{V_j}$ and $\epsilon = f(k)/2$, which we claim we may. To see this, note that, since $\mathcal{Q}$ is weak $f(k)/2k$-regular by item (i), the partition $\mathcal{Q}|_{V_i}\cup\mathcal{Q}|_{V_j}$ of $G[V_i,V_j]$ is indeed weak ϵ-regular by Claim 3.6; moreover, we may assume that $|V_i|\,,|V_j| \geq 8/f^4(k)$ as required by Lemma 2.2, since otherwise $|V| \leq O(k/f^4(k))$ is at most the bound in the statement of Theorem 9, so we may instead take the partition of V into parts of size one. Thus, we transform each bipartite subgraph as above into an $f(k)$-regular graph by modifying some of its edges. It therefore follows that for the modified graph, the partition $\mathcal{P}$ is f-regular. By Lemma 2.2, the total number of edge modifications thus made is at most

$$\ell_1(\mathcal{Q},\mathcal{P}) = \frac{1}{2}\sum_{i,j=1}^{k} \sum_{\substack{U\in\mathcal{Q}|_{V_i}\\ U'\in\mathcal{Q}|_{V_j}}} |U||U'||d(U,U') - d(V_i,V_j)|\ .$$

Since $\mathcal{H}(\mathcal{Q}) - \mathcal{H}(\mathcal{P}) \leq 2p(\delta/2)^2$, it follows from Lemma 3.4 that $\ell_1(\mathcal{Q},\mathcal{P}) \leq (\delta/2)p\,|V|^2 = \delta\,|E|$. This completes the proof. ☐

4. THE REMOVAL LEMMA VIA SRAL

In this section we prove Theorem 1. For the proof we will need a counting lemma that corresponds to SRAL. Call G an *(ϵ, p)-graph* if G is multipartite, and the bipartite graph between any pair of classes is either empty or ϵ-regular of density at least p. The following "approximate counting lemma" shows that the usual counting lemma approximately holds for any graph that is sufficiently dense in an (ϵ, p)-graph.

Lemma 4.1. *Let H be an h-vertex graph with m edges, let G' be a k-partite (ϵ, p)-graph on $(U_1, \dots, U_k)$ with $|U_i| = n$, and let G be a graph on $V(G')$ such that for every $i \neq j$ with $d_{G'}(U_i, U_j) > 0$, $G[U_i, U_j]$ is δ-close to $G'[U_i, U_j]$. Suppose that $\delta \le 1/2m$, $\epsilon \le (p^{m+1}/32h^4)^2$ and $n \ge 4h^{h+3}/p^m$. If G' contains a copy of H, then the number of copies of H in G is at least*

$$\frac{1 - \delta m}{2} \cdot p^m n^h \ .$$

The proof of Lemma 4.1, which is a souped-up version of the standard proof of the graph counting lemma (see, e.g., the survey [12]), appears in Subsection 4.1. Let us now show how to prove the graph removal lemma by relying on the sparse regular approximation lemma (Theorem 2) and the above Lemma 4.1.

Proof of Theorem 1. Let H be a graph with $h \ge 3$ vertices and $m \ge 1$ edges. Put

$$\epsilon' = \left(\frac{\epsilon}{h}\right)^{h^2}, \quad \delta = \left(\frac{1}{4m}\right)^2, \quad d = \left(\frac{1}{\epsilon'}\right)^2 .$$

We will prove the bound

$$\mathrm{Rem}_H(\epsilon) \le [d \cdot S(\epsilon', \delta, \epsilon)]^h \ . \tag{16}$$

It is well known [21] that $\mathrm{Rem}_H(\epsilon) \ge (1/\epsilon)^{c \log(1/\epsilon)}$ for some $c = c(H)$. Assuming $\epsilon \le \epsilon_0(h)$ is small enough, this means $\mathrm{Rem}_H(\epsilon) \ge d^{2h}$. Thus, proving (16) would imply $d \le S(\epsilon', \delta, \epsilon)$, and therefore $\mathrm{Rem}_H(\epsilon) \le [S(\epsilon', \delta, \epsilon)]^{2h}$, which proves (1).

Let G be an n-vertex graph that is ϵ-far from being H-free. Observe that G contains at least $\Delta := \epsilon n^2/m$ edge-disjoint copies of H. Let G_1 be a subgraph of G on $V(G)$ that only consists of Δ such copies of H. Note that the density p of G_1 is given by $\frac{1}{2}pn^2 = m\Delta$, or equivalently, $p = 2\epsilon$. Using Definition 1.1, there is a graph G_2 on $V(G)$ with $|E(G_1) \triangle E(G_2)| \le \delta |E(G_1)|$ that has an ϵ'-regular equipartition $\mathcal{P}$ with $1/\epsilon' \le |\mathcal{P}| \le S(\epsilon', \delta, p)$. Since G_1 is a subgraph of G, in order to prove (16) it suffices to show that G_1 contains at least $n^h/(d|\mathcal{P}|)^h$ copies of H.

We next construct a subgraph G_3 of G_2 which would facilitate the embedding of H. We obtain G_3 by removing all edges between each pair $V, V' \in \mathcal{P}$ that satisfies either one of the following:

(i) $V = V'$,
(ii) $G_2[V, V']$ is not ϵ'-regular,
(iii) $d_{G_2}[V, V'] < \sqrt{\delta}\epsilon$,
(iv) $G_1[V, V']$ is not $\sqrt{\delta}$-close to $G_2[V, V']$, that is, $|E_{G_1}(V, V') \triangle E_{G_2}(V, V')| > \sqrt{\delta} \cdot |E_{G_2}(V, V')|$.

The number of edges removed from G_2 to obtain G_3 is smaller than

$$\begin{aligned}&\sum_{V\in\mathcal{P}}\frac{1}{2}|V|^2+\sum_{\substack{V,V'\in\mathcal{P}\colon (V,V')\\ \text{not } \epsilon'\text{-regular}}}|V||V'|+\sum_{V,V'\in\mathcal{P}}\sqrt{\delta}\epsilon|V||V'|\\ &\qquad+\sum_{V,V'\in\mathcal{P}}\frac{1}{\sqrt{\delta}}\big|E_{G_2}(V,V')\mathbin{\triangle}E_{G_1}(V,V')\big|\\ &\le\frac{n^2}{2|\mathcal{P}|}+\epsilon'n^2+\sqrt{\delta}\epsilon n^2+\frac{1}{\sqrt{\delta}}|E(G_2)\triangle E(G_1)|\le 2\epsilon'n^2+\sqrt{\delta}\epsilon n^2+\sqrt{\delta}\epsilon n^2\le 3\sqrt{\delta}\epsilon n^2.\end{aligned}$$

Thus, one obtains G_3 from G_1 by modifying fewer than $(3\sqrt{\delta}+\delta)\epsilon n^2\le 4\sqrt{\delta}\epsilon n^2 = \epsilon n^2/m = \Delta$ edges. Recalling that G_1 contains Δ edge-disjoint copies of H, we deduce that G_3 contains a copy of H.

Put $k=|\mathcal{P}|$. From items (i), (ii), and (iii) above it follows that G_3 contains an h-partite (ϵ',q)-graph with $q=\sqrt{\delta}\epsilon=\epsilon/4m$, having n/k vertices in each vertex class,[10] containing a copy of H. From item (iv) it follows that for every pair of its vertex classes $U,U'\in\mathcal{P}$ with $d_{G_3}(U,U')>0$, $G_1[U,U']$ is $\sqrt{\delta}$-close to $G_3[U,U']$. Recall that our claim is that G_1 (and hence G) contains at least $n^h/(dk)^h$ copies of H. Assume $n\ge dk$, as otherwise we are done since already $\Delta\ge 1$. We apply the approximate counting lemma (Lemma 4.1), noting that, as required, $\epsilon'\ge(q^{m+1}/32h^4)^2$ and $n/k\ge d=(h/\epsilon)^{2h^2}\ge 4h^{h+3}/q^m$. We deduce that the number of copies of H in G_1 is at least

$$\frac{1-\sqrt{\delta}m}{2}q^m\cdot\left(\frac{n}{k}\right)^h=\frac{3}{8}q^m\cdot\left(\frac{n}{k}\right)^h\ge\frac{1}{d}\cdot\left(\frac{n}{k}\right)^h\ge\frac{n^h}{(dk)^h}\,,$$

proving our claim and thus completing the proof. □

4.1. **Approximate counting lemma.** In order to prove Lemma 4.1 we will need the following well-known properties of ϵ-regular graphs. For completeness, we prove these properties in the appendix. Throughout, we say that (A,B) is an *(ϵ,d)-regular pair* if the bipartite graph between the vertex subsets A,B is ϵ-regular of density d. Furthermore, we use the notation $x\pm\epsilon$ for a number lying in the interval $[x-\epsilon,\ x+\epsilon]$.

Fact 4.2. If (A,B) is an (ϵ,d)-regular pair, all vertices of B but at most $2\epsilon|B|$ have degree $(d\pm\epsilon)|A|$.

Fact 4.3. Let $\alpha\ge\epsilon>0$. Let (A,B) be an $(\epsilon$,d$)$-regular pair. If $A'\subseteq A$, $B'\subseteq B$ are of size $|A|\ge\alpha|A|$, $|B|\ge\alpha|B|$, then the pair (A',B') is $(2\epsilon/\alpha,\ d\pm\epsilon)$-regular.

Fact 4.4. Let the pairs $(A,C),(B,C)$ be (ϵ,d)-regular and (ϵ,d')-regular, respectively. Write $\operatorname{codeg}(a,b)$ for the number of common neighbors of a,b in C, and put $\epsilon'=6\epsilon/d$. All pairs $(a,b)\in A\times B$ but at most $\epsilon'|A||B|$ satisfy $\operatorname{codeg}(a,b)=(dd'\pm\epsilon')|C|$.

Lemma 4.5 (Counting lemma, Lemma 1.6 in [4]). *Let let H be be a graph on $[h]$, and let G be an h-partite graph on $(V_1,\dots,V_h)$. If all pairs (V_i,V_j) are ϵ-regular, then the number of induced copies of H in G where vertex $i\in[h]$ is embedded in V_i is*

$$\prod_{i=1}^{h}|V_i|\bigg(\prod_{1\le i<j\le h}p'_{i,j}\ \pm\ \sqrt{h^3\epsilon}\bigg)\,,$$

---

[10] We assume, as we may, that all parts of the equipartition $\mathcal{P}$ have exactly the same size.

where $p'_{i,j} = d_G(V_i, V_j)$ *if* $(i,j) \in E(H)$ *and* $p'_{i,j} = 1 - d_G(V_i, V_j)$ *otherwise.*

We are now ready to prove the approximate counting lemma.

