REGULARITY LEMMAS FOR GRAPHS

VOJTĚCH RÖDL AND MATHIAS SCHACHT

Abstract. Szemerédi's regularity lemma proved to be a fundamental result in modern graph theory. It had a number of important applications and is a widely used tool in extremal combinatorics. For some applications variants of the regularity lemma were considered. Here we discuss several of those variants and their relation to each other.

1. Introduction

Szemerédi's regularity lemma is one of the most important tools in extremal graph theory. It has many applications not only in graph theory, but also in combinatorial number theory, discrete geometry, and theoretical computer science. The first form of the lemma was invented by Szemerédi [47] as a tool for the resolution of a famous conjecture of Erdős and Turán [9] stating that any sequence of integers with positive upper density must contain arithmetic progressions of any finite length.

The regularity lemma roughly states that every graph may be approximated by a union of induced random-like (quasi-random) bipartite subgraphs. Since the quasi-randomness brings important additional information, the regularity lemma proved to be a useful tool. The regularity lemma allows one to import probabilistic intuition to deterministic problems. Moreover, there are many applications where the original problem did not suggest a probabilistic approach.

Motivated especially by questions from computer science, several other variants of Szemerédi's regularity lemma were considered. In Section 2 we focus mainly on the lemmas proved by Frieze and Kannan [12] and by Alon, Fischer, Krivelevich, and M. Szegedy [2]. We show how these lemmas compare to Szemerédi's original lemma and how they relate to some other variants. Most proofs stated here appeared earlier in the literature and here we just give an overview. A thorough discussion of the connections of those regularity lemmas, from an analytical and geometrical perspective was given recently by Lovász and B. Szegedy in [30]. In Section 3 we discuss the so-called *counting lemmas* and the *removal lemma* and its generalizations. We close with a brief discussion of the *limit approach* of Lovász and B. Szegedy and its relation to the regularity lemmas from Section 4.

There are several surveys devoted to Szemerédi regularity lemma and its applications. The reader is recommended to consult Komlós and Simonovits [26] and Komlós, Shoukoufandeh, Simonovits, and Szemeredi [25], where many applications of the regularity lemma are discussed.

Another line of research, which we will not discuss here, concerns sparse versions of the regularity lemma. Since Szemeredi's lemma is mainly suited for addressing

---

Date: August 11, 2009.

The first author was supported by NSF grant DMS 0800070.

problems involving "dense" graphs, that is graphs with at least $\Omega(|V|^2)$ edges, it is natural to ask for similar statements that would apply to "sparse graphs", i.e., graphs with $o(|V|^2)$ edges. It turns out that a regularity lemma applicable to certain classes of sparse graphs can be proved [22, 34] (see also [1]). Such a lemma was first applied by Kohayakawa and his collaborators to address extremal and Ramsey-type problems for subgraphs of random graphs (see, e.g., [19, 20, 21]). Here we will not further discuss this line of research and we refer the interested reader to the surveys [15, 23, 31] and the references therein.

2. Regularity lemmas

In this section we discuss several regularity lemmas for graphs. We start our discussion with the regularity lemma of Frieze and Kannan [12] in the next section. In Section 2.2 we show how Szemerédi's regularity lemma [48] can be deduced from the weaker lemma of Frieze and Kannan by iterated applications. In Section 2.3 we discuss the (ε, r)-regularity lemma, whose analog for 3-uniform hypergraphs was introduced by Frankl and Rödl [11]. We continue in Section 2.4 with the regularity lemma of Alon, Fischer, Krivelevich, and M. Szegedy [2], which can be viewed as an iterated version of Szemerédi's regularity lemma. In Section 2.5 we introduce the regular approximation lemma whose hypergraph variant was developed in [37]. Finally, in Section 2.6 we briefly discuss the original regularity lemma of Szemerédi [47] for bipartite graphs and a multipartite version of it from [8].

2.1. The regularity lemma of Frieze and Kannan.

The following variant of Szemerédi's regularity lemma was introduced by Frieze and Kannan [12] for the design of an efficient approximation algorithm for the MAX-CUT problem in dense graphs.

Theorem 1. *For every $\varepsilon > 0$ and every $t_0 \in \mathbb{N}$ there exist $T_{\mathrm{FK}} = T_{\mathrm{FK}}(\varepsilon, t_0)$ and n_0 such that for every graph $G = (V, E)$ with at least $|V| = n \geq n_0$ vertices the following holds. There exists a partition $V_1 \dot\cup \ldots \dot\cup V_t = V$ such that*

- *(i) $t_0 \leq t \leq T_{\mathrm{FK}}$,*
- *(ii) $|V_1| \leq \cdots \leq |V_t| \leq |V_1| + 1$, and*
- *(iii) for every $U \subseteq V$*

$$\left| e(U) - \sum_{i=1}^{t-1} \sum_{j=i+1}^{t} d(V_i, V_j)|U \cap V_i||U \cap V_j| \right| \leq \varepsilon n^2 \,, \tag{1}$$

where $e(U)$ denotes the number of edges contained in U and $d(V_i, V_j) = e(V_i, V_j)/(|V_i||V_j|)$ denotes the density of the bipartite graph induced on V_i and V_j.

Definition 2. *A partition satisfying property (ii) of Theorem 1 will be called* equitable *and a partition satisfying all three properties (i)-(iii) will be referred to as* $(\varepsilon, t_0, T_{\mathrm{FK}})$-FK-partition. *Sometimes we may omit t_0 and T_{FK} and simply refer to such a partition as* ε-FK-partition.

The essential properties of the partition provided by Theorem 1 are property (i) and (iii). Property (i) bounds the number of partition classes by a constant independent of G and n and, roughly speaking, property (iii) asserts that the number of edges of any large set U can be fairly well approximated by the densities $d(V_i, V_j)$ given by the partition $V_1 \dot\cup \ldots \dot\cup V_t = V$. More precisely, $e(U) \approx e(U')$ for any choice of U and U' satisfying for example $|U \cap V_i| \approx |U' \cap V_i|$ for all $i \in [t]$. Moreover, we note that conclusion (iii) can be replaced by the following:

(iii') for all (not necessarily disjoint) sets $U, W \subseteq V$

$$\left| e(U,W) - \sum_{i=1}^{t} \sum_{j\in[t]\setminus\{i\}} d(V_i,V_j)|U\cap V_i||W\cap V_j| \right| \le 4\varepsilon n^2\,, \tag{2}$$

where edges contained in $U \cap W$ are counted twice in $e(U,W)$.

Indeed, if (iii) holds, then we infer (iii') from the identity

$$e(U,W) = e(U\cup W) - e(U\setminus W) - e(W\setminus U) + e(U\cap W)\,.$$

The proof of Theorem 1 relies on the *index* of a partition, a concept which was first introduced and used by Szemerédi.

Definition 3. *For a partition $\mathcal{P} = (V_1,\dots,V_t)$ of the vertex sets of a graph $G = (V,E)$, i.e., $V_1\dot\cup\dots\dot\cup V_t = V$ we define the* index *of $\mathcal{P}$ by*

$$\operatorname{ind}(\mathcal{P}) = \frac{1}{\binom{|V|}{2}} \sum_{i=1}^{t-1}\sum_{j=i+1}^{t} d^2(V_i,V_j)|V_i||V_j|\,.$$

Note that it follows directly from the definition of the index that for any partition $\mathcal{P}$ we have

$$0 \le \operatorname{ind}(\mathcal{P}) \le 1\,.$$

For the proof of Theorem 1 we will use the following consequence of the Cauchy-Schwarz inequality.

Lemma 4. *Let $1 \le M < N$, let $\sigma_1,\dots,\sigma_N$ be positive and $d_1,\dots,d_N$, and d be reals. If $\sum_{i=1}^N \sigma_i = 1$ and $d = \sum_{i=1}^N d_i\sigma_i$ then*

$$\sum_{i=1}^{N} d_i^2\sigma_i \ge d^2 + \left(d - \frac{\sum_{i=1}^M d_i\sigma_i}{\sum_{i=1}^M \sigma_i}\right)^2 \frac{\sum_{i=1}^M \sigma_i}{1-\sum_{i=1}^M \sigma_i}\,.$$

For completeness we include the short proof of Lemma 4.