Proof of Lemma 4.1. Assume G' contains a copy of H. Fix an embedding, and let $f : [h] \to [k]$ be such that vertex $i \in [h]$ of H is embedded in $U_{f(i)}$. We henceforth refer to a homomorphic copy of H where vertex $i \in [h]$ is embedded in $U_{f(i)}$ as an *f-copy.* Put $p_{i,j} = d_{G'}(U_{f(i)}, U_{f(j)})$. Let $(a,b) \in E(H)$. We will prove that for all edges e of G' between $U_{f(a)}$ and $U_{f(b)}$ but at most γn^2, the number of f-copies in G' containing e is

$$n^{h-2}\Big(\prod_{\substack{(i,j)\in E(H)\\ (i,j)\neq(a,b)}} p_{i,j} \pm \gamma\Big) , \tag{17}$$

where $\gamma = p^m/4h^2$. We will also show that the total number of f-copies in G' is

$$n^{h}\Big(\prod_{(i,j)\in E(H)} p_{i,j} \pm \gamma\Big) . \tag{18}$$

Applying (17) for each edge $e \in E(G') \setminus E(G)$ (and each $(a,b) \in E(H)$), we deduce using (18) that in G the total number of f-copies is at least

$$\begin{aligned} &n^{h}\Big(\prod_{(i,j)\in E(H)} p_{i,j} - \gamma\Big) - m\left(\delta p_{a,b} n^2 \cdot n^{h-2}\Big(\prod_{\substack{(i,j)\in E(H)\\ (i,j)\neq(a,b)}} p_{i,j} + \gamma\Big) + \gamma n^h\right) \\ &\geq n^h\left((1-\delta m) \prod_{(i,j)\in E(H)} p_{i,j} - (2m+1)\gamma\right) \geq n^h\Big((1-\delta m)p^m - (2m+1)\gamma\Big) , \end{aligned}$$

where in the last inequality we used the fact that $p_{i,j} \geq p$ for every $(i,j) \in E(H)$, which follows from the lemma's assumptions that G' contains an (f-)copy of H, meaning $p_{i,j} \neq 0$ for $(i,j) \in E(H)$, and that G' is an (ϵ, p)-graph. However, some of these homomorphic copies may not be proper copies. The number of mappings from $V(H)$ to $V(G)$ that are not injective is

$$(hn)^h - \prod_{i=0}^{h-1}(hn - i) \leq h^2(hn)^{h-1} = (h^{h+1}/n) \cdot n^h \leq \gamma n^h ,$$

where in the last inequality we used the assumption that $n \geq 4h^{h+3}/p^m = h^{h+1}/\gamma$. We deduce that, as desired, the number of (proper) copies of H in G is at least

$$n^h\big((1-\delta m)p^m - h^2\gamma\big) \geq n^h \cdot \frac{1}{2}(1-\delta m)p^m ,$$

where we used the fact that $h^2\gamma = p^m/4$ and $\delta \leq 1/2m$.

It remains to prove (17) and (18). Let F be the h-partite graph obtained from G' by replacing each vertex class U_t by $|f^{-1}(t)|$ copies $(V_i)_{i\in f^{-1}(t)}$, so that $F[V_i, V_j] = G'[U_{f(i)}, U_{f(j)}]$ if $(i,j) \in E(H)$ and $F[V_i, V_j]$ is empty otherwise (so in particular, each $F[V_i, V_j]$ is ϵ-regular). Observe that the number of f-copies in G' is equal to the number of (induced) copies of H in F where vertex $i \in [h]$ is embedded in V_i. Therefore, Lemma 4.5 implies (18). As for proving (17), let us fix $(a,b) \in E(H)$. For $(x,y) \in V_a \times V_b$ and $i \notin \{a,b\}$ let

$$V'_i = \{z \in V_i \,:\, (i,a) \in E(H) \Rightarrow (z,x) \in E(F),\ (i,b) \in E(H) \Rightarrow (z,y) \in E(F)\} .$$

Put $\epsilon' = 6\epsilon/p$, and put $p'_{i,j} = d(V_i, V_j)$ if $(i,j) \in E(H)$ and $p'_{i,j} = 1$ otherwise. It follows from Fact 4.4 that all pairs $(x,y) \in V_a \times V_b$ but at most $h\epsilon' n^2$ satisfy, for all $i \notin \{a,b\}$, that $|V'_i| \geq n(p'_{i,a}p'_{i,b} - \epsilon')$ $(\geq np^2/2)$. For a "good" such pair (x,y), let $F_{x,y}$ be the $(h-2)$-partite subgraph of F induced on the V'_i. By Fact 4.3, all pairs (V'_i, V'_j) are $4\epsilon/p^2$-regular of density $d(V_i, V_j) \pm \epsilon$. Letting H' be the induced subgraph of H obtained by removing a and b, it follows from Lemma 4.5 that the number of copies of H' in $F_{x,y}$, where vertex $i \in [h]$ is embedded in V'_i, is

$$\prod_{\substack{i\in[h]:\\ i\neq a,b}} n(p'_{i,a}p'_{i,b} \pm \epsilon')\Bigg(\prod_{\substack{(i,j)\in E(H)\\ i,j\notin\{a,b\}}} (p_{i,j} \pm \epsilon) \,\pm\, \sqrt{h^3 \cdot 4\epsilon/p^2}\Bigg)$$
$$= n^{h-2}\Bigg(\prod_{\substack{(i,j)\in E(H)\\ (i,j)\neq(a,b)}} p_{i,j} \pm \big(2m\epsilon' + 2h^2\sqrt{\epsilon}/p\big)\Bigg).^{11}$$

As the error above is at most $8h^2\sqrt{\epsilon}/p \leq p^m/4h^2 = \gamma$ we deduce (17), completing the proof. ☐

5. Lower bound for SRAL

5.1. **Proof overview.** In this section we prove Theorem 4. First, we give a short description of the density-p graph G witnessing our lower bound, followed by an overview of the proof of correctness.

5.1.1. *Construction.* We construct a bipartite graph G iteratively as follows. Starting with $G_0 = K_{k,k}$, where $k = \mathrm{poly}(1/p)$, we define G_{i+1} as follows: we take a blow-up of G_i, inflating each vertex into $2^{\Omega(|V(G_i)|)}$ vertices and replacing each edge by a complete bipartite graph. For each such complete bipartite graph on vertex sets (X,Y) we randomly bipartition $X = X_1 \cup X_2$ and $Y = Y_1 \cup Y_2$; we then remove all edges between X_1, Y_2 and between X_2, Y_1, therefore removing half of all edges. We repeat the above process $s = \log\frac{1}{p}$ times, thus obtaining a graph G_s of density p. As our final graph G we take any blow-up of G_s. Note that $V(G_i)$ naturally defines a partition $\mathcal{X}_i$ of $V(G)$. Clearly, $|\mathcal{X}_i| = \mathrm{twr}(\Omega(i))$, and it has the property that a 2^{-i}-fraction of its pairs have density $2^i p$ in G while the rest have density 0. We note that the idea behind the above construction is for it to be "hard" for the iterative upper bound proof of Theorem 2 based on Scott's method from [20] that we mentioned in Subsection 1.3.

5.1.2. *Proof of correctness.* As in most lower bound proofs for regularity lemmas pioneered by Gowers [8], we use the following notions (more precise definitions appear in Subsection 5.5). We say that S is γ*-contained* in T if all but a γ-fraction of S is contained in T (i.e., $|S\setminus T| \leq \gamma|S|$). We say that an equipartition $\mathcal{Z}$ γ*-refines* $\mathcal{X}$ if all but $\gamma|\mathcal{Z}|$ clusters $Z \in \mathcal{Z}$ are γ-contained in some $X \in \mathcal{X}$. For simplicity, one may think of γ as a fixed small constant for the rest of this subsection.

We now give an overview of the crux of the proof. Let G' be any graph obtained from G by adding/removing $\delta_0|E(G)|$ edges, where again one may think of δ_0 as a fixed small constant. Our proof shows that any ϵ-regular equipartition $\mathcal{Z}$ of G'

with $\epsilon = p^5$ must γ-refine $\mathcal{X}_s$, where we assume for simplicity that $\mathcal{Z}$ refines $\mathcal{X}_0$.[12] This implies that $|\mathcal{Z}| = \Omega(|\mathcal{X}_s|) \geq \mathrm{twr}(\Omega(s)) = \mathrm{twr}(\Omega(\log \frac{1}{p}))$, proving the lower bound on $S(\epsilon, \delta_0, p)$ that is stated in Theorem 4.

The proof proceeds by assuming towards contradiction that $\mathcal{Z}$ does not γ-refine $\mathcal{X}_s$, meaning there are at least $\gamma|\mathcal{Z}|$ parts $Z \in \mathcal{Z}$ that are not γ-contained in a member of $\mathcal{X}_s$. Letting Z be such a part, there must be $1 \leq r \leq s$ such that Z is γ-contained in $X \in \mathcal{X}_{r-1}$ yet is not γ-contained in any one member of $\mathcal{X}_r$. Using the randomness in the choice of the bipartitions $X = X_1 \cup X_2$ in the construction of G, it can be shown that an $\Omega(2^{-r})$-fraction of the bipartitions satisfy $\min\{|Z \setminus X_1|, |Z \setminus X_2|\} \geq \Omega(\gamma|Z|)$. Fix one such bipartition $X_1 \cup X_2$, and let $Y \in \mathcal{X}_{r-1}$ be the part that "induces" it (so Y is also bipartitioned, $Y = Y_1 \cup Y_2$). Assume without loss of generality that $|Z \cap X_1| \geq \frac{1}{2}|Z \cap X| \approx \frac{1}{2}|Z|$. The structure of G is then used to argue that the density between $Z \cap X_1$ and Y_1 is $2^r p$, while the density between $Z \setminus X_1$ and Y_1 is only a fraction of $2^r p$. Importantly, as explained above, both $Z \cap X_1$ and $Z \setminus X_1$ are linear-size subsets of Z.

Next, we use an important property of G which is that G is a (somewhat) quasirandom bipartite graph. Together with the assumption that Y is essentially partitioned into clusters $Z' \in \mathcal{Z}$ and the assumption that all pairs (Z, Z') are regular in G', one can deduce that, roughly speaking, "almost" all pairs of the form (Z, Y_1) must be regular in G' as well. Since (Z, Y_1) was already shown to be irregular in G in a strong sense—having linear-size witnesses and an $\Omega(2^r p)$ density discrepancy—it follows that $G'[Z, Y_1]$, the bipartite subgraph of G' induced by $Z \cup Y_1$, must have been obtained from $G[Z, Y_1]$ by adding or removing $\Omega(2^r p|Z||Y_1|)$ edges. Summing over all $\Omega(2^{-r}|\mathcal{X}_{r-1}|)$ parts $Y \in \mathcal{X}_{r-1}$ as above, we deduce that the number of modifications in the edges adjacent to Z is $\Omega(|Z|pn)$. Next, by summing over all $Z \in \mathcal{Z}$ as above, it follows that the total number of modifications is $\Omega(|E(G)|)$, a contradiction that completes the proof.

The overview above clearly hides quite a few assumptions, steps, and subtleties. As one example, the fact that $\mathcal{Z}$ is not exactly a refinement of $\mathcal{X}_{r-1}$ introduces an error term to our lower bound on the number of edge modifications between Z and Y_1. While this error term can certainly "kill off" the main term for some of the Y_1's, the quasirandomess of G implies that it has a negligible effect when summing over sufficiently many $Y \in \mathcal{X}_{r-1}$. Of course, we must also guarantee that our bipartite G is not *too* quasirandom, as otherwise it would have had a p^5-regular partition of order 2. The proof of Theorem 4 spans the rest of this section. Specifically, the next three subsections contain the construction of G and the proofs of its various properties, and the following two subsections contain the technical parts of the proof described here. The last subsection completes the proof of Theorem 4 by showing that the complete bipartite graph can be decomposed into copies of G.

5.2. **Preliminary lemmas.** We use the standard definitions and notation given in Section 2. In this section all graphs are bipartite. We note that our actual construction differs slightly from the one described in Subsection 5.1 in that the random bipartitions are replaced by a sequence of deterministic bipartitions having pseudorandom properties. This will allow us to more easily control the measure of quasirandomness as we go through the iterative construction.

---

[12] This is in fact how we proceed in the actual proof. Namely, the main technical part of the proof (Theorem 10) essentially makes this assumption, and we reduce to this case in the proof of Theorem 11 below.

5.2.1. *Pseudorandom bipartitions.* Let $\mathcal{B} = (X_{1,0}, X_{1,1}), \dots, (X_{d,0}, X_{d,1})$ be a sequence of equitable bipartitions of a set X with $|X|$ even. We say that $\mathcal{B}$ is α-*orthogonal* if for every $1 \le i < j \le d$ and $\ell, \ell' \in \{0,1\}$ we have

$$|X_{i,\ell} \cap X_{j,\ell'}| \le \Big(\frac{1}{4} + \alpha\Big)|X| \ . \tag{19}$$

We say that $\mathcal{B}$ is β-*balanced* if for every $x \ne y \in X$, the number of bipartitions $(X_{i,0}, X_{i,1})$ with $x, y \in X_{i,0}$ or $x, y \in X_{i,1}$ is at most $(\frac{1}{2} + \beta)d$. We say that $\mathcal{B}$ is an (n,d,α,β)-*sequence* if $|X| = n$, $|\mathcal{B}| = d$, and $\mathcal{B}$ is both α-orthogonal and β-balanced.

Lemma 5.1. *For every $d \ge 200$ and every even $n \le 2^{\lfloor d/200 \rfloor}$ there is an (n,d,α,β)-sequence with $\alpha = \sqrt{2\ln(d)/n}$, $\beta = 1/16$.*

Proof. Choose d equitable bipartitions $(X_{i,0}, X_{i,1})$ of X independently and uniformly at random, where $|X| = n$ is even. Fix a pair $1 \le i < j \le d$. The random variable $|X_{i,0} \cap X_{j,1}|$ follows a hypergeometric distribution. Thus, we may apply the Chernoff bound (see Section 6 in [10]), meaning the probability that $|X_{i,0} \cap X_{j,0}| = (\frac{1}{4} \pm \alpha)n$ (which is equivalent to all four inequalities in (19) with $\ell, \ell' \in \{0,1\}$) does not hold is at most $2\exp(-2\alpha^2 n)$. Next, fix a pair $x \ne y \in X$. The probability that a given $1 \le i \le d$ satisfies $x, y \in X_{i,0}$ or $x, y \in X_{i,1}$ is $2\binom{n-2}{n/2-2}/\binom{n}{n/2} \le 1/2$. Since these events are mutually independent we may apply the Chernoff bound, meaning the probability that there are more than $(\frac{1}{2} + \beta)d$ values of i for which the above holds is at most $\exp(-2\beta^2 d)$.

By the union bound, the probability that at least one of the two events above holds for some choice of $1 \le i < j \le d$ or $x \ne y \in X$ is at most

$$\binom{d}{2} \cdot 2\exp(-2\alpha^2 n) + \binom{n}{2} \cdot \exp(-2\beta^2 d) \ .$$

This probability is smaller than 1 when taking d, n, α, β as in the statement, completing the proof. $\square$

We will later need the following trivial fact.

Fact 5.2. Any subsequence of length d' of an (n,d,α,β)-sequence is an $(n,d',\alpha,1/2)$-sequence.

Our proof will critically rely on the following lemma from [14], which improved upon a similar lemma from [8].

Lemma 5.3. *If $(X_{1,0}, X_{1,1}), \dots, (X_{d,0}, X_{d,1})$ is a sequence of bipartitions of X that is $\frac{1}{16}$-balanced, then for every $\lambda = (\lambda_1, \dots, \lambda_{|X|})$ with $\lambda_t \ge 0$ and $\|\lambda\|_1 = 1$, at least $d/6$ of the bipartitions $(X_{i,0}, X_{i,1})$ satisfy $\min\{\sum_{t \in X_{i,0}} \lambda_t, \sum_{t \in X_{i,1}} \lambda_t\} \ge \frac{1}{8}(1 - \|\lambda\|_\infty)$.*

5.2.2. *Common refinement.* Henceforth, the common refinement of partitions $\mathcal{Z}, \mathcal{X}$ is denoted $\mathcal{Z} \cap \mathcal{X}$; that is,

$$\mathcal{Z} \cap \mathcal{X} = \{Z \cap X \ : \ Z \in \mathcal{Z}, \, X \in \mathcal{X}\} \ .$$

We will need the definition of a regular partition when the partition is not necessarily equitable. A vertex partition $\mathcal{Z}$ of an n-vertex graph is said to be ϵ-regular if

$$\sum_{\substack{(Z,Z')\in\mathcal{Z}^2\\ \text{not } \epsilon\text{-regular}}} |Z||Z'| \le \epsilon n^2 \,.$$

Claim 5.4. Let $\mathcal{Z}$ be an ϵ-regular partition of a graph G. For any partition $\mathcal{X}$ of order k, the common refinement $\mathcal{Z}\cap\mathcal{X}$ is a $\sqrt{8k\epsilon}$-regular partition of G.

Proof. Put $\alpha = \sqrt{\epsilon/2k}$ and $n = |V(G)|$. For each ϵ-regular pair $(Z,Z')\in\mathcal{Z}^2$, it follows from Fact B.2 that for every $X,X'\in\mathcal{X}$ with $|Z\cap X|\ge\alpha|Z|$ and $|Z'\cap X'|\ge\alpha|Z'|$, the pair $(Z\cap X,\, Z'\cap X')$ is ϵ'-regular with $\epsilon' = (2/\alpha)\epsilon = \sqrt{8k\epsilon}$. Call $Z\cap X\in\mathcal{Z}\cap\mathcal{X}$ *small* if $|Z\cap X|<\alpha|Z|$. We have

$$\sum_{\substack{A,A'\in\mathcal{Z}\cap\mathcal{X}:\\ A \text{ small}}} |A||A'| \le \sum_{\substack{A\in\mathcal{Z}\cap\mathcal{X}:\\ A \text{ small}}} |A|n = \sum_{\substack{Z\in\mathcal{Z},X\in\mathcal{X}:\\ |Z\cap X|<\alpha|Z|}} |Z\cap X|n \le \sum_{\substack{Z\in\mathcal{Z},X\in\mathcal{X}:\\ |Z\cap X|<\alpha|Z|}} \alpha|Z|n \le k\alpha n^2 \,.$$

Call $(Z\cap X, Z'\cap X')\in(\mathcal{Z}\cap\mathcal{X})^2$ *bad* if (Z,Z') is not ϵ-regular. Recall that if $(A,A')\in(\mathcal{Z}\cap\mathcal{X})^2$ is not bad and A,A' are both not small, then (A,A') is ϵ'-regular. Therefore,

$$\begin{aligned}\sum_{\substack{(A,A')\in(\mathcal{Z}\cap\mathcal{X})^2\\ \text{not } \epsilon'\text{-regular}}} |A||A'| &\le \sum_{\substack{A,A'\in\mathcal{Z}\cap\mathcal{X}:\\ (A,A') \text{ bad}}} |A||A'| + 2\sum_{\substack{A,A'\in\mathcal{Z}\cap\mathcal{X}:\\ A \text{ small}}} |A||A'|\\ &\le \sum_{\substack{(Z,Z')\in\mathcal{Z}^2\\ \text{not } \epsilon\text{-regular}}} |Z||Z'| + 2k\alpha n^2 \le (\epsilon+2k\alpha)n^2 \le \epsilon' n^2 \,.\end{aligned}$$

This proves that $\mathcal{Z}\cap\mathcal{X}$ is ϵ'-regular, as needed. ☐

5.2.3. *Quasirandom graphs.* A bipartite graph $G=(U,V;E)$ of density p is said to be (ϵ)-*regular* if all sets $A\subseteq U$, $B\subseteq V$ with $|A|\ge\epsilon|U|$, $|B|\ge\epsilon|V|$ satisfy

$$|d_G(A,B)-p| \le \epsilon p \,. \tag{20}$$

Definition 5.5 ((p,δ)-quasirandom graph)**.** A regular bipartite graph $G=(U,V;E)$ of density p is (p,δ)-*quasirandom* if all but $\delta|U|^2$ pairs $(u,u')\in U^2$ satisfy $\operatorname{codeg}(u,u') \le (1+\delta)p^2|V|$.

As is well known, "regular" codegrees imply quasirandomness. However, we need a somewhat different version with specific parameters, which we prove below for completeness.

Lemma 5.6. *Every $(p,\epsilon p)$-quasirandom graph is $(2\epsilon^{1/7})$-regular.*

Proof. Let the bipartite graph $G=(U,V;E)$ be $(p,\epsilon p)$-quasirandom. We will prove that all sets $A\subseteq U$, $B\subseteq V$ of size $|A|=\alpha|U|$, $|B|=\beta|V|$ satisfy

$$|d_G(A,B)-p| \le 3(\epsilon/\alpha^3\beta)^{1/3}p \,.$$

This would complete the proof since $\alpha,\beta\ge 2\epsilon^{1/7}$ would imply $|d_G(A,B)-p|\le 2\epsilon^{1/7}p$, as needed. Let $D=e(v,A)$, where v is chosen uniformly at random from V. Then

$$\mathbb{E}[D] = \frac{1}{|V|}\sum_{v\in V} e(v,A) = \frac{1}{|V|}\sum_{u\in A}\deg_G(u) = p|A| \,,$$

where in the last equality we used the assumption that the U side is regular. Moreover,

$$\mathbb{E}[D^2] = \frac{1}{|V|}\sum_{v\in V} e(v,A)^2 = \frac{1}{|V|}\sum_{u,u'\in A} \mathrm{codeg}(u,u') \le \epsilon p|U|^2 \cdot p + \sum_{u,u'\in A}(1+\epsilon p)p^2$$
$$= p^2|A|^2(\epsilon/\alpha^2 + 1 + \epsilon p) \le p^2|A|^2(1+2\epsilon/\alpha^2)\ ,$$

where in the first inequality we used the fact that G is $(p, \epsilon p)$-quasirandom together with the regularity of the U side. It follows that

$$\mathrm{Var}[D] = \mathbb{E}[D^2] - \mathbb{E}[D]^2 \le p^2|A|^2(1+2\epsilon/\alpha^2) - (p|A|)^2 = (2\epsilon/\alpha^2)p^2|A|^2\ .$$

By Chebyshev's inequality, for any $\lambda > 0$ we have

$$\mathbb{P}\Big(\big|D - p|A|\big| \ge \lambda p|A|\Big) \le \frac{\mathrm{Var}[D]}{(\lambda p|A|)^2} \le \frac{2\epsilon}{\lambda^2\alpha^2}\ ,$$

and since $e_G(A,B) = \sum_{v\in B}\deg_A(v)$, we have

$$(|B| - (2\epsilon/\lambda^2\alpha^2)|V|)\cdot(1-\lambda)p|A| \le e_G(A,B) \le (2\epsilon/\lambda^2\alpha^2)|V|\cdot p|U| + |B|\cdot(1+\lambda)p|A|\ ,$$

where in the right inequality we used the assumption that the V side is regular. Therefore,

$$(1 - 2\epsilon/\lambda^2\alpha^2\beta)\cdot(1-\lambda)p \le d_G(A,B) \le (2\epsilon/\lambda^2\alpha^3\beta + 1 + \lambda)p\ ,$$

implying that

$$(1 - \lambda - \lambda^{-2}\cdot 2\epsilon/\alpha^2\beta)p \le d_G(A,B) \le (1 + \lambda + \lambda^{-2}\cdot 2\epsilon/\alpha^3\beta)p\ .$$

Taking $\lambda = (2\epsilon/\alpha^2\beta)^{1/3}$ for the lower bound and $\lambda = (2\epsilon/\alpha^3\beta)^{1/3}$ for the upper bound implies

$$(1 - 2(2\epsilon/\alpha^2\beta)^{1/3})p \le d_G(A,B) \le (1 + 2(2\epsilon/\alpha^3\beta)^{1/3})p\ .$$

In particular, this implies the desired bound,

$$|d_G(A,B) - p| \le 3(\epsilon/\alpha^3\beta)^{1/3}\cdot p\ ,$$

which completes the proof. ☐

5.3. **Modified blow-up.** Here we show how to execute the iterative process described in Subsection 5.1, given the pseudorandom bipartitions constructed in Subsection 5.2.

Let G be a d-regular graph. Let $n \in \mathbb{N}$, $\alpha, \beta \in [0,1]$ be such that there exists an (n,d,α,β)-sequence. We define $G(n,d,\alpha,\beta)$ as any graph obtained as follows, where here we use $N(x)$ to denote the neighbors of vertex x in G. We first replace each vertex x of G by a set X of n new vertices. For this paragraph, if $y \in N(x)$ and we replaced x with X and y with Y, then we will say that $Y \in N(X)$. For each X and $Y \in N(X)$, we *associate* with Y a bipartition $(X_{Y,0}, X_{Y,1})$ of X, so that the sequence of bipartitions $\{(X_{Y,0}, X_{Y,1})\}_{Y\in N(X)}$ is an (n,d,α,β)-sequence. For each edge $e = (x,y)$ of G we do either one of the following:

(i) we put two copies of $K_{n/2,n/2}$, between $(X_{Y,0}, Y_{X,0})$ and between $(X_{Y,1}, Y_{X,1})$, or
(ii) we put two copies of $K_{n/2,n/2}$, between $(X_{Y,0}, Y_{X,1})$ and between $(X_{Y,1}, Y_{X,0})$.