Proof. For $M = 1$ and $N = 2$ the statement follows from the identity

$$\hat d_1^2\hat\sigma_1 + \hat d_2^2\hat\sigma_2 = \hat d^2 + (\hat d - \hat d_1)^2\frac{\hat\sigma_1}{\hat\sigma_2}\,. \tag{3}$$

which is valid for positive $\hat\sigma_1, \hat\sigma_2$ with $\hat\sigma_1 + \hat\sigma_2 = 1$ and $\hat d = \hat d_1\hat\sigma_1 + \hat d_2\hat\sigma_2$.

For general $1 \le M < N$ we infer from the Cauchy-Schwarz inequality applied twice in the form $(\sum d_i\sigma_i)^2 \le \sum d_i^2\sigma_i\sum\sigma_i$

$$\begin{aligned}
\sum_{i=1}^N d_i^2\sigma_i &= \sum_{i=1}^M d_i^2\sigma_i + \sum_{i=M+1}^N d_i^2\sigma_i\\
&\ge \frac{\left(\sum_{i=1}^M d_i\sigma_i\right)^2}{\sum_{i=1}^M\sigma_i} + \frac{\left(\sum_{i=M+1}^N d_i\sigma_i\right)^2}{\sum_{i=M+1}^N\sigma_i}\\
&= \left(\frac{\sum_{i=1}^M d_i\sigma_i}{\sum_{i=1}^M\sigma_i}\right)^2\sum_{i=1}^M\sigma_i + \left(\frac{\sum_{i=M+1}^N d_i\sigma_i}{\sum_{i=M+1}^N\sigma_i}\right)^2\sum_{i=M+1}^N\sigma_i\,.
\end{aligned}$$

Setting

$$\hat\sigma_1=\sum_{i=1}^{M}\sigma_i,\quad \hat\sigma_2=\sum_{i=M+1}^{N}\sigma_i,$$

$$\hat d_1=\frac{\sum_{i=1}^{M}d_i\sigma_i}{\sum_{i=1}^{M}\sigma_i},\quad \hat d_2=\frac{\sum_{i=M+1}^{N}d_i\sigma_i}{\sum_{i=M+1}^{N}\sigma_i},\quad\text{and}\quad \hat d=\hat d_1\hat\sigma_1+\hat d_2\hat\sigma_2$$

we have $\hat d=\sum_{i=1}^{N}d_i\sigma_i=d$ and from (3) we infer

$$\sum_{i=1}^{N}d_i^2\sigma_i\geq\left(\sum_{i=1}^{N}d_i\sigma_i\right)^2+\left(\sum_{i=1}^{N}d_i\sigma_i-\frac{\sum_{i=1}^{M}d_i\sigma_i}{\sum_{i=1}^{M}\sigma_i}\right)^2\frac{\sum_{i=1}^{M}\sigma_i}{\sum_{i=M+1}^{N}\sigma_i},$$

which is what we claimed. □

After those preparations we prove Theorem 1.

Proof of Theorem 1. The proof is based on the following idea already present in the original work of Szemerédi. Starting with an arbitrary equitable vertex partition $\mathcal{P}_0$ with t_0 classes, we consider a sequence of partitions $\mathcal{P}_0,\mathcal{P}_1,\dots$ such that $\mathcal{P}_j$ always satisfies properties (i) and (ii). As soon as $\mathcal{P}_j$ also satisfies (iii) we can stop. On the other hand, if $\mathcal{P}_j$ does not satisfy (iii) we will show that there exists a partition $\mathcal{P}_{j+1}$ whose index increased by $\varepsilon^2/2$. Since $\operatorname{ind}(\mathcal{P})\leq 1$ for any partition $\mathcal{P}$, we infer that after at most $2/\varepsilon^2$ steps this procedure must end with a partition satisfying properties (i), (ii), and (iii) of the theorem.

So suppose $\mathcal{P}_j=\mathcal{P}=(V_1,\dots,V_t)$ is a partition of V which satisfies (i) and (ii), but there exists a set $U\subseteq V$ such that (1) fails. We are going to construct a partition $\mathcal{R}=\mathcal{P}_{j+1}$ satisfying

$$\operatorname{ind}(\mathcal{R})\geq\operatorname{ind}(\mathcal{P})+\varepsilon^2/2\,. \tag{4}$$

For that set

$$U_i=V_i\cap U\quad\text{and}\quad \bar U_i=V_i\setminus U\,.$$

We define a new partition $\mathcal{Q}$ by replacing every vertex class V_i by U_i and $\bar U_i$

$$\mathcal{Q}=(U_1,\bar U_1\dots,U_t,\bar U_t)\,.$$

Next we show that the index of $\mathcal{Q}$ increased by ε^2 compared to $\operatorname{ind}(\mathcal{P})$. For every $1\leq i<j\leq t$ we set

$$\varepsilon_{ij}=d(U_i,U_j)-d(V_i,V_j)\,.$$

Since we may assume $t\geq t_0\geq 1/\varepsilon$, which yields $\sum_{i=1}^{t}e(V_i)\leq\varepsilon n^2/2$, we infer from the assumption that (1) fails, that

$$\Big|\sum_{i<j}\varepsilon_{ij}|U_i||U_j|\Big|\geq\varepsilon n^2-\sum_{i=1}^{t}e(U_i)\geq\varepsilon n^2-\sum_{i=1}^{t}e(V_i)\geq\frac{\varepsilon}{2}n^2\,, \tag{5}$$

Since $V_i=U_i\dot\cup\bar U_i$ for every $i\in[t]$ we obtain

$$\begin{aligned}d(V_i,V_j)|V_i||V_j|=d(U_i,U_j)|U_i||U_j|&+d(\bar U_i,U_j)|\bar U_i||U_j|\\&+d(U_i,\bar U_j)|U_i||\bar U_j|+d(\bar U_i,\bar U_j)|\bar U_i||\bar U_j|\end{aligned}$$

and

$$|V_i||V_j|=|U_i||U_j|+|\bar U_i||U_j|+|U_i||\bar U_j|+|\bar U_i||\bar U_j|\,.$$

Combining those identities with Lemma 4, we obtain

$$d^2(U_i,U_j)|U_i||U_j| + d^2(\bar{U}_i,U_j)|\bar{U}_i||U_j| + d^2(U_i,\bar{U}_j)|U_i||\bar{U}_j| + d^2(\bar{U}_i,\bar{U}_j)|\bar{U}_i||\bar{U}_j|$$
$$\geq d^2(V_i,V_j)|V_i||V_j| + \varepsilon_{ij}^2\left(\frac{|U_i||U_j|}{1-\frac{|U_i||U_j|}{|V_i||V_j|}}\right) \geq d^2(V_i,V_j)|V_i||V_j| + \varepsilon_{ij}^2|U_i||U_j|\,.$$

Summing over all $1 \leq i < j \leq t$ we obtain

$$\begin{aligned}\operatorname{ind}(\mathcal{Q}) &\geq \operatorname{ind}(\mathcal{P}) + \frac{1}{\binom{n}{2}}\sum_{i<j}\varepsilon_{ij}^2|U_i||U_j| \\ &\geq \operatorname{ind}(\mathcal{P}) + \frac{\big(\sum_{i<j}\varepsilon_{ij}|U_i||U_j|\big)^2}{\binom{n}{2}\sum_{i<j}|U_i||U_j|} \overset{(5)}{\geq} \operatorname{ind}(\mathcal{P}) + \frac{(\varepsilon n^2/2)^2}{\binom{n}{2}\binom{n}{2}} \geq \operatorname{ind}(\mathcal{P}) + \varepsilon^2. \quad (6)\end{aligned}$$