We note that for the proof of Theorem 4 the reader may assume that choice (i) is used for all edges; both choices will be used in Subsection 5.7. Since $v(G(n,d,\alpha,\beta)) = v(G)\cdot n$ and since $G(n,d,\alpha,\beta)$ is $d\cdot\frac{1}{2}n$-regular, we have

$$d_{G(n,d,\alpha,\beta)} = \frac{1}{2}d_G \ . \tag{21}$$

Claim 5.7. Let $G' = G(n,d,\alpha,\beta)$. Let $x \neq y, w \in V(G)$, where w is a common neighbor of x, y, and denote by X, Y, W the sets replacing them in G', respectively. For every $x' \in X$ and $\ell \in \{0,1\}$ we have $e_{G'}(x', W_{Y,\ell}) \le (\frac{1}{4}+\alpha)n$.

Proof. By construction, the set of neighbors of x' in W is precisely $W_{X,\ell'}$ for some $\ell' \in \{0,1\}$. This implies that $e_{G'}(x', W_{Y,\ell}) = |W_{X,\ell'} \cap W_{Y,\ell}| \le (\frac{1}{4}+\alpha)n$, where in the inequality we used the fact that $W_{X,\ell'}$ and $W_{Y,\ell}$ belong to two distinct bipartitions in an α-orthogonal sequence of bipartitions. ☐

Claim 5.8. Let $G' = G(n,d,\alpha,\beta)$. If G is (p,ϵ)-quasirandom, then any blow-up of G' is $(\frac{1}{2}p, \epsilon')$-quasirandom with $\epsilon' = \epsilon + \max\{8\alpha,\ 2/v(G)\}$.

Proof. Let G° be a blow-up of G', and note that $d_{G^\circ} = \frac{1}{2}p$ follows from (21). Put $G = (U,V;E)$ and $G^\circ = (U',V';E')$, and put $|U'|/|U| = |V'|/|V| = k$. Suppose $u, v \in V(G^\circ)$ lie in the blow-up of $x, y \in V(G)$, respectively, with $x \neq y \in U$. We claim that

$$\text{codeg}_{G^\circ}(u,v) \le \Big(\frac{1}{4}+\alpha\Big)k \cdot \text{codeg}_G(x,y) \ .$$

This would imply that all but

$$\epsilon|U|^2k^2 + |U'|^2/|U| = |U'|^2(\epsilon + 1/|U|) = |U'|^2(\epsilon + 2/v(G))$$

pairs $(u,v) \in U'^2$ satisfy

$$\begin{aligned}\text{codeg}_{G^\circ}(u,v) &\le \Big(\frac{1}{4}+\alpha\Big)k\cdot(1+\epsilon)p^2|V| = (1+4\alpha)(1+\epsilon)\Big(\frac{1}{2}p\Big)^2|V'|\\ &\le (1+\epsilon+8\alpha)\Big(\frac{1}{2}p\Big)^2|V'| \ ,\end{aligned}$$

which would complete the proof.

To prove the claim above, first note that if a vertex of G° that lies in the blow-up of $w \in V(G)$ is a common neighbor of u and v in G°, then, by construction, w must be a common neighbor of x and y in G. It follows from Claim 5.7 that the number of common neighbors of u and v in the blow-up of w is at most $(\frac{1}{4}+\alpha)k$. This implies that $\text{codeg}_{G^\circ}(u,v) \le (\frac{1}{4}+\alpha)k\cdot\text{codeg}_G(x,y)$, proving our claim above. ☐

5.3.1. *Iterated modified blow-up.* Let G be a d_0-regular graph. Let $n_i \in \mathbb{N}$, $\alpha_i, \beta_i \in [0,1]$ be such that for every $1 \le i \le r$ there exists an $(n_i, d_{i-1}, \alpha_i, \beta_i)$-sequence where $d_{i-1} = d_0\prod_{j=1}^{i-1}(n_j/2)$. For every $1 \le i \le r$ put $\rho_i = (n_i, d_{i-1}, \alpha_i, \beta_i)$. We define $G(\rho_1,\ldots,\rho_r)$ as any graph recursively obtained as

$$G(\rho_1,\ldots,\rho_i) = [G(\rho_1,\ldots,\rho_{i-1})](n_i, d_{i-1}, \alpha_i, \beta_i) \ ,$$

with G as the base case. This is well defined since for every $1 \le i \le r$ the graph $G(\rho_1,\ldots,\rho_{i-1})$ is d_{i-1}-regular. We have the following by (21).

Fact 5.9. The bipartite graph $K_{n_0,n_0}(\rho_1,\ldots,\rho_r)$ is regular of density $1/2^r$.

In order prove the (ϵ)-regularity of an iterated modified blow-up, we analyze the effect of each iteration on its (ϵ,p)-quasirandomness, and then finally apply Lemma 5.6.

Claim 5.10. Any blow-up of $K_{n_0,n_0}(\rho_1,\dots,\rho_r)$ is $(1/n_0^{1/14})$-regular, provided $\alpha_i \le 1/(8n_0\cdots n_{i-1})$ for every $1 \le i \le r$ and $n_0 \ge 4^{r+8}$.

Proof. By definition, K_{n_0,n_0} is $(1,0)$-quasirandom. By Claim 5.8 and Fact 5.9, any blow-up H of $K_{n_0,n_0}(\rho_1,\dots,\rho_r)$ is (p,ϵ)-quasirandom with $p = 1/2^r$ and

$$\epsilon \le \sum_{i=1}^{r} 1/(n_0\cdots n_{i-1}) \le (1/n_0)\sum_{i=1}^{r} 1/2^{i-1} \le 2/n_0 \ ,$$

where in the first inequality we used the fact that $v(K_{n_0,n_0}(\rho_1,\dots,\rho_{r-1})) = 2n_0\cdots n_{i-1}$ and in the second inequality we used the fact that $n_j \ge 2$ for $j \ge 1$. It follows from Lemma 5.6 that, since H is $(p,\epsilon' p)$-quasirandom with $\epsilon' = 2^{r+1}/n_0$, it is also (ϵ'')-regular with

$$\epsilon'' = 2\epsilon'^{1/7} = (2^{r+8}/n_0)^{1/7} \ .$$

By the claim's assumption that $2^{r+8} \le \sqrt{n_0}$ we have $\epsilon'' \le 1/n_0^{1/14}$, which completes the proof. □

5.4. **The graph G_s°.** We are now ready to formally define the graph that will be used to prove Theorem 4. Let $s \in \mathbb{N}$ be even with

$$s \ge 400 \ , \tag{22}$$

and put $n_0 = 4^{s+8}$. Our graph, which we denote by G_s°, will be of density $p := 1/2^s$. First, for every $1 \le r \le s$ put $n_r = 2^{\lfloor n_{r-1}/200 \rfloor}$. Note that

$$n_s \ge \mathrm{twr}(s/2) \ , \tag{23}$$

since $n_{r+2} \ge 2^{n_r}$ (as $n_r \ge n_0$ is sufficiently large). Moreover, for $1 \le r \le s$ put $\alpha_r = 1/(8n_0\cdots n_{r-1})$ and $d_{r-1} = n_0\prod_{j=1}^{r-1}(n_j/2)$.

We recursively construct graphs $G_0, G_1, \dots, G_s$, starting from $G_0 = K_{n_0,n_0}$, in the same manner described in the previous subsection. More precisely, setting $\rho_r = (n_r, d_{r-1}, \alpha_r, 1/16)$ for each $1 \le r \le s$, we let

$$G_r = K_{n_0,n_0}(\rho_1,\dots,\rho_r) \ . \tag{24}$$

Importantly, (24) is well defined since there exists a ρ_r-sequence for every $1 \le r \le s$. Indeed, this follows from Lemma 5.1 since $d_{r-1} \ge n_0 \ge 200$, n_r is even, $n_r \le 2^{\lfloor d_{r-1}/200 \rfloor}$ (as $n_{r-1} \le d_{r-1}$), and

$$\sqrt{2\ln(d_{r-1})/n_r} \le 1/n_{r-1}^2 \le 1/(8n_0\cdots n_{r-1}) = \alpha_r \ .$$

We let our final graph G_s° be any blow-up of G_s. Note that by Fact 5.9, G_s° is a regular bipartite graph of density $p = 1/2^s$.

5.4.1. *Properties of* G_s°. Recall that in the process of constructing G_s°, each vertex of G_r is repeatedly replaced by a set of new vertices. For $0 \le r \le s$ let $\mathcal{X}_r$ be the partition of $V(G_s^\circ)$ whose parts correspond to the vertices of G_r. Therefore, in what follows, we will interchangeably refer to $X \in \mathcal{X}_r$ also as a vertex of G_r or as a cluster of vertices in one of the graphs $G_{r+1},\dots,G_s$.

Observe that each $\mathcal{X}_r$ refines $\mathcal{X}_{r-1}$ and that $\mathcal{X}_r$ is an equipartition of order

$$|\mathcal{X}_r| = v(G_r) = 2\prod_{i=0}^{r} n_i \ .$$

In particular, we have

$$|\mathcal{X}_0| = 2n_0 = 2^{17}\cdot 4^s \ , \tag{25}$$

and moreover, using (23),

$$|\mathcal{X}_s| \geq n_s \geq \mathrm{twr}(s/2)\ . \tag{26}$$

If $X, Y \in \mathcal{X}_r$ with $r < s$ and $(X, Y) \in E(G_r)$, then we denote by $(X_{Y,0}, X_{Y,1})$ the bipartition of X that is *associated* with Y in the construction of G_{r+1} from G_r (recall the definition of a modified blow-up in Subsection 5.3). Thus, $X_{Y,0}$ and $X_{Y,1}$ are each a union of parts in $\mathcal{X}_{r+1}$. Similarly, we denote by $(Y_{X,0}, Y_{X,1})$ the bipartition of Y that is associated with X. We will need the following properties of G_s°. We first note that from (22) we have

$$p = 1/2^s \leq 2^{-400}\ . \tag{27}$$

Claim 5.11. Let $1 \leq r \leq s$, and let $X, Y \in \mathcal{X}_{r-1}$ with $(X, Y) \in E(G_{r-1})$. For every $\ell \in \{0, 1\}$ there is $\ell' \in \{0, 1\}$ such that:

- $d_{G_s^{\circ}}(X_{Y,\ell}, Y_{X,\ell'}) \neq 0$. In particular, $d_{G_s^{\circ}}(X, Y) \neq 0$.
- Every $v \in X_{Y,\ell}$ satisfies $d_{G_s^{\circ}}(v, Y_{X,\ell'}) = 2^r p$ and $d_{G_s^{\circ}}(v, Y_{X,1-\ell'}) = 0$. In particular, $d_{G_s^{\circ}}(v, Y) = 2^{r-1}p$.