We now find an equitable partition $\mathcal{R}$ which is a refinement of $\mathcal{P}$ (and almost a refinement of $\mathcal{Q}$) for which (4) holds. For that subdivide each vertex class V_i of $\mathcal{P}$ into sets $W_{i,a}$ of size $\lfloor\varepsilon^2 n/(5t)\rfloor$ or $\lfloor\varepsilon^2 n/(5t)\rfloor + 1$ in such a way that for all but at most one of these sets either $W_{i,a} \subseteq U_i$ or $W_{i,a} \subseteq \bar{U}_i$ holds. For every $i \in [t]$ let $W_{i,0}$ denote the exceptional set if it exists and let $W_{i,0}$ be arbitrary otherwise. Let $\mathcal{R}$ be the resulting partition. Moreover, we consider the partition $\mathcal{R}^*$ which is a refinement of $\mathcal{R}$ obtained by replacing $W_{i,0}$ by possibly two classes $U_i \cap W_{i,0}$ and $\bar{U}_i \cap W_{i,0}$. Since the contribution of the index of $\mathcal{R}$ and $\mathcal{R}^*$ may differ only on pairs with at least one vertex in $W_{i,0}$ for some $i \in [t]$ and since $|W_{i,0}| \leq \lfloor\varepsilon^2 n/(5t)\rfloor + 1$ for every $i \in [t]$ we infer that

$$\operatorname{ind}(\mathcal{R}^*) - \operatorname{ind}(\mathcal{R}) \leq \binom{n}{2}^{-1}\sum_{i=1}^{t}\left(\frac{\varepsilon^2 n}{5t}+1\right)n \leq \frac{\varepsilon^2}{2}\,.$$

for sufficiently large n. Furthermore, since $\mathcal{R}^*$ is a refinement of $\mathcal{Q}$ it follows from the Cauchy-Schwarz inequality that $\operatorname{ind}(\mathcal{Q}) \leq \operatorname{ind}(\mathcal{R}^*)$ and, consequently,

$$\operatorname{ind}(\mathcal{R}) \geq \operatorname{ind}(\mathcal{R}^*) - \frac{\varepsilon^2}{2} \geq \operatorname{ind}(\mathcal{Q}) - \frac{\varepsilon^2}{2} \overset{(6)}{\geq} \operatorname{ind}(\mathcal{P}) + \frac{\varepsilon^2}{2}\,,$$

which concludes the proof of the theorem. □

The proof of Theorem 1 shows that choosing

$$T_{\mathrm{FK}}(\varepsilon,t_0) = \max\{t_0, 1/\varepsilon\}\cdot(6/\varepsilon^2)^{2/\varepsilon^2} = t_0 2^{\mathrm{poly}(1/\varepsilon)}$$

suffices. In fact, in each refinement step we split the vertex classes V_i into at most $\lfloor 5/\varepsilon^2 + 1\rfloor \leq 6/\varepsilon^2$ classes $W_{i,a}$, when we construct $\mathcal{R}$. Hence, each time property (*iii*) fails the number of vertex classes of the new partition increases by a factor of $6/\varepsilon^2$ and in total there are at most $2/\varepsilon^2$ iterations.

On the other hand, it was shown by Lovász and B. Szegedy [30] that for every $0 < \varepsilon \leq 1/3$ there are graphs for which every partition into t classes satisfying property (*iii*) of Theorem 1 requires $t \geq 2^{1/(8\varepsilon)}/4$ and, hence, $t \gg 1/\varepsilon$. As a consequence Theorem 1 does not allow to obtain useful bounds for $e(U\cap V_i, U\cap V_j)$, since for such a graph $\varepsilon n^2 \gg n^2/t^2 = |V_i||V_j|$. Property (*iii*) of Theorem 1 only implies $e(U\cap V_i, U\cap V_j) \approx d(V_i,V_j)|U\cap V_i||U\cap V_j|$ on average over all pairs $i<j$ for every "large" set U. However, Szemerédi's regularity lemma (which was proved long before Theorem 1) allows to control $e(U\cap V_i, U\cap V_j)$ for most $i<j$. The price

of this is, however, a significantly larger upper bound for the number of partition classes t.

2.2. **Szemerédi's regularity lemma.** In this section we show how Szemerédi's regularity lemma from [48] can be obtained from Theorem 1 by iterated applications. For that we consider the following simple corollary of Theorem 1, which was first considered by Tao [49].

Corollary 5. *For all ν, $\varepsilon > 0$, every function $\delta\colon \mathbb{N} \to (0,1]$, and every $t_0 \in \mathbb{N}$ there exist $T_0 = T_0(\nu, \varepsilon, \delta(\cdot), t_0)$ and n_0 such that for every graph $G = (V, E)$ with at least $|V| = n \geq n_0$ vertices the following holds. There exists a vertex partition $\mathcal{P} = (V_i)_{i\in[t]}$ with $V_1 \dot\cup \ldots \dot\cup V_t = V$ and a refinement $\mathcal{Q} = (W_{i,j})_{i\in[t],j\in[s]}$ with $W_{i,1} \dot\cup \ldots \dot\cup W_{i,s} = V_i$ for every $i \in [t]$ such that*

(i) $\mathcal{P}$ is an (ε, t_0, T_0)-FK-partition,
(ii) $\mathcal{Q}$ is a $(\delta(t), t_0, T_0)$-FK-partition, and
(iii) $\mathrm{ind}(\mathcal{Q}) \leq \mathrm{ind}(\mathcal{P}) + \nu$.

Before we deduce Corollary 5 from Theorem 1, we discuss property (*iii*). Roughly speaking, if two refining partitions $\mathcal{P}$ and $\mathcal{Q}$ satisfy property (*iii*), then this implies that $d(W_{i,a}, W_{j,b})$ and $d(V_i, V_j)$ are "relatively close" for "most" choices of $i < j$ and $a, b \in [s]$. More precisely, we have the following, which was already observed by Alon, Fischer, Krivelevich, and M. Szegedy [2].

Lemma 6. *Let γ, $\nu > 0$, let $G = (V, E)$ be a graph with n vertices, and for some positive integers t and s let $\mathcal{P} = (V_i)_{i\in[t]}$ with $V_1 \dot\cup \ldots \dot\cup V_t = V$ be a vertex partition and let $\mathcal{Q} = (W_{i,j})_{i\in[t],j\in[s]}$ be a refinement with $W_{i,1} \dot\cup \ldots \dot\cup W_{i,s} = V_i$ for every $i \in [t]$. If $\mathrm{ind}(\mathcal{Q}) \leq \mathrm{ind}(\mathcal{P}) + \nu$, then*

$$\sum_{1\leq i<j\leq t} \sum_{a,b\in[s]} \Big\{ |W_{i,a}||W_{j,b}|\colon\ |d(W_{i,a}, W_{j,b}) - d(V_i, V_j)| \geq \gamma \Big\} \leq \frac{\nu}{\gamma^2} n^2 \,.$$

Proof. For $1 \leq i < j \leq t$ let $A^+_{ij} = \{(a, b) \in [s] \times [s]\colon\ d(W_{i,a}, W_{j,b}) - d(V_i, V_j) \geq \gamma\}$. Since

$$\begin{aligned} d(V_i, V_j)|V_i||V_j| &= \sum_{a,b\in[s]} d(W_{i,a}, W_{j,b})|W_{i,a}||W_{j,b}| \\ &= \sum_{(a,b)\in A^+_{ij}} d(W_{i,a}, W_{j,b})|W_{i,a}||W_{j,b}| + \sum_{(a,b)\notin A^+_{ij}} d(W_{i,a}, W_{j,b})|W_{i,a}||W_{j,b}| \,, \end{aligned}$$

we obtain from the defect form of Cauchy-Schwarz (Lemma 4), that

$$\sum_{a,b\in[s]} d^2(W_{i,a}, W_{j,b})|W_{i,a}||W_{j,b}| \geq d^2(V_i, V_j)|V_i||V_j| + \gamma^2 \sum_{(a,b)\in A^+_{ij}} |W_{i,a}||W_{j,b}| \,.$$

Summing over all $1 \leq i < j \leq t$ we get

$$\mathrm{ind}(\mathcal{Q}) \geq \mathrm{ind}(\mathcal{P}) + \frac{\gamma^2}{\binom{n}{2}} \sum_{1\leq i<j\leq t} \sum_{(a,b)\in A^+_{ij}} |W_{i,a}||W_{j,b}| \,.$$

Since, by assumption $\mathrm{ind}(\mathcal{Q}) \leq \mathrm{ind}(\mathcal{P}) + \nu$, we have

$$\sum_{1\leq i<j\leq t} \sum_{(a,b)\in A^+_{ij}} |W_{i,a}||W_{j,b}| \leq \frac{\nu}{\gamma^2} \binom{n}{2} \leq \frac{\nu n^2}{2\gamma^2} \,.$$

Repeating the argument with the appropriate definition of A^-_{ij} yields the claim. □

Proof of Corollary 5. For the proof of the corollary we simply iterate Theorem 1. Without loss of generality we may assume that $\delta(t) \leq \varepsilon$ for every $t \in \mathbb{N}$. For given ν, ε, $\delta(\cdot)$, and t_0, we apply Theorem 1 and obtain an (ε, t_0, T_0)-FK-partition $\mathcal{P}$ with t classes. Since in the proof of Theorem 1 the initial partition was an arbitrary equitable partition, we infer that after another application of Theorem 1 with $\delta(t)$ (in place of ε) and t_0 we obtain an equitable refinement $\mathcal{Q}$ of $\mathcal{P}$ which is a $(\delta(t), t_0, T_0)$-FK-partition with st classes. In other words, $\mathcal{P}$ and $\mathcal{Q}$ satisfy properties (i) and (ii) of Corollary 5 and if (iii) also holds, then we are done. On the other hand, if (iii) fails, then we replace $\mathcal{P}$ by $\mathcal{Q}$ and iterate, i.e., we apply Theorem 1 with $\delta(ts)$ (in place of ε) and $t_0 = ts$ to obtain an equitable refinement $\mathcal{Q}'$ of $\mathcal{P}' = \mathcal{Q}$. Since we only iterate as long as (iii) of Corollary 5 fails and since ν is fixed throughout the proof, this procedure must end after at most $1/\nu$ iterations. Therefore the upper bound T_0 on the number of classes is in fact independent of G and n and can be given by a recursive formula depending on ν, ε, $\delta(\cdot)$, and t_0. □

We now show that Corollary 5 applied with the right choice of parameters yields the following theorem, which is essentially Szemerédi's regularity lemma from [48].

Theorem 7. *For every $\varepsilon > 0$ and every $t_0 \in \mathbb{N}$ there exist $T_{\mathrm{Sz}} = T_{\mathrm{Sz}}(\varepsilon, t_0)$ and n_0 such that for every graph $G = (V, E)$ with at least $|V| = n \geq n_0$ vertices the following holds. There exists a partition $V_1 \dot\cup \ldots \dot\cup V_t = V$ such that*

- (i) *$t_0 \leq t \leq T_{\mathrm{Sz}}$,*
- (ii) *$|V_1| \leq \cdots \leq |V_t| \leq |V_1| + 1$, and*
- (iii) *for all but at most εt^2 pairs (V_i, V_j) with $i < j$ we have that for all subsets $U_i \subseteq V_i$ and $U_j \subseteq V_j$*

$$\left|e(U_i, U_j) - d(V_i, V_j)|U_i||U_j|\right| \leq \varepsilon |V_i||V_j|\,. \tag{7}$$

We note that the usual statement of Szemerédi's regularity lemma is slightly different from the one above. Usually $\varepsilon|V_i||V_j|$ on the right-hand side of (7) is replaced by $\varepsilon|U_i||U_j|$ and for (iii) it is assumed that $|U_i| \geq \varepsilon|V_i|$ and $|U_j| \geq \varepsilon|V_j|$. However, applying Theorem 7 with $\varepsilon' = \varepsilon^3$ would yield a partition with comparable regular properties.

Definition 8. *Pairs (V_i, V_j) for which (7) holds for every $U_i \subseteq V_i$ and $U_j \subseteq V_j$ are called* ε-regular. *Partitions satisfying all three properties (i)-(iii) of Theorem 7, we will refer to as* $(\varepsilon, t_0, T_{\mathrm{Sz}})$-Szemerédi-partition. *Again we may sometimes omit t_0 and T_{Sz} and simply refer to such partitions as* ε-Szemerédi-partitions.

Below we deduce Theorem 7 from Corollary 5 and Lemma 6.

Proof of Theorem 7. For given $\varepsilon > 0$ and t_0, we apply Corollary 5 with

$$\nu' = \frac{\varepsilon^4}{36^2}, \quad \varepsilon' = 1, \quad \delta'(t) = \frac{\varepsilon}{36t^2}, \quad \text{and} \quad t_0' = t_0$$

and obtain constants T_0' and n_0' which define $T_{\mathrm{Sz}} = T_0'$ and $n_0 = n_0'$. (We remark that the choice for ε' has no bearing for the proof and therefore we set it equal to 1.) For a given graph $G = (V, E)$ with n vertices Corollary 5 yields two partitions $\mathcal{P} = (V_i)_{i\in[t]}$ and $\mathcal{Q} = (W_{i,j})_{i\in[t],j\in[s]}$ satisfying properties (i)-(iii) of Corollary 5. We will show that, in fact, the coarser partition $\mathcal{P}$ also satisfies properties (i)-(iii) of Theorem 7. Since $\mathcal{P}$ is an $(\varepsilon', t_0', T_0')$-FK-partition by our choice of $t_0' = t_0$ and

$T_{\mathrm{Sz}} = T_0'$ the partition $\mathcal{P}$ obviously satisfies properties (i) and (ii) of Theorem 7 and we only have to verify property (iii).

For that we consider for every $1 \le i < j \le t$ the set

$$A_{ij} = \big\{(a,b) \in [s] \times [s] \colon\ |d(W_{i,a}, W_{j,b}) - d(V_i, V_j)| \ge \varepsilon/6\big\}$$

and we let

$$I = \Big\{\{i,j\} \colon\ 1 \le i < j \le t \text{ such that } \textstyle\sum_{(a,b)\in A_{ij}} |W_{i,a}||W_{j,b}| \ge \varepsilon |V_i||V_j|/6\Big\}\,.$$

We will first show that $|I| \le \varepsilon t^2$ and then we will verify that if $\{i,j\} \notin I$, then (7) holds. Indeed, due to property (iii) of Corollary 5 we have $\mathrm{ind}(\mathcal{Q}) \le \mathrm{ind}(\mathcal{P}) + \nu'$ and, consequently, it follows from Lemma 6 (applied with $\nu' = \varepsilon^4/36^2$ and $\gamma' = \varepsilon/6$) that

$$\frac{\varepsilon^2 n^2}{36} \ge \sum_{i<j} \sum_{(a,b)\in A_{ij}} |W_{i,a}||W_{j,b}| \ge \sum_{\{i,j\}\in I} \sum_{(a,b)\in A_{ij}} |W_{i,a}||W_{j,b}| \ge \frac{\varepsilon}{6} \sum_{\{i,j\}\in I} |V_i||V_j|\,.$$

Moreover, since $|V_i| \ge \lfloor n/t \rfloor \ge n/(2t)$ for every $i \in [t]$ we have $\varepsilon n^2/6 \ge |I| n^2/(4t^2)$ and, consequently,

$$|I| \le \frac{2}{3}\varepsilon t^2 < \varepsilon t^2\,. \tag{8}$$

Next we will show that if $\{i,j\} \notin I$ then the pair (V_i, V_j) is ε-regular, i.e., we show that (7) holds for every $U_i \subseteq V_i$ and $U_j \subseteq V_j$. For given sets $U_i \subseteq V_i$ and $U_j \subseteq V_j$ and a, $b \in [s]$ we set

$$U_{i,a} = U_i \cap W_{i,a} \quad \text{and} \quad U_{j,b} = U_j \cap W_{j,b}$$

and have

$$e(U_i, U_j) = \sum_{a,b\in[s]} e(U_{i,a}, U_{j,b})\,.$$

Appealing to the fact that $\mathcal{Q}$ is a $(\delta'(t), t_0', T_0')$-FK-partition we obtain from (2) that

$$e(U_i, U_j) = \sum_{a,b\in[s]} d(W_{i,a}, W_{j,b})|U_{i,a}||U_{j,b}| \pm 6\delta'(t) n^2\,.$$