Proof. As the first item follows from the second, we prove the latter. By construction, the edge (X, Y) of G_{r-1} is replaced in G_r by two copies of $K_{k,k}$ (with $k = \frac{1}{2}n_r$) and two copies of its complement $\overline{K_{k,k}}$. Specifically, $G_r[X_{Y,\ell}, Y_{X,\ell'}] \simeq K_{k,k}$ and $G_r[X_{Y,\ell}, Y_{X,1-\ell'}] \simeq \overline{K_{k,k}}$, where $\ell' = \ell$ if choice (i) in Section 5.3 is used, and $\ell' = 1 - \ell$ if choice (ii) is used. In the construction of G_{r+1}, the above copy of $K_{k,k}$ is turned into a modified blow-up of $K_{k,k}$; that is, $G_{r+1}[X_{Y,\ell}, Y_{X,\ell'}] \simeq K_{k,k}(\rho'_{r+1})$ with $\rho'_{r+1} = (n_{r+1}, k, \alpha_{r+1}, 1/2)$. This follows from the fact that $G_{r+1}[X_{Y,\ell}, Y_{X,\ell'}]$ is a subgraph of G_{r+1} together with Fact 5.2. Indeed, for each vertex, its associated sequence of bipartitions in $G_{r+1}[X_{Y,\ell}, Y_{X,\ell'}]$ is a subsequence of its associated sequence in G_{r+1}. Continuing in this manner, we deduce that $G_s[X_{Y,\ell}, Y_{X,\ell'}] \simeq K_{k,k}(\rho'_{r+1}, \ldots, \rho'_s)$ (with $\rho'_i = (n_i, \prod_{j=r}^{i-1}(n_j/2), \alpha_i, 1/2)$), which is regular of density $1/2^{s-r} = 2^r p$ by Fact 5.9. This completes the proof. ☐

Claim 5.12. For every $1 \leq r \leq s$, every $X, Y \in \mathcal{X}_{r-1}$ with $d_{G_s^{\circ}}(X, Y) \neq 0$, every $v \in V(G_s^{\circ}) \setminus X$, and every $\ell \in \{0, 1\}$, we have $d_{G_s^{\circ}}(v, Y_{X,\ell}) \leq \frac{5}{8}2^r p$.

Proof. Suppose $v \in A' \subseteq A$ with $A \in \mathcal{X}_{r-1}$ and $A' \in \mathcal{X}_r$, where by assumption $A \neq X$. Recall that $G_r = G_{r-1}(n_r, d_{r-1}, \alpha_r, 1/16)$. Apply Claim 5.7 on $G = G_{r-1}$, $G' = G_r$ and with x, y, w, x' corresponding to A, X, Y, A', respectively. It follows that the fraction of $Y' \in \mathcal{X}_r$ with $Y' \subseteq Y_{X,\ell}$ that satisfy $d_{G_s^{\circ}}(A', Y') \neq 0$ is at most $2(\frac{1}{4} + \alpha_r) \leq 5/8$, where the last inequality uses the fact that, by construction, $\alpha_r \leq 1/16$. By the second item in Claim 5.11 we have $d_{G_s^{\circ}}(v, Y') \leq 2^r p$ for each of the Y' above, hence $d_{G_s^{\circ}}(v, Y_{X,\ell}) \leq \frac{5}{8}2^r p$, as needed. ☐

Summarizing Claims 5.11 and 5.12, we have the following regarding the degrees in G_s°.

Claim 5.13. Let $1 \leq r \leq s$, and let $X, Y \in \mathcal{X}_{r-1}$ with $d_{G_s^{\circ}}(X, Y) \neq 0$. If $d_{G_s^{\circ}}(X_{Y,\ell}, Y_{X,\ell'}) \neq 0$ then for every vertex $v \in V(G_s^{\circ})$ we have

$$d_{G_s^{\circ}}(v, Y_{X,\ell'}) = \begin{cases} \leq \frac{5}{8}2^r p & \text{if } v \notin X, \\ 2^r p & \text{if } v \in X_{Y,\ell}, \\ 0 & \text{if } v \in X_{Y,1-\ell}. \end{cases}$$

Claim 5.14. For $0 \leq r \leq s$ and $X \in \mathcal{X}_r$, the number of $Y \in \mathcal{X}_r$ with $d_{G_s^\circ}(X,Y) \neq 0$ is $|\mathcal{X}_r|/2^{r+1}$.

Proof. By (24) and Fact 5.9, every vertex of G_r has precisely $\frac{1}{2}|V(G_r)|/2^r$ neighbors. Recalling that the parts of $\mathcal{X}_r$ correspond to the vertices of G_r, it follows that the number of $Y \in \mathcal{X}_r$ with $d_{G_s^\circ}(X,Y) \neq 0$ is, using the first item in Claim 5.11, $\frac{1}{2}|\mathcal{X}_r|/2^r = |\mathcal{X}_r|/2^{r+1}$. ☐

We will also need the following two pseudorandom properties of G_s°.

Claim 5.15. Let $Z \subseteq V(G_s^\circ)$, and let $1 \leq r \leq s$. Suppose $|Z \setminus X| \leq \zeta|Z|$ for some $X \in \mathcal{X}_{r-1}$, while $|Z \setminus X'| \geq \zeta'|Z|$ for every $X' \in \mathcal{X}_r$. For at least $\frac{1}{6}|\mathcal{X}_{r-1}|/2^r$ clusters $Y \in \mathcal{X}_{r-1}$ we have

$$\min\{|Z \cap X_{Y,0}|, |Z \cap X_{Y,1}|\} \geq \frac{1}{8}(\zeta' - \zeta)|Z| \,.$$

Proof. Let $\widetilde{X} = \{X_1, \ldots, X_{n_r}\}$ be the partition of X into parts of $\mathcal{X}_r$. Recall that each of the $|\mathcal{X}_{r-1}|/2^r$ clusters $Y \in \mathcal{X}_{r-1}$ with $d_{G_s^\circ}(X,Y) \neq 0$ (see Claim 5.14) is associated with a bipartition of $\widetilde{X}$, and that the sequence of these bipartitions is 1/16-balanced. Apply Lemma 5.3 on this sequence with $\lambda_t = |Z \cap X_t|/|Z \cap X|$. Thus, for at least $\frac{1}{6}|\mathcal{X}_{r-1}|/2^r$ clusters $Y \in \mathcal{X}_r$ we have

$$\min\{|Z \cap X_{Y,0}|, |Z \cap X_{Y,1}|\} \geq \frac{1}{8}\Big(|Z \cap X| - \max_t |Z \cap X_t|\Big) \geq \frac{1}{8}(\zeta' - \zeta)|Z| \,,$$

where the last inequality uses the fact that, by the assumptions in the statement, $|Z \cap X| \geq (1-\zeta)|Z|$, while $|Z \cap X_t| \leq (1-\zeta')|Z|$ for every t. ☐

We write $n = |V(G_s^\circ)|$ and, recalling that G_s° is bipartite, we write $G_s^\circ = (U, V; E)$.

Claim 5.16. Let $A \subseteq U$, $B \subseteq V$. If $|A| \geq p^{1/7}n$ and $|B| \leq \frac{1}{512}n$, then $e_{G_s^\circ}(A,B) \leq \frac{1}{256}pn|A|$.

Proof. First, we prove that G_s, and hence G_s°, is (ϵ)-regular with $\epsilon \leq p^{1/7}$ (recall (20)). Recalling $G_s = K_{n_0,n_0}(\rho_1, \ldots, \rho_s)$, we apply Claim 5.10 on G_s using the fact that $\alpha_i = 1/8n_0 \cdots n_{i-1}$ for every $1 \leq i \leq r$, and the fact that $n_0 = 4^{s+8} \geq 4^{r+8}$. It follows that G_s is (ϵ)-regular with $\epsilon \leq 1/n_0^{1/14} \leq 1/2^{(s+8)/7} \leq 1/2^{s/7} = p^{1/7}$, as desired.

Now, if $|B| \geq \epsilon|V|$, then $e_{G_s^\circ}(A,B) \leq (1+\epsilon)p|A||B| \leq \frac{1}{256}p|A|n$, as needed. Suppose otherwise that $|B| \leq \epsilon|V|$. Note that $\epsilon \leq 2^{-8}$ by (27). We have $|V \setminus B| \geq (1-\epsilon)|V| \geq \epsilon|V|$, and thus, by the (ϵ)-regularity of G_s°, we get $e_{G_s^\circ}(A, V \setminus B) \geq (1-\epsilon)p|A||V \setminus B|$. Therefore,

$$\begin{aligned} e_{G_s^\circ}(A,B) &= e_{G_s^\circ}(A,V) - e_{G_s^\circ}(A, V \setminus B) \leq p|A|(|V| - (1-\epsilon)|V \setminus B|) \\ &\leq p|A|(|B| + \epsilon|V|) \leq \frac{1}{256}pn|A| \,, \end{aligned}$$

where we used the fact that $e_{G_s^\circ}(A,V) = p|A||V|$ since G_s° is regular of density p. This completes the proof. ☐

5.5. **Lower bound proof.** For sets S, T we write $S \subseteq_\beta T$ if $|S \setminus T| \le \beta|S|$. For a partition $\mathcal{P}$ we write $S \in_\beta \mathcal{P}$ if $S \subseteq_\beta P$ for some $P \in \mathcal{P}$. For partitions $\mathcal{P}, \mathcal{Q}$ of the same set of size n, we write $\mathcal{Q} \preceq_\beta \mathcal{P}$ if

$$\sum_{Q \in \mathcal{Q} \colon Q \notin_\beta \mathcal{P}} |Q| \le \beta n \, .$$

Note that for $\mathcal{Q}$ equitable, $\mathcal{Q} \preceq_\beta \mathcal{P}$ if and only if all but $\beta|\mathcal{Q}|$ parts $Q \in \mathcal{Q}$ satisfy $Q \in_\beta \mathcal{P}$ (as mentioned in Subsection 5.1).

Our main technical result towards proving Theorem 4 is the following, where we recall that G_s° denotes the graph of density $p = 2^{-s}$ constructed in Subsection 5.4. We say that a partition $\mathcal{Z}$ is *perfectly ϵ-regular* if all pairs of $\mathcal{Z}$ are ϵ-regular.

Theorem 10. *Let $\delta \le 2^{-32}$, and put $\gamma = \max\{64\sqrt{\delta}, p^{1/7}\}$. Let $\mathcal{Z} \preceq \mathcal{X}_0$ be a perfectly $\frac{1}{16}p$-regular partition of a graph that is δ-close to G_s°. Then $\mathcal{Z} \preceq_\gamma \mathcal{X}_s$.*

The proof of Theorem 10 appears in Subsection 5.6. Our goal in the rest of this subsection is to use Theorem 10 in order to prove the lower bound (4) in Theorem 4. For convenience, we restate the lower bound statement below.

Theorem 11 (SRAL lower bound, restated). *There are fixed constants $\delta_0, c > 0$ such that the following holds. If $\mathcal{Z}$ is a p^5-regular partition of a graph that is δ_0-close to G_s°, then $|\mathcal{Z}| \ge \operatorname{twr}(c \log \frac{1}{p})$.*

Note that Theorem 11 does not require the partition $\mathcal{Z}$ to be equitable. To prove Theorem 11 we will need the following corollary of Theorem 10.

Corollary 12. *Let $p^{2/7} \le \delta \le 2^{-33}$. If $\mathcal{Z}$ is a p^5-regular partition of a graph that is δ-close to G_s°, then $\mathcal{Z} \cap \mathcal{X}_0 \preceq_\gamma \mathcal{X}_s$ with $\gamma = 128\sqrt{\delta}$.*

Proof. Put $\mathcal{Z}_0 = \mathcal{Z} \cap \mathcal{X}_0$. Recall that $|\mathcal{X}_0| = 2^{17}p^{-2}$ by (25). By Claim 5.4, the partition $\mathcal{Z}_0$ is ϵ-regular with $\epsilon = \sqrt{8|\mathcal{X}_0|p^5} = 2^{10}p^{3/2}$. Note that

$$\epsilon \le \frac{1}{4}p^{9/7} \le \frac{1}{4}\delta p \le \frac{1}{16}p \, , \tag{28}$$

where the first inequality uses (27) to bound $2^{10}p^{1/2} \le \frac{1}{4}p^{2/7}$, and the second and third inequalities use the assumed bounds on δ. Since $\mathcal{Z}_0$ is an ϵ-regular partition of a graph that is δ-close to G_s°, it is also a perfectly $\frac{1}{16}p$-regular partition of a graph that is 2δ-close to G; indeed, such a graph is obtained by removing all edges between pairs of $\mathcal{Z}_0$ that are not ϵ-regular, of which there are, by (28), at most $\frac{1}{4}\delta p|V(G)|^2 = \delta e(G_s^\circ)$. We apply Theorem 10 with (the not necessarily equitable) $\mathcal{Z}_0$, using the fact that $\mathcal{Z}_0 \preceq \mathcal{X}_0$ and $2\delta \le 2^{-32}$. It follows that $Z_0 \preceq_\gamma \mathcal{X}_s$ with $\gamma \le \max\{128\sqrt{\delta}, p^{1/7}\} = 128\sqrt{\delta}$, where we again used the assumed lower bound on δ. ☐

We will also need the following fact about the order of an approximate refinement.

Claim 5.17. If $\mathcal{Q} \preceq_{1/4} \mathcal{P}$ and $\mathcal{P}$ is equitable, then $|\mathcal{Q}| \ge \frac{1}{2}|\mathcal{P}|$.

Proof. Let the function π map the parts $Q \in \mathcal{Q}$ satisfying $Q \subseteq_{1/4} P$ for some $P \in \mathcal{P}$ to that (unique) P. Denoting by n the number of elements in the underlying set, observe that the total number of elements in the parts $P \in \mathcal{P}$ that are not in the image of π is at most $\frac{1}{4}n + \sum_{Q \in \mathcal{Q}} \frac{1}{4}|Q| = \frac{1}{2}n$. As $\mathcal{P}$ is equitable, this means there are at least $\frac{1}{2}|\mathcal{P}|$ parts $P \in \mathcal{P}$ in the image of π, and therefore $|\mathcal{Q}| \ge \frac{1}{2}|\mathcal{P}|$, as claimed. ☐

Proof of Theorem 11. Suppose $\mathcal{Z}$ is a p^5-regular partition of a graph that is 2^{-33}-close to G_s°. By Corollary 12, the common refinement $\mathcal{Z}_0 := \mathcal{Z} \cap \mathcal{X}_0$ satisfies $\mathcal{Z}_0 \preceq_{1/4} \mathcal{X}_s$. By Claim 5.17, $|\mathcal{Z}_0| \geq \frac{1}{2}|\mathcal{X}_s|$. Since $|\mathcal{Z}_0| \leq |\mathcal{Z}||\mathcal{X}_0|$ we get $|\mathcal{Z}| \geq \frac{1}{2}|\mathcal{X}_s|/|\mathcal{X}_0|$, which completes the proof by (25) and (26). $\square$

5.6. **Proof of Theorem 10.**

Proof of Theorem 10. Put $G = G_s^\circ$ and $n = |V(G)|$. Let $\mathcal{Z}$ be a perfectly $\frac{1}{16}p$-regular partition of a graph G' on $V(G)$, and suppose $\mathcal{Z} \preceq \mathcal{X}_0$ yet $\mathcal{Z} \not\preceq_\gamma \mathcal{X}_s$. Our goal is to prove that G' is not δ-close to G, that is, $|E(G) \triangle E(G')| > \delta \cdot p(n/2)^2$.

Let $1 \leq R \leq s$ be the smallest integer such that $\mathcal{Z} \not\preceq_\gamma \mathcal{X}_R$. For each $1 \leq r \leq R$ let

$$\mathcal{D}_r = \{Z \in \mathcal{Z} \,:\, Z \notin_\gamma \mathcal{X}_r \text{ and } Z \in_\gamma \mathcal{X}_{r-1}\}\ .$$

We let $\mathcal{B}_r$, with $1 \leq r \leq s$, be the set of vertices that either lie in some $Z \notin_\gamma \mathcal{X}_{r-1}$ or lie in some $Z \setminus X$ with $Z \subseteq_\gamma X \in \mathcal{X}_{r-1}$. More formally,

$$\mathcal{B}_r = V(G) \setminus \bigcup_{\substack{X \in \mathcal{X}_{r-1}, Z \in \mathcal{Z}:\\ Z \in_\gamma X}} (Z \cap X)\ .$$

Note that since $\mathcal{X}_s \preceq \cdots \preceq \mathcal{X}_0$ we have that $\mathcal{B}_1 \subseteq \cdots \subseteq \mathcal{B}_s$ and, furthermore, that $Z \in \mathcal{D}_r$ for at most one value of r. Throughout, if $\mathcal{F}$ is a family of disjoint sets we denote by $\|\mathcal{F}\| = \left|\bigcup_{F \in \mathcal{F}} F\right|$ the "total" size of $\mathcal{F}$. Put $\mathcal{D} = \bigcup_{r=1}^{R} \mathcal{D}_r$. Since $\mathcal{Z} \not\preceq_\gamma \mathcal{X}_R$ yet $\mathcal{Z} \preceq_\gamma \mathcal{X}_{R-1}$, we have

$$\|\mathcal{D}\| > \gamma n \quad \text{yet} \quad |\mathcal{B}_R| \leq 2\gamma n \leq 2^{-9} n\ , \tag{29}$$

where the first inequality uses the fact that $\mathcal{D} = \{Z \in \mathcal{Z} \,|\, Z \notin_\gamma \mathcal{X}_R\}$ as every $Z \in \mathcal{Z}$ satisfies $Z \in_0 \mathcal{X}_0$ by assumption, and the last inequality uses the assumed bound on δ as well as (27) in order to bound

$$\gamma \leq \max\{64 \cdot 2^{-16}, p^{1/7}\} = 2^{-10}\ . \tag{30}$$

Let $Z \in \mathcal{D}_r$. Let $X \in \mathcal{X}_{r-1}$ be the unique cluster such that $Z \subseteq_\gamma X$ (recall that $\gamma < 1/2$). Let $Y \in \mathcal{X}_{r-1}$ be one of the $|\mathcal{X}_{r-1}|/2^r$ clusters with $d_{G_s^\circ}(X, Y) \neq 0$ (recall Claim 5.14). Call Y *good* if $Z \nsubseteq_{\frac{1}{16}\gamma} X_{Y,i}$ for each $i \in \{0, 1\}$. Denote by $g(Z)$ the set of all clusters that are good for Z. We claim that for every $Z \in \mathcal{D}_r$ we have

$$|g(Z)| \geq \frac{1}{6}|\mathcal{X}_{r-1}|/2^r\ . \tag{31}$$

Indeed, if $Z \nsubseteq_{\frac{1}{2}\gamma} X$ this is clear (actually in this case $|g(Z)| = |\mathcal{X}_{r-1}|/2^r$), and otherwise this follows from Claim 5.15 with $\zeta' = \gamma$ and $\zeta = \frac{1}{2}\gamma$.

Put $\alpha = \frac{1}{16}\gamma$. Fix $Y \in g(Z)$, and let $\ell \in \{0, 1\}$ satisfy $|Z \cap X_{Y,\ell}| \geq |Z \cap X_{Y,1-\ell}|$. Since $Z \subseteq_\gamma X$ we have

$$|Z \cap X_{Y,\ell}| \geq \frac{1}{2}|Z \cap X| \geq \frac{1}{2}(1-\gamma)|Z| \geq \frac{7}{16}|Z|\ , \tag{32}$$

where the last inequality uses (30). Furthermore, since $Z \nsubseteq_{\frac{1}{16}\gamma} X_{Y,\ell}$ we have

$$|Z \setminus X_{Y,\ell}| \geq \alpha|Z|\ . \tag{33}$$

Let Z_1 be an arbitrary subset of $Z \cap X_{Y,\ell}$ of size $\alpha|Z|$, and let Z_2 be an arbitrary subset of $Z \setminus X_{Y,\ell}$ of size $\alpha|Z|$ (both choices are possible by (32) and (33)). By Claim 5.11 there is $\ell' \in \{0, 1\}$ with $d_G(X_{Y,\ell}, Y_{X,\ell'}) \neq 0$. Put

$$Y' = Y_{X,\ell'} \quad \text{and} \quad Y^* = Y' \setminus \mathcal{B}_r\ .$$

Notice that $\alpha \geq \frac{1}{16}p^{1/7} \geq \frac{1}{16}p$. As $\mathcal{Z}$ is a perfectly $\frac{1}{16}p$-regular partition of G', for every $Z' \in \mathcal{Z}$ we have

$$\begin{aligned} e_{G'}(Z_1, Z'\cap Y') - e_{G'}(Z_2, Z'\cap Y') &= (d_{G'}(Z_1, Z'\cap Y') - d_{G'}(Z_2, Z'\cap Y'))\alpha|Z||Z'\cap Y'| \\ &\leq \frac{1}{8}p\alpha|Z||Z'| \,, \end{aligned}$$

where, denoting $W = Z' \cap Y'$, the last inequality bounds $d_{G'}(Z_1, W) - d_{G'}(Z_2, W)$ by $2 \cdot \frac{1}{16}p$ if $|W| \geq \frac{1}{16}p|Z'|$, and otherwise bounds $d_{G'}(Z_1, W) - d_{G'}(Z_2, W)$ by 1. Summing over all $Z' \subseteq_\gamma Y$ gives

$$\begin{aligned} e_{G'}(Z_1, Y^*) - e_{G'}(Z_2, Y^*) &= \sum_{\substack{Z' \in \mathcal{Z}: \\ Z' \subseteq_\gamma Y}} (e_{G'}(Z_1, Z' \cap Y') - e_{G'}(Z_2, Z' \cap Y')) \\ &\leq \frac{1}{4}p \cdot \alpha|Z||Y^*| \leq \frac{1}{4}p\alpha|Z||Y| \,, \end{aligned} \tag{34}$$

where the equality uses the fact that $Y^* = \bigcup_{Z' \subseteq_\gamma Y} Z' \cap Y'$ and the first inequality uses the fact that $|Z'| \leq 2|Z' \cap Y|$ for every $Z' \subseteq_\gamma Y$ as $\gamma \leq \frac{1}{2}$. On the other hand, in G we have that (notice $Y' \cap \mathcal{B}_r = Y' \setminus Y^*$)

$$\begin{aligned} e_G(Z_1, Y^*) - e_G(Z_2, Y^*) &\geq (d_G(Z_1, Y') - d_G(Z_2, Y'))\alpha|Z||Y'| - e_G(Z_1, Y' \cap \mathcal{B}_r) \\ &\geq \frac{3}{8}2^r p \cdot \alpha|Z| \cdot \frac{1}{2}|Y| - e_G(Z_1, Y \cap \mathcal{B}_R) \,, \end{aligned} \tag{35}$$

where the second inequality uses Claim 5.13 (all three cases) and the fact that $Y' \subseteq Y$ and $\mathcal{B}_r \subseteq \mathcal{B}_R$.