From the assumption $\{i,j\} \notin I$ we infer

$$\sum_{(a,b)\in A_{ij}} d(W_{i,a}, W_{j,b})|U_{i,a}||U_{j,b}| \le \sum_{(a,b)\in A_{ij}} |W_{i,a}||W_{j,b}| \le \frac{\varepsilon}{6}|V_i||V_j|$$

and, furthermore, for $(a,b) \notin A_{ij}$ we have

$$d(W_{i,a}, W_{j,b})|U_{i,a}||U_{j,b}| = \Big(d(V_i, V_j) \pm \frac{\varepsilon}{6}\Big)\,|U_{i,a}||U_{j,b}|\,.$$

Combining, those three estimates we infer

$$e(U_i, U_j) = \sum_{a,b\in[s]} d(V_i, V_j)|U_{i,a}||U_{j,b}| \pm \frac{\varepsilon}{6}|U_i||U_j| \pm \frac{\varepsilon}{6}|V_i||V_j| \pm 6\delta'(t) n^2\,.$$

Hence from our choice of $\delta'(t)$ and $V_i \ge \lfloor n/t \rfloor \ge n/(2t)$ we deduce

$$\big|e(U_i, U_j) - d(V_i, V_j)|U_i||U_j|\big| \le \frac{\varepsilon}{3}|V_i||V_j| + \frac{\varepsilon}{6}\left(\frac{n}{t}\right)^2 \le \varepsilon |V_i||V_j|\,,$$

which concludes the proof of Theorem 7. $\square$

In contrast to Theorem 1 the upper bound $T_{\mathrm{Sz}} = T_{\mathrm{Sz}}(\varepsilon, t_0)$ we obtain from the proof of Theorem 7 is not exponential, but of tower-type. In fact, we use Corollary 5 with $\nu = \varepsilon^4/36^2$ and $\delta(t) = \varepsilon/(36t^2)$. Due to the choice of ν we iterate Theorem 1 at most $36^2/\varepsilon^4$ times and each time the number of classes grows exponentially, i.e., t_i classes from the i-th iteration may split into $2^{O(t_i^4/\varepsilon^2)}$ classes for the next step. As a consequence, the upper bound $T_{\mathrm{Sz}} = T_{\mathrm{Sz}}(\varepsilon, t_0)$, which we obtain from this proof, is a tower of 4's of height $O(\varepsilon^{-4})$ with t_0 as the last exponent. The proof of Szemerédi's regularity lemma from [48] yields a similar upper bound of a tower of 2's of height proportional to ε^{-5}. However, recall that the statement from [48] is slightly different from the version proved here, by having a smaller error term in (7). A lower bound of similar type was obtained by Gowers [17]. In fact, Gowers showed an example of a graph for which any partition satisfying even only a considerably weaker version of property *(iii)* requires at least t classes, where t is a tower of 2's of height proportional to $1/\varepsilon^{1/16}$.

2.3. **The (ε, r)-regularity lemma.** As we have just discussed in the previous section, the example of Gowers shows that we cannot prevent the situation when the number of parts t of a Szemerédi-partition is much larger than, say, $1/\varepsilon$. For several applications this presents an obstacle which one would like to overcome. More precisely one would like to obtain some control of the densities of subgraphs which are of size much smaller than, say, n/t^2. The (ε, r)-regularity lemma (Theorem 9), the regularity lemma of Alon, Fischer, Krivelevich, and M. Szegedy (Theorem 10), and the regular approximation lemma (Theorem 11), were partly developed to address such issues.

A version for 3-uniform hypergraphs of the following regularity lemma was obtained by Frankl and Rödl in [11].

Theorem 9. *For every $\varepsilon > 0$, every function $r\colon \mathbb{N} \to \mathbb{N}$, and every $t_0 \in \mathbb{N}$ there exist $T_{\mathrm{FR}} = T_{\mathrm{FR}}(\varepsilon, r(\cdot), t_0)$ and n_0 such that for every graph $G = (V, E)$ with at least $|V| = n \geq n_0$ vertices the following holds. There exists a partition $V_1 \dot\cup \dots \dot\cup V_t = V$ such that*

(i) $t_0 \leq t \leq T_{\mathrm{FR}}$,
(ii) $|V_1| \leq \dots \leq |V_t| \leq |V_1| + 1$, *and*
(iii) *for all but at most εt^2 pairs (V_i, V_j) with $i < j$ we have that for all sequences of subsets $U_i^1, \dots, U_i^{r(t)} \subseteq V_i$ and $U_j^1, \dots, U_j^{r(t)} \subseteq V_j$*

$$\Big|\big|\bigcup_{q=1}^{r(t)} E(U_i^q, U_j^q)\big| - d(V_i, V_j)\big|\bigcup_{q=1}^{r(t)} U_i^q \times U_j^q\big|\Big| \leq \varepsilon |V_i||V_j| \,. \tag{9}$$

Note that if $r(t) \equiv 1$ then Theorem 9 is identical to Theorem 7 and if $r(t) \equiv k$ for some constant $k \in \mathbb{N}$ (independent of t), then it is a direct consequence of Theorem 7. We remark that for arbitrary functions $r(\cdot)$, Theorem 9 can be proved along the lines of Szemerédi's proof of Theorem 7 from [48]. Below we deduce Theorem 9, using a slightly different approach, namely we infer Theorem 9 from Corollary 5 in a similar way as we proved Theorem 7.

Proof. For given ε, $r(\cdot)$, and t_0 we follow the lines of the proof of Theorem 7. This time we apply Corollary 5 with a smaller choice of $\delta'(\cdot)$

$$\nu' = \frac{\varepsilon^4}{36^2}, \quad \varepsilon' = 1, \quad \delta'(t) = \frac{\varepsilon}{36t^2(4^{r(t)} - 3^{r(t)})}, \quad \text{and} \quad t_0' = t_0$$

and obtain T_0' and n_0', which determines T_{FR} and n_0. We define the sets A_{ij} and I identical as in the proof of Theorem 7, i.e., for $1 \le i < j \le t$ we set

$$A_{ij} = \big\{(a,b) \in [s] \times [s] \colon\ |d(W_{i,a}, W_{j,b}) - d(V_i, V_j)| \ge \varepsilon/6\big\}$$

and we let

$$I = \Big\{\{i,j\} \colon\ 1 \le i < j \le t \text{ such that } \textstyle\sum_{(a,b)\in A_{ij}} |W_{i,a}||W_{j,b}| \ge \varepsilon |V_i||V_j|/6\Big\}\,.$$

Again we obtain (8) and the rest of the proof requires some small straightforward adjustments.

We set $r = r(t)$ and we will show that if $\{i,j\} \notin I$, then (9) holds for every sequence $\hat{U}_i^1, \dots, \hat{U}_i^r \subseteq V_i$ and $\hat{U}_j^1, \dots, \hat{U}_j^r \subseteq V_j$. For such given sequences we consider new sequences $U_i^1, \dots, U_i^R \subseteq V_i$ and $U_j^1, \dots, U_j^R \subseteq V_j$ satisfying the disjointness property (see (10) below). For that let $R = 4^r - 3^r$ and for a non-empty set $\emptyset \neq L \subseteq [r]$ let

$$\hat{U}_i(L) = \bigcap_{\ell \in L} \hat{U}_i^\ell \setminus \bigcup_{\ell \in L} \hat{U}_i^\ell \quad \text{and} \quad \hat{U}_j(L) = \bigcap_{\ell \in L} \hat{U}_j^\ell \setminus \bigcup_{\ell \in L} \hat{U}_j^\ell$$

and for two sets L, L' with non-empty intersection we let $U_i(L, L') = \hat{U}_i(L)$ and $U_j(L, L') = \hat{U}_j(L')$. Note that there are $R = 4^r - 3^r$ such pairs of sets L, L' and we can relabel the sequences $(U_i(L,L'))_{L\cap L' \neq \emptyset}$ and $(U_j(L,L'))_{L\cap L'\neq\emptyset}$ to $U_i^1, \dots, U_i^R \subseteq V_i$ and $U_j^1, \dots, U_j^R \subseteq V_j$. Note that for all $p \neq q$ the sets U_i^p and U_i^q may either be equal or disjoint. Moreover, due to this definition we obtain for all $1 \le p < q \le R$

$$(U_i^q \times U_j^q) \cap (U_i^p \times U_j^p) = \emptyset \quad \text{and} \quad \dot{\bigcup}_{q\in[R]} U_i^q \times U_j^q = \bigcup_{q \in [r]} \hat{U}_i^q \times \hat{U}_j^q\,. \tag{10}$$

Furthermore, for $q \in [R]$ and a, $b \in [s]$ we set

$$U_{i,a}^q = U_i^q \cap W_{i,a} \quad \text{and} \quad U_{j,b}^q = U_j^q \cap W_{j,b}$$

and we get for every $q \in [R]$

$$e(U_i^q, U_j^q) = \sum_{a,b\in[s]} e(U_{i,a}^q, U_{j,b}^q)\,.$$

Appealing to the fact that $\mathcal{Q}$ is a $(\delta'(t), t_0', T_0')$-FK-partition we obtain from (2) that

$$e(U_i^q, U_j^q) = \sum_{a,b\in[s]} d(W_{i,a}, W_{j,b}) |U_{i,a}^q||U_{j,b}^q| \pm 6\delta'(t) n^2\,.$$

From the assumption $\{i,j\} \notin I$ and the disjointness property from (10) we infer

$$\sum_{(a,b)\in A_{ij}} \sum_{q\in[R]} d(W_{i,a}, W_{j,b}) |U_{i,a}^q||U_{j,b}^q| \le \sum_{(a,b)\in A_{ij}} |W_{i,a}||W_{j,b}| \le \frac{\varepsilon}{6} |V_i||V_j|$$

and, furthermore, for $(a,b) \notin A_{ij}$ we have

$$d(W_{i,a}, W_{j,b}) |U_{i,a}^q||U_{j,b}^q| = \Big(d(V_i, V_j) \pm \frac{\varepsilon}{6}\Big)\, |U_{i,a}^q||U_{j,b}^q|$$

for every $q \in [R]$. Combining, those three estimates we infer

$$\begin{aligned}\left|\bigcup_{q\in[r]} E(\hat{U}_i^q,\hat{U}_j^q)\right| &= \left|\dot{\bigcup}_{q\in[R]} E(U_i^q,U_j^q)\right| \\ &= \Big(d(V_i,V_j)\pm\frac{\varepsilon}{6}\Big)\sum_{q\in[R]}\sum_{a,b\in[s]}|U_{i,a}^q||U_{j,b}^q| \pm \frac{\varepsilon}{6}|V_i||V_j| \pm 6R\delta'(t)n^2 \\ &= d(V_i,V_j)\left|\bigcup_{q\in[R]} E(U_i^q,U_j^q)\right| \pm \frac{\varepsilon}{3}|V_i||V_j| \pm 6R\delta'(t)n^2\,.\end{aligned}$$

Hence from our choice of $\delta'(t)$, $R = (4^r - 3^r)$, and $V_i \geq \lfloor n/t \rfloor \geq n/(2t)$ we deduce from (10)

$$\Big|\big|\bigcup_{q=1}^r E(\hat{U}_i^q,\hat{U}_j^q)\big| - d(V_i,V_j)\big|\bigcup_{q=1}^r \hat{U}_i^q\times\hat{U}_j^q\big|\Big| \leq \varepsilon|V_i||V_j|\,,$$

which concludes the proof of Theorem 9. □

2.4. **The regularity lemma of Alon, Fischer, Krivelevich, and M. Szegedy.** In the last two sections we iterated the regularity lemma of Frieze and Kannan and obtained Corollary 5, from which we deduced Szemerédi's regularity lemma (Theorem 7) and the (ε,r)-regularity lemma (Theorem 9). From this point of view it seems natural to iterate these stronger regularity lemmas. This was indeed first carried out by Alon, Fischer, Krivelevich, and M. Szegedy [2] who iterated Szemerédi's regularity lemma for an application in the area of property testing.

Theorem 10. *For every ν, $\varepsilon > 0$, every function $\delta\colon \mathbb{N}\to(0,1]$, and every $t_0\in\mathbb{N}$ there exist $T_{\mathrm{AFKS}} = T_{\mathrm{AFKS}}(\nu,\varepsilon,\delta(\cdot),t_0)$ and n_0 such that for every graph $G=(V,E)$ with at least $|V| = n \geq n_0$ vertices the following holds. There exists a vertex partition $\mathcal{P} = (V_i)_{i\in[t]}$ with $V_1\dot\cup\dots\dot\cup V_t = V$ and a refinement $\mathcal{Q} = (W_{i,j})_{i\in[t],j\in[s]}$ with $W_{i,1}\dot\cup\dots\dot\cup W_{i,s} = V_i$ for every $i\in[t]$ such that*

- (*i*) *$\mathcal{P}$ is an $(\varepsilon,t_0,T_{\mathrm{AFKS}})$-Szemerédi-partition,*
- (*ii*) *$\mathcal{Q}$ is a $(\delta(t),t_0,T_{\mathrm{AFKS}})$-Szemerédi-partition, and*
- (*iii*) $\mathrm{ind}(\mathcal{Q}) \leq \mathrm{ind}(\mathcal{P}) + \nu$.

Proof. The proof is identical to the proof of Corollary 5 with the only adjustment that we iterate Theorem 7 instead of Theorem 1. □

The price for the stronger properties of the partitions $\mathcal{P}$ and $\mathcal{Q}$, in comparison to Szemerédi's regularity lemma, is again in the bound T_{AFKS}. In general T_{AFKS} can be expressed as a recursive formula in $\nu,\varepsilon,\delta(\cdot)$, and t_0, and for example, if $\delta(t)$ is given by a polynomial in $1/t$, then T_{AFKS} is an iterated tower-type function, which is sometimes referred to as a wowzer-type function.

Theorem 9 relates to Theorem 10 in the following way. It is a direct consequence of (9) that if (V_i,V_j) is not one of the exceptional pairs in (*iii*) of Theorem 9, then for any partition of V_i and V_j into at most $\sqrt{r(t)}$ parts of equal size, "most" of the $r(t)$ pairs have the density "close" (up to an error of $O(\sqrt{\varepsilon})$) to $d(V_i,V_j)$. Hence, if we set at the beginning $r(t) = (T_{\mathrm{Sz}}(\delta(t),t))^2$ and then apply Theorem 7 to obtain a $(\delta(t),t,T_{\mathrm{Sz}}(\delta(t),t))$-Szemerédi-partition $\mathcal{Q}$, which refines the given partition, then we arrive to a similar situation as in Theorem 10. In fact, we have two Szemerédi-partitions satisfying (*i*) and (*ii*) of Theorem 10 and (*iii*) would be replaced by the fact that $d(W_{i,a},W_{j,b}) \approx d(V_i,V_j)$ for "most" pairs from the finer partition $\mathcal{Q}$.

2.5. **The regular approximation lemma.** The following regularity lemma is another byproduct of the hypergraph generalization of the regularity lemma and appeared in general form in [37]. In a different context, Theorem 11 appeared in the work of Lovász and B. Szegedy [30, Lemma 5.2].

Theorem 11. *For every $\nu > 0$, every function $\varepsilon\colon \mathbb{N} \to (0,1]$, and every $t_0 \in \mathbb{N}$ there exist $T_0 = T_0(\nu, \varepsilon(\cdot), t_0)$ and n_0 such that for every graph $G = (V,E)$ with at least $|V| = n \geq n_0$ vertices the following holds. There exists a partition $\mathcal{P} = (V_i)_{i\in[t]}$ with $V_1 \dot\cup \ldots \dot\cup V_t = V$ and a graph $H = (V, E')$ on the same vertex set V as G such that*

- *(a) $\mathcal{P}$ is an $(\varepsilon(t), t_0, T_0)$-Szemerédi-partition for H and*
- *(b) $|E \triangle E'| = |E \setminus E'| + |E' \setminus E| \leq \nu n^2$.*

The main difference between Theorem 11 and Theorem 7 is in the choice of ε being a function of t. As already mentioned, it follows from the work of Gowers [17] (or alternatively from the work of Lovász and B. Szegedy [30, Proposition 7.1]) that it is not possible to obtain a Szemerédi (or even a Frieze-Kannan) partition for certain graphs G with ε of order $1/t$. Property (a) of Theorem 11 asserts, however, that by adding and deleting at most νn^2 edges from/to G we can obtain another graph H which admits a "much more" regular partition, e.g., with $\varepsilon(t) \ll 1/t$.