For every pair of disjoint subsets $S, T \subseteq V(G)$, denote

$$\Delta(S,T) = |E_G(S,T) \triangle E_{G'}(S,T)|.$$

Note that $\Delta(S,T) \geq |E_G(S,T) - E_{G'}(S,T)|$. We get

$$\begin{aligned} \Delta(Z,Y) &\geq \Delta(Z_1, Y^*) + \Delta(Z_2, Y^*) \\ &\geq (e_G(Z_1, Y^*) - e_{G'}(Z_1, Y^*)) + (e_{G'}(Z_2, Y^*) - e_G(Z_2, Y^*)) \\ &\geq \frac{3}{16}2^r p\alpha|Z||Y| - e_G(Z_1, Y \cap \mathcal{B}_R) - \frac{1}{4}p\alpha|Z||Y| \\ &\geq \frac{1}{16}2^r p\alpha|Z||Y| - e_G(Z_1, Y \cap \mathcal{B}_R), \end{aligned} \tag{36}$$

where the third inequality uses (34) and (35), and the last equality bounds $p \leq \frac{1}{2}2^r p$ as $r \geq 1$. Recall that the above applies for every choice of a subset Z_1 of $Z \cap X_{Y,\ell}$ of size $\alpha|Z|$. Note that by choosing such Z_1 uniformly at random, we have

$$\mathbb{E}[e_G(Z_1, Y \cap \mathcal{B}_R)] = \frac{|Z_1|}{|Z \cap X_{Y,\ell}|} \cdot e_G(Z \cap \mathcal{X}_{Y,\ell}, Y \cap \mathcal{B}_R) \leq \frac{16}{7}\alpha \cdot e_G(Z, Y \cap \mathcal{B}_R) \,,$$

where the inequality uses (32). Thus, there is Z_1 for which $e_G(Z_1, Y \cap \mathcal{B}_R) \leq \frac{16}{7}\alpha \cdot e_G(Z, Y \cap \mathcal{B}_R)$. Substituting into (36) implies that

$$\Delta(Z,Y) \geq \alpha\Big(\frac{1}{16}2^r p \cdot |Z||Y| - \frac{16}{7}e_G(Z, Y \cap \mathcal{B}_R)\Big) \,.$$

Summarizing, for every $1 \le r \le R$ and every $Z \in \mathcal{D}_r$ we have, using (31), that

$$\Delta(Z, V(G)) \ge \sum_{Y \in g(Z)} \Delta(Z, Y) \ge \alpha\Big(\frac{pn|Z|}{96} - \frac{16}{7} e_G(Z, \mathcal{B}_R)\Big) . \tag{37}$$

As G is bipartite, let U, V denote the vertex classes of G, and note that every $Z \in \mathcal{Z}$ is contained in either U or V, since $\mathcal{Z} \preceq \mathcal{X}_0$. Assume without loss of generality that $\mathcal{D}' := \{Z \in \mathcal{D} : Z \subseteq U\}$ satisfies $\|\mathcal{D}'\| \ge \frac{1}{2}\|\mathcal{D}\|$. We can now prove a lower bound on $|E(G) \triangle E(G')|$:

$$\begin{aligned} |E(G)\triangle E(G')| &\ge \sum_{r=1}^{R} \sum_{Z \in \mathcal{D}_r} \Delta(Z, V(G)) \ge \alpha \sum_{Z \in \mathcal{D}'} \Big(\frac{pn|Z|}{96} - \frac{16}{7} e_G(Z, \mathcal{B}_R)\Big) \\ &= \alpha\Big(\frac{pn\,\|\mathcal{D}'\|}{96} - \frac{16}{7} e_G\Big(\bigcup_{Z \in \mathcal{D}'} Z, \mathcal{B}_R\Big)\Big) \ge \alpha\,\|\mathcal{D}'\|\,\Big(\frac{pn}{96} - \frac{16}{7}\frac{pn}{256}\Big) \\ &\ge \frac{1}{32}\gamma^2 \cdot \frac{1}{672} pn^2 > \frac{(64\sqrt{\delta})^2}{2^{16}} pn^2 = \delta p (n/2)^2 , \end{aligned}$$

where the second inequality uses (37), the third inequality uses Claim 5.16 (with $A = \bigcup_{Z \in \mathcal{D}'} Z$ and $B = \mathcal{B}_r$ while relaying on (29) to bound $|A| \ge \gamma n \ge p^{1/7} n$ and $|B| \le 2^{-9} n$), the fourth inequality uses (29) to bound $\|\mathcal{D}'\|$ from below, and the last inequality uses the fact that $\gamma \ge 64\sqrt{\delta}$. Thus, we have shown that G' is not δ-close to G, completing the proof. □

5.7. **SRAL and lower bounds for hypergraph regularity.** We start with proving Theorem 4 by constructing a decomposition of $K_{N,N}$ into graphs witnessing (4). First, we generalize the definition of a modified blow-up of a graph to a definition of an edge coloring of a graph.

5.7.1. *Multicolored modified blow-up.* Let $\mathcal{G}$ be a q-edge-colored graph whose q graphs are each d-regular. Let $n \in \mathbb{N}$, $\alpha, \beta \in [0,1]$ be such that there exists an (n, d, α, β)-sequence. We define a $2q$-edge-colored graph $\mathcal{G}' = \mathcal{G}(n, d, \alpha, \beta)$ as follows. Each vertex x of $\mathcal{G}$ is replaced by a set of n new vertices X. Each edge (x, y) of $\mathcal{G}$ in color i is replaced by (using the notation of Subsection 5.3) two copies of $K_{n/2,n/2}$ in color i_1, between $(X_{Y,0}, Y_{X,0})$ and between $(X_{Y,1}, Y_{X,1})$, as well as two copies of $K_{n/2,n/2}$ in color i_2, between $(X_{Y,0}, Y_{X,1})$ and between $(X_{Y,1}, Y_{X,0})$. Here, i_1 and i_2 are two new colors, hence $\mathcal{G}'$ is indeed $2q$-edge-colored. Importantly, by using both choices available in the definition of a modified blow-up (see Subsection 5.3), the graphs of color i_1 and of color i_2 are each a modified blow-up of the graph of $\mathcal{G}$ of color i.

5.7.2. *Multicolored construction.* Consider the 2^s-edge-colored bipartite graph obtained by iterating the above s times starting from the graph K_{n_0,n_0}, where s, n_0 and the parameters (n, d, α, β) for each iteration are chosen as in Subsection 5.4. Let $\mathcal{G}^*_s$ be any blow-up of the colored graph above, meaning that each edge in color i is replaced by a complete bipartite graph in color i. It follows from the definition of a multicolored modified blow-up above that each of the 2^s graphs of $\mathcal{G}^*_s$ is of the form G°_s. In particular, each is a bipartite graph of density $p = 2^{-s}$, and together they form a partition of the edges of a $K_{N,N}$.

Proof of Theorem 4. Follows from the construction above together with Theorem 11. □

Let us now explain the relevance of the multicolored lower bound to lower bounds for hypergraph regularity. As part of the usual proof of the 3-graph regularity lemma, one is confronted with the following task; given a complete bipartite graph $K_{N,N}$ whose edges are partitioned into sparse graphs—or equivalently, are colored by many different colors—find a partition that is ϵ-regular (with ϵ depending on the density) for all graphs simultaneously.[13] As explained before Proposition 1.2, it in fact suffices to solve this task with the additional flexibility of modifying δN^2 of the edges. This raises the question of whether the additional flexibility allows one to do better than a tower-type bound. Using the multicolored graph $\mathcal{G}_s^*$ constructed above, Theorem 13 below shows that this task remains hard even if edge modifications are allowed. In fact, it remains hard even if the partition is required to be regular only for a *negligible* fraction of the graphs. We emphasize that Theorem 13 does not follow from (4), but rather requires the fact that $K_{N,N}$ can be decomposed into sparse bipartite graphs, all of which are hard for SRAL.

Theorem 13. *Let* $p^{4/7} \le \delta \le 2^{-66}$. *Let* $\mathcal{Z}$ *be a partition of* $V(\mathcal{G}_s^*)$, *and suppose that one can swap the colors of at most* δN^2 *edges of* $\mathcal{G}_s^*$ *so that* $\mathcal{Z}$ *is a* p^5*-regular partition for at least* $\sqrt{\delta} \cdot 2^s$ *of its graphs (over* $V(\mathcal{G}_s^*)$*). Then* $\mathcal{Z} \cap \mathcal{X}_0 \preceq_\gamma \mathcal{X}_s$ *with* $\gamma = 128\sqrt[4]{\delta}$. *In particular,*

$$|\mathcal{Z}| \ge \operatorname{twr}(\Omega(\log(1/p))) \,.$$

Proof. By averaging, there are fewer than $\sqrt{\delta} \cdot 2^s$ graphs G of $\mathcal{G}_s^*$ for which the number of edges that are added/removed is greater than $\sqrt{\delta} \cdot e(G)$. Therefore, there exists a graph that is $\sqrt{\delta}$-close to a graph of $\mathcal{G}_s^*$ for which $\mathcal{Z}$ is a p^5-regular partition. Since every graph of $\mathcal{G}_s^*$ is of the form G_s°, Corollary 12 implies that $\mathcal{Z} \cap \mathcal{X}_0 \preceq_\gamma \mathcal{X}_s$ with $\gamma = 128\sqrt[4]{\delta}$, as desired. In particular, $|\mathcal{Z}| \ge |\mathcal{Z} \cap \mathcal{X}_0|/|\mathcal{X}_0| \ge \frac{1}{2}|\mathcal{X}_s|/|\mathcal{X}_0| = \operatorname{twr}(\Omega(\log \frac{1}{p}))$, by Claim 5.17 and (25), (26). ☐

Appendix A. Proof of the (stronger) weak regularity lemma

Here we give a proof of Theorem 6, which closely follows the proof in [18].

Proof of Theorem 6. Let $G = (V, E)$ be a graph. Suppose the partition $\mathcal{P} = \{V_1, \dots, V_k\}$ of V is not weak ϵ-regular, and let $S, T \subseteq V$ be disjoint sets that witness this. Then, recalling the notation $S_i = S \cap V_i$ and $T_j = T \cap V_j$, we have

$$(38) \qquad |S|\,, |T| \ge \epsilon\,|V| \qquad \text{and} \qquad \sum_{i,j=1}^{k} \frac{|S_i|\,|T_j|}{|S|\,|T|}\,|d(S_i, T_j) - d(V_i, V_j)| > \epsilon \,.$$

Let $\mathcal{Q}$ be the refinement of $\mathcal{P}$ obtained by subdividing each V_i into three parts, S_i, T_i, and $W_i := V_i \setminus (S_i \cup T_i)$. Put differently, $\mathcal{Q}|_{V_i} = \{S_i, T_i, W_i\}$, where $\mathcal{Q}|_{V_i}$ denotes the partition of V_i that $\mathcal{Q}$ induces. We claim that $q(\mathcal{Q}) > q(\mathcal{P}) + \epsilon^4$, where q denotes the mean square density of a partition, that is,

$$q(\{Z_1, \dots, Z_r\}) = \sum_{i,j=1}^{r} \frac{|Z_i|\,|Z_j|}{|V|^2} d^2(Z_i, Z_j)$$

---

[13] This task is iterated in the proofs of the 3-graph regularity lemma; combined with the fact that $M(\epsilon) \ge \operatorname{twr}(\operatorname{poly}(1/\epsilon))$, this explains their wowzer-type bounds.

(where the sum is over ordered pairs (i,j)). Indeed,

$$\begin{aligned}
|V|^2\,(q(\mathcal{Q})-q(\mathcal{P})) &= \sum_{i,j=1}^{k}\Bigg(\sum_{\substack{U\in\mathcal{Q}|_{V_i},\\ U'\in\mathcal{Q}|_{V_j}}}|U||U'|d^2(U,U')-|V_i|\,|V_j|\,d^2(V_i,V_j)\Bigg)\\
&= \sum_{i,j=1}^{k}|V_i|\,|V_j|\Bigg(\sum_{\substack{U\in\mathcal{Q}|_{V_i},\\ U'\in\mathcal{Q}|_{V_j}}}\frac{|U||U'|}{|V_i|\,|V_j|}d^2(U,U')-d^2(V_i,V_j)\Bigg)\\
&= \sum_{i,j=1}^{k}|V_i|\,|V_j|\Bigg(\sum_{\substack{U\in\mathcal{Q}|_{V_i},\\ U'\in\mathcal{Q}|_{V_j}}}\frac{|U||U'|}{|V_i|\,|V_j|}(d(U,U')-d(V_i,V_j))^2\Bigg)\\
&\ge \sum_{i,j=1}^{k}|V_i|\,|V_j|\cdot\frac{|S_i|\,|T_j|}{|V_i|\,|V_j|}(d(S_i,T_j)-d(V_i,V_j))^2\\
&= |S|\,|T|\sum_{i,j=1}^{k}\frac{|S_i|\,|T_j|}{|S|\,|T|}(d(S_i,T_j)-d(V_i,V_j))^2\\
&\ge |S|\,|T|\Bigg(\sum_{i,j=1}^{k}\frac{|S_i|\,|T_j|}{|S|\,|T|}\,|d(S_i,T_j)-d(V_i,V_j)|\Bigg)^2 > \epsilon^4\,|V|^2\ ,
\end{aligned}$$

where in the first inequality we used the fact that $S_i\in\mathcal{Q}|_{V_i}$ and $T_j\in\mathcal{Q}|_{V_j}$, in the second inequality we used Jensen's inequality, and in the third inequality we used (38).