Below we show how Theorem 11 can be deduced from the iterated regularity lemma of Alon, Fischer, Krivelevich and M. Szegedy (Theorem 10). The idea is to apply Theorem 10 with appropriate parameters to obtain Szemerédi-partitions $\mathcal{P} = (V_i)_{i\in[t]}$ and $\mathcal{Q} = (W_{i,j})_{i\in[t],j\in[s]}$ for which $\mathcal{Q}$ refines $\mathcal{P}$ and $\mathrm{ind}(\mathcal{Q}) \leq \mathrm{ind}(\mathcal{P}) + \nu'$. The last condition and Lemma 6 imply that $d(W_{i,a}, W_{j,b}) \approx d(V_i, V_j)$ (with an error depending on ν') for "most" $i < j$ and $a, b \in [s]$. The strong regularity of the finer partition $\mathcal{Q}$ will then be used to adjust G (by adding and removing a few edges randomly) to obtain H for which $\mathcal{P}$ will have the desired properties. We now give the details of this outline.

Proof of Theorem 11. For given ν, $\varepsilon(\cdot)$, and t_0 we apply Theorem 10 with $\nu' = \nu^3/16$, some arbitrary ε', say $\varepsilon' = 1$, $\delta'(t) = \min\{\varepsilon(t)/2, \nu/4\}$, and $t'_0 = t_0$. We also fix an auxiliary constant $\gamma' = \nu/2$. We then set $T_0 = T'_{\mathrm{AFKS}}$ and $n_0 = n'_0$. After we apply Theorem 10 to the given graph $G = (V,E)$, we obtain an (ε', t_0, T_0)-Szemerédi-partition $\mathcal{P}$ and a $(\delta'(t), t_0, T_0)$-Szemerédi-partition $\mathcal{Q}$ which refines $\mathcal{P}$ such that $\mathrm{ind}(\mathcal{Q}) \leq \mathrm{ind}(\mathcal{P}) + \nu'$.

Next we will change G and obtain the graph H, which will satisfy (a) and (b) of Theorem 11. For that:

- (A) we replace every subgraph $G[W_{i,a}, W_{j,b}]$ which is not $\delta'(t)$-regular by a random bipartite graph of density $d(V_i, V_j)$ and
- (B) for every $1 \leq i < j \leq t$ and a, $b \in [s]$ we add or remove edges randomly to change the density of $G[W_{i,a}, W_{j,b}]$ to $d(V_i, V_j) + o(1)$.

It follows from the Chernoff bound that the resulting graph $H = (V, E')$ has the property that for *every* $1 \leq i < j \leq t$ and a, $b \in [s]$ the induced subgraph $H[W_{i,a}, W_{j,b}]$ is $(\delta'(t) + o(1))$-regular and $d_H(W_{i,a}, W_{j,b}) = d_H(V_i, V_j) + o(1)$, where $o(1) \to 0$ as $n \to \infty$. (Recall that for G from Lemma 6 we can only infer that $d_G(W_{i,a}, W_{j,b}) = d_G(V_i, V_j) \pm \gamma'$ for "most" pairs for some $\gamma' \gg \delta'(t)$.) Hence for every $1 \le i < j \le t$ and arbitrary sets $U_i \subseteq V_i$ and $U_j \subseteq V_j$ we have

$$\begin{aligned} &e_H(U_i, U_j) \\ &\quad = \sum_{a,b\in[s]} \Big(d_H(W_{i,a}, W_{j,b})|U_i \cap W_{i,a}||U_j \cap W_{j,b}| \pm (\delta'(t) + o(1))|W_{i,a}||W_{j,b}| \Big) \\ &\qquad\qquad\qquad\qquad\qquad\qquad = d_H(V_i, V_j)|U_i||U_j| \pm 2\delta'(t)|V_i||V_j| \,. \end{aligned}$$

In other words, the partition $\mathcal{P}$ is a $(2\delta'(t) \le \varepsilon(t), t_0, T_0)$-Szemerédi-partition for H, which is assertion (a) of Theorem 11. For part (b) we will estimate the symmetric difference of E and E'. Since $\mathcal{Q}$ is a $(\delta'(t), t_0, T_0)$-Szemerédi-partition for G the changes in Step (A) contributed at most

$$\delta'(t)t^2s^2 \left\lceil \frac{n}{ts} \right\rceil^2 \le \frac{\nu}{2} n^2 \tag{11}$$

to that difference.

For estimating the changes introduced in Step (B) we appeal to Lemma 6. From that we infer that, since $\operatorname{ind}(\mathcal{Q}) \le \operatorname{ind}(\mathcal{P}) + \nu'$, we "typically" changed only $\gamma'|W_{i,a}|W_{j,b}|$ pairs. More precisely, in Step (B) we changed at most

$$\begin{aligned} &\sum_{i<j} \sum_{a,b\in[s]} \gamma'|W_{i,a}||W_{j,b}| \\ &\qquad + \sum_{i<j} \sum_{a,b\in[s]} \Big\{ |W_{i,a}||W_{j,b}| \colon \ |d_G(W_{i,a}, W_{j,b}) - d_G(V_i, V_j)| \ge \gamma' \Big\} \\ &\qquad\qquad\qquad\qquad\qquad \le \left(\frac{\gamma'}{2} + \frac{\nu'}{(\gamma')^2} \right) n^2 \le \frac{\nu}{2} n^2 \,. \end{aligned} \tag{12}$$

Finally, from (11) and (12) we infer $|E \triangle E'| \le \nu n^2$, which shows that H satisfies property (b) of Theorem 11. ☐

2.6. **An early version of the regularity lemma.** We close this section with the statement of an early version of Szemerédi's regularity lemma, which was introduced in [47] and one of the key components in the proof of the Erdős-Turán conjecture concerning the upper density of subsets of the integers containing no arithmetic progression of fixed length. Another application of that lemma lead to the upper bound for the Ramsey-Turán problem for K_4 due to Szemerédi [46] and to the resolution of the $(6, 3)$-problem, which was raised by Brown, Erdős and Sós [6, 45], and solved by Ruzsa and Szemerédi [43].

Theorem 12. *For all positive ε_1, ε_2, δ, ϱ, and σ there exist T_0, S_0, M, and N such that for every bipartite graph $G = (X \dot\cup Y, E)$ satisfying $|X| = m \ge M$ and $|Y| = n \ge N$ there exists a partition $X_0 \dot\cup X_1 \dot\cup \dots \dot\cup X_t = X$ with $t \le T_0$ and for every $i = 1, \dots, t$ there exists a partition $Y_{i,0} \dot\cup Y_{i,1} \dot\cup \dots \dot\cup Y_{i,s_i} = Y$ with $s_i \le S_0$ such that*

- (a) *$|X_0| \le \varrho m$ and $|Y_{i,0}| \le \sigma n$ for every $i = 1, \dots, t$, and*
- (b) *for every $i = 1, \dots, t$, every $j = 1, \dots, s_i$, and all sets $U \subseteq X_i$ and $W \subseteq Y_{i,j}$ with $|U| \ge \varepsilon_1 |X_i|$ and $|W| \ge \varepsilon_2 |Y_{i,j}|$ we have $d(U, W) \ge d(X_i, Y_{i,j}) - \delta$.* ☐

Note that this lemma does not ensure such an elegant and easy to use structure of the partition as the later lemmas. More precisely, the partitions of Y may be very different for every $i = 1, \dots, t$. On the other hand, the upper bounds T_0 and S_0 are of similar type as those of Theorem 1, i.e., we have $T_0, S_0 = 2^{\mathrm{poly}(1/\min\{\varepsilon_1,\varepsilon_2,\delta,\varrho,\sigma\})}$. We also point out that for example in [43] Theorem 12 was applied iteratively, which in turn lead to a tower-type bound for the $(6,3)$-problem and up to now no better bound was found. A multipartite version of Theorem 12 was developed by Duke, Lefmann, and Rödl [8] for efficiently approximating the subgraph frequencies in a given graph G on n vertices for subgraphs of up to $\Omega(\sqrt{\log\log(n)})$ vertices.