Suppose now that $\mathcal{P}$ is also equitable. We will use $\mathcal{Q}$ in order to construct an equitable refinement $\mathcal{P}'$ of $\mathcal{P}$ satisfying

$$q(\mathcal{P}')\ge q(\mathcal{P})+\epsilon^4/2\ . \tag{39}$$

Put $s=|V|/bk$ with $b=\lceil 8/\epsilon^4\rceil\in\mathbb{N}$. Let $\mathcal{P}'$ be the equipartition obtained from $\mathcal{P}$ by subdividing each $V_i\in\mathcal{P}$ into parts of size $\lfloor s\rfloor$ or $\lfloor s\rfloor+1$[14] in such a way that every part $U\in\mathcal{P}'|_{V_i}$ satisfies either $U\subseteq S_i$, $U\subseteq T_i$, or $U\subseteq W_i$ except for at most three parts U_i',U_i'',U_i''' in $\mathcal{P}'|_{V_i}$. Note that $\mathcal{P}'$ refines $\mathcal{P}$, but not $\mathcal{Q}$ (because of the sets U_i',U_i'',U_i'''). To prove (39), let $\mathcal{P}^*$ be an auxiliary partition obtained from $\mathcal{P}'$ by subdividing each U_i' into the three parts $\{U_i'\cap S_i,\ U_i'\cap T_i,\ U_i'\cap W_i\}$, and similarly for U_i'',U_i'''. Observe that $\mathcal{P}^*$ refines $\mathcal{Q}$. Furthermore,

$$q(\mathcal{P}^*)-q(\mathcal{P}')\le\sum_{i=1}^{k}\Big(\frac{|U_i'|\,|V|}{|V|^2}+\frac{|U_i''|\,|V|}{|V|^2}+\frac{|U_i'''|\,|V|}{|V|^2}\Big)\le k\frac{3(\lfloor s\rfloor+1)}{|V|}\le\frac{4}{b}\le\epsilon^4/2\ .$$

Since $\mathcal{P}^*$ refines $\mathcal{Q}$ we have $q(\mathcal{P}^*)\ge q(\mathcal{Q})$ by Jensen's inequality. Therefore,

$$q(\mathcal{P}')\ge q(\mathcal{P}^*)-\epsilon^4/2\ge q(\mathcal{Q})-\epsilon^4/2\ge q(\mathcal{P})+\epsilon^4/2\ ,$$

which proves (39). Note that $|\mathcal{P}'|\le bk\le(16/\epsilon^4)\,|\mathcal{P}|$.

Starting with the equipartition $\mathcal{P}_0$ given in the statement, we iteratively apply the above argument as long as the current partition $\mathcal{P}$ is not weak ϵ-regular. It follows from (39), together with the fact that the potential function q is at most 1,

---

[14] Simply divide $|V_i|$ by $\lfloor s\rfloor$; write $|V_i|=a\lfloor s\rfloor+m=(a-m)\lfloor s\rfloor+m(\lfloor s\rfloor+1)$ and observe that $m\le b\le\lfloor |V_i|/s\rfloor\le a$.

that a weak ϵ-regular equipartition is obtained after at most $2/\epsilon^4$ iterations. Since the order of the partition increases in each iteration by a factor of at most $16/\epsilon^4$, the order of the final partition increases by a factor of at most

$$(16/\epsilon^4)^{2/\epsilon^4} = (2/\epsilon)^{8/\epsilon^4} \leq 2^{16/\epsilon^5} .$$

This completes the proof. ☐

Appendix B. Properties of ϵ-regular graphs

For completeness, here we give proofs for the well-known properties used in Section 4. Recall that we say that (A, B) is an (ϵ, d)-*regular pair* if the bipartite graph between the vertex subsets A, B is ϵ-regular of density d. First, we have the following degree property.

Fact B.1. If (A, B) is an (ϵ, d)-regular pair, all vertices of B but at most $2\epsilon|B|$ have degree $(d \pm \epsilon)|A|$.

Proof. Otherwise there is a set $B' \subseteq B$ of at least $\epsilon|B|$ vertices whose degrees are, without loss of generality, greater than $(d+\epsilon)|A|$. Thus $d(A, B') > d + \epsilon$, a contradiction. ☐

Next is the so-called *slicing lemma*.

Fact B.2. Let $\alpha \geq \epsilon > 0$, and let (A, B) be an (ϵ, d)-regular pair. If $A' \subseteq A$, $B' \subseteq B$ are of size $|A'| \geq \alpha|A|$, $|B'| \geq \alpha|B|$, then the pair (A', B') is $(2\epsilon/\alpha,\ d \pm \epsilon)$-regular.

Proof. First, $|d(A', B') - d| \leq \epsilon$ is immediate as G is ϵ-regular and $\alpha \geq \epsilon$. Next, if $X \subseteq A'$ and $Y \subseteq B'$ satisfy $|X| \geq (\epsilon/\alpha)|A'|$ and $|Y| \geq (\epsilon/\alpha)|B'|$, then $|X| \geq \epsilon|A|$ and $|Y| \geq \epsilon|B|$. Since (A, B) is (ϵ, d)-regular we have $|d(X, Y) - d(A', B')| \leq |d(X, Y) - d| + |d - d(A', B')| \leq 2\epsilon \leq 2\epsilon/\alpha$. ☐

Finally, we have the following codegree property.

Fact B.3. Let the pairs $(A, C), (B, C)$ be (ϵ, d)-regular and (ϵ, d')-regular, respectively. Write $\text{codeg}(a, b)$ for the number of common neighbors of a, b in C, and put $\epsilon' = 6\epsilon/d$. All pairs $(a, b) \in A \times B$ but at most $\epsilon'|A||B|$ satisfy $\text{codeg}(a, b) = (dd' \pm \epsilon')|C|$.

Proof. Assume $d \geq 6\epsilon$, as otherwise there is nothing to prove. Let $a \in A$ with $e(a, C) = (d \pm \epsilon)|C|$ $(\geq \epsilon|C|)$, noting that by Fact B.1 there are at most $2\epsilon|A|$ vertices of A not satisfying this condition. By Fact B.2, the graph between B and the vertices of $e(a, C)$ is of density $d'' := d' \pm \epsilon$ and is $2\epsilon/d$-regular. Thus, again by Fact B.1, all vertices $b \in B$ but at most $(4\epsilon/d)|B|$ satisfy $\text{codeg}(a, b) = (d'' \pm 2\epsilon/d)e(a, C) = (d' \pm 3\epsilon/d)(d \pm \epsilon)|C| = (dd' \pm \epsilon')|C|$. Therefore, the number of pairs (a, b) not satisfying $\text{codeg}(a, b) = (dd' \pm \epsilon')|C|$ is at most $2\epsilon|A| \cdot |B| + |A| \cdot (4\epsilon/d)|B| \leq \epsilon'|A||B|$, as needed. ☐

Acknowledgment

The first author would like to thank V. Rödl for many helpful discussions regarding this work. In particular, Proposition 1.2 was obtained in joint discussions with him.

References

[1] Noga Alon, Eldar Fischer, Michael Krivelevich, and Mario Szegedy, *Efficient testing of large graphs*, Combinatorica **20** (2000), no. 4, 451–476, DOI 10.1007/s004930070001. MR1804820

[2] David Conlon and Jacob Fox, *Bounds for graph regularity and removal lemmas*, Geom. Funct. Anal. **22** (2012), no. 5, 1191–1256, DOI 10.1007/s00039-012-0171-x. MR2989432

[3] David Conlon and Jacob Fox, *Graph removal lemmas*, Surveys in combinatorics 2013, London Math. Soc. Lecture Note Ser., vol. 409, Cambridge Univ. Press, Cambridge, 2013, pp. 1–49. MR3156927

[4] Richard A. Duke, Hanno Lefmann, and Vojtěch Rödl, *A fast approximation algorithm for computing the frequencies of subgraphs in a given graph*, SIAM J. Comput. **24** (1995), no. 3, 598–620, DOI 10.1137/S0097539793247634. MR1333857

[5] Jacob Fox, *A new proof of the graph removal lemma*, Ann. of Math. (2) **174** (2011), no. 1, 561–579, DOI 10.4007/annals.2011.174.1.17. MR2811609

[6] Peter Frankl and Vojtěch Rödl, *Extremal problems on set systems*, Random Structures Algorithms **20** (2002), no. 2, 131–164, DOI 10.1002/rsa.10017.abs. MR1884430

[7] Alan Frieze and Ravi Kannan, *Quick approximation to matrices and applications*, Combinatorica **19** (1999), no. 2, 175–220, DOI 10.1007/s004930050052. MR1723039

[8] W. T. Gowers, *Lower bounds of tower type for Szemerédi's uniformity lemma*, Geom. Funct. Anal. **7** (1997), no. 2, 322–337, DOI 10.1007/PL00001621. MR1445389

[9] W. T. Gowers, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. of Math. (2) **166** (2007), no. 3, 897–946, DOI 10.4007/annals.2007.166.897. MR2373376

[10] Wassily Hoeffding, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc. **58** (1963), 13–30. MR0144363

[11] Robert M. Gray, *Entropy and information theory*, 2nd ed., Springer, New York, 2011. MR3134681

[12] J. Komlós and M. Simonovits, *Szemerédi's regularity lemma and its applications in graph theory*, Combinatorics, Paul Erdős is eighty, Vol. 2 (Keszthely, 1993), Bolyai Soc. Math. Stud., vol. 2, János Bolyai Math. Soc., Budapest, 1996, pp. 295–352. MR1395865

[13] László Lovász and Balázs Szegedy, *Szemerédi's lemma for the analyst*, Geom. Funct. Anal. **17** (2007), no. 1, 252–270, DOI 10.1007/s00039-007-0599-6. MR2306658

[14] Guy Moshkovitz and Asaf Shapira, *A short proof of Gowers' lower bound for the regularity lemma*, Combinatorica **36** (2016), no. 2, 187–194, DOI 10.1007/s00493-014-3166-4. MR3516883

[15] Brendan Nagle, Vojtěch Rödl, and Mathias Schacht, *The counting lemma for regular k-uniform hypergraphs*, Random Structures Algorithms **28** (2006), no. 2, 113–179, DOI 10.1002/rsa.20117. MR2198495

[16] M. S. Pinsker, *Information and information stability of random variables and processes*, Translated and edited by Amiel Feinstein, Holden-Day, Inc., San Francisco, Calif.-London-Amsterdam, 1964. MR0213190

[17] Vojtěch Rödl and Mathias Schacht, *Regular partitions of hypergraphs: regularity lemmas*, Combin. Probab. Comput. **16** (2007), no. 6, 833–885. MR2351688

[18] Vojtěch Rödl and Mathias Schacht, *Regularity lemmas for graphs*, Fete of combinatorics and computer science, Bolyai Soc. Math. Stud., vol. 20, János Bolyai Math. Soc., Budapest, 2010, pp. 287–325, DOI 10.1007/978-3-642-13580-4_11. MR2798368

[19] Vojtěch Rödl and Jozef Skokan, *Regularity lemma for k-uniform hypergraphs*, Random Structures Algorithms **25** (2004), no. 1, 1–42, DOI 10.1002/rsa.20017. MR2069663

[20] Alexander Scott, *Szemerédi's regularity lemma for matrices and sparse graphs*, Combin. Probab. Comput. **20** (2011), no. 3, 455–466, DOI 10.1017/S0963548310000490. MR2784637

[21] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Colloq. Math. Soc. János Bolyai, vol. 18, North-Holland, Amsterdam-New York, 1978, pp. 939–945. MR519318 (1978), 399–401.

[22] Endre Szemerédi, *Regular partitions of graphs* (English, with French summary), Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401. MR540024

[23] Terence Tao, *A variant of the hypergraph removal lemma*, J. Combin. Theory Ser. A **113** (2006), no. 7, 1257–1280, DOI 10.1016/j.jcta.2005.11.006. MR2259060

[24] Terence Tao, *Szemerédi's regularity lemma revisited*, Contrib. Discrete Math. **1** (2006), no. 1, 8–28. MR2212136

School of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel
Email address: `guymosko@tau.ac.il`

School of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel
Email address: `asafico@tau.ac.il`