Theorem 13. *For every $\varepsilon > 0$ and every integer $k \geq 2$ there exist $T_0 = 4^{k^2/\varepsilon^5}$ such that for every k-partite graph $G = (V, E)$ with vertex classes $V_1 \dot\cup \ldots \dot\cup V_k = V$ and $|V_1| = \cdots = |V_k| = N$ there exists a partition $\mathcal{P}$ of $V_1 \times \cdots \times V_k$ such that*

(i) the number of elements $W_1 \times \cdots \times W_k$ in $\mathcal{P}$ is at most T_0,

(ii) $|W_i| \geq \varepsilon^{k^2/\varepsilon^5} N$ for every $i = 1, \ldots, k$ and every $W_1 \times \cdots \times W_k$ in $\mathcal{P}$, and

(iii) we have

$$\sum_{W_1\times\cdots\times W_k \in \mathcal{P}_{\mathrm{irr}}} \prod_{i=1}^{k} |W_i| \leq \varepsilon N^k .$$

for the subfamily $\mathcal{P}_{\mathrm{irr}} \subseteq \mathcal{P}$ containing those elements $W_1 \times \cdots \times W_k$ from $\mathcal{P}$ which contain an irregular pair (W_i, W_j), i.e., a pair (W_i, W_j) with $i < j$ for which there exist subsets $U_i \subseteq W_i$ and $U_j \subseteq W_j$ with $|U_i| \geq \varepsilon|W_i|$ and $|U_j| \geq \varepsilon|W_j|$ such that $|d(U_i, U_j) - d(W_i, W_j)| > \varepsilon$. □

The main advantage of Theorem 13, in comparison to Szemerédi's regularity lemma (Theorem 7), is the smaller upper bound T_0. The partition in Theorem 13 still conveys information if $1/\varepsilon$ and k tend slowly to infinity with $n = |V|$, for example, if $1/\varepsilon$ and k are of order $\log^c(n)$ for some small constant $c > 0$. Due to the tower-type bound of Theorem 7 there $1/\varepsilon$ can be at most of order $\log^*(n)$, where $\log^*$ denotes the iterated logarithm function.

On the other hand, the upper bound T_0 in Theorem 13 is comparable to the one from Theorem 1 and as we will see in the next section Theorem 1 would be also well suited for the main application of Theorem 13 in [8]. Moreover, the structure of the partition provided by Theorem 1 seems to be simpler and easier to work with.

3. Reduced graph and counting lemmas

In this section we show how regular properties of the partitions given by the regularity lemmas from Section 2 can be applied to approximate the number of subgraphs of fixed isomorphism type of a given graph G. More precisely, for graphs G and F let $N_F(G)$ denote the number of labeled copies of F in G. Roughly speaking, we will show that $N_F(G)$ can be fairly well approximated by only studying the so-called *reduced graph* (or *cluster-graph*) of a regular partition.

Definition 14. *Let $\varepsilon > 0$, $G = (V, E)$ be a graph, and let $\mathcal{P} = (V_i)_{i\in[t]}$ be a partition of V.*

(i) If $\mathcal{P}$ is an ε-FK-partition, then the reduced graph *$R = R_G(\mathcal{P})$ is defined to be the weighted, complete, undirected graph with vertex set $V(R) = [t]$ and edge weights $w_R(i,j) = d(V_i, V_j)$.*

(ii) If $\mathcal{P}$ is an ε-Szemerédi-partition, then the reduced graph *$R = R_G(\mathcal{P}, \varepsilon)$ is defined to be the weighted, undirected graph with vertex set $V(R) = [t]$, edge set $E(R) = \{\{i,j\}\colon (V_i, V_j) \text{ is } \varepsilon\text{-regular}\}$, and edge weights $w_R(i,j) = d(V_i, V_j)$.*

The reduced graph carries a lot of the structural information of the given graph G. In fact, in many applications of the regularity lemma, the original problem for G one is interested in can be turned into a "simpler" problem for the reduced graph.

Remark 15. Below we will consider (labeled) copies F_R of a given graph F in a reduced graph R. If R is the reduced graph of an FK-partition, then R is an edge-weighting of the complete graph and, consequently, any ordered set of $|V(F)|$ vertices of $V(R)$ spans a copy of F. On the other hand, if R is the reduced graph of an ε-Szemerédi-partition, then R is not a complete graph and for a labeled copy F_R of F in R with $V(F_R) = \{i_1, \dots, i_\ell\}$ we will have that (V_{i_j}, V_{i_k}) is ε-regular for every edge $\{i_j, i_k\} \in E(F_R)$.

3.1. **The global counting lemma.** Here by a *counting lemma* we mean an assertion which enables us to deduce directly from the reduced graph some useful information on the number $N_F(G)$ of labeled copies of a fixed graph F in a large graph G. We will distinguish between two different settings here. The first counting lemma will yield an estimate on $N_F(G)$ in the context of Theorem 1. Since $N_F(G)$ concerns the total number of copies, we regard this result as a *global counting lemma.*

In contrast, for an ℓ-vertex graph F the *local counting lemma* (Theorem 18) will yield estimates on $N_F(G[V_{i_1}, \dots, V_{i_\ell}])$ for an induced ℓ-partite subgraph of G given by the regular partition $\mathcal{P}$. However, for such a stronger assertion we will require that $\mathcal{P}$ be a Szemerédi-partition.

Theorem 16. *Let F be a graph with vertex set $V(F) = [\ell]$. For every $\gamma > 0$ there exists $\varepsilon > 0$ such that for every $G = (V, E)$ with $|V| = n$ and every ε-FK-partition $\mathcal{P} = (V_i)_{i\in[t]}$ with reduced graph $R = R_G(\mathcal{P})$ we have*

$$N_F(G) = \sum_{F_R} \prod_{\{i_j,i_k\}\in E(F_R)} w_R(i_j, i_k) \prod_{i_j\in V(F_R)} |V_{i_j}| \pm \gamma n^\ell \,, \tag{13}$$

where the sum runs over all labeled copies F_R of F in R (cf. Remark 15).

For a simpler notation we denote here and below the vertices $V(F_R)$ of a given copy of F_R of the ℓ-vertex graph F in R by $\{i_1, \dots, i_\ell\}$ and omit the dependence of F_R.

Proof. We follow an argument of Lovász and B. Szegedy from [29]. We prove Theorem 18 by induction on the number of edges of F. Clearly, the theorem holds for graphs with no edges and for graphs with one edge it follows from the definition of ε-FK-partition with $\varepsilon = \gamma$.

For given F and γ we let $\varepsilon \le \gamma/12$ be sufficiently small, so that the statement for the induction assumption holds with $\gamma' = \gamma/2$. For two vertices $x, y \in V$ we set

$$d_{\mathcal{P}}(x, y) = \begin{cases} 0 & \text{if } x,\, y \in V_i \text{ for some } i \in [t], \\ d(V_i, V_j) & \text{if } x \in V_i \text{ and } y \in V_j \text{ for some } 1 \le i < j \le t \end{cases}$$

and we denote by $\mathbb{1}_E(x, y)$ the indicator function for E, i.e., $\mathbb{1}_E(x, y)$ equals 1 if $\{x, y\} \in E$ and it equals 0 otherwise. We consider the difference of the left-hand side and the main term of the right-hand side in (13) and obtain

$$\left| N_F(G) - \sum_{F_R} \prod_{\{i_j,i_k\}\in E(F_R)} w_R(i_j,i_k) \prod_{i_j\in V(F_R)} |V_{i_j}| \right| \\ = \left| \sum_{x_1,\dots,x_\ell\in(V)_\ell} \left(\prod_{\{i,j\}\in E(F)} \mathbb{1}_E(x_i,x_j) - \prod_{\{i,j\}\in E(F)} d_{\mathcal{P}}(x_i,x_j) \right) \right| , \quad (14)$$

where $x_1,\dots,x_\ell \in (V)_\ell$ is an arbitrary sequence of ℓ distinct vertices in V. Without loss of generality we may assume that $\{\ell-1,\ell\}$ is an edge in F and we denote by F^- the spanning subgraph of F with the edge $\{\ell-1,\ell\}$ removed. Then, applying the identity $\alpha_1\alpha_2 - \beta_1\beta_2 = \beta_2(\alpha_1-\beta_1) + \alpha_1(\alpha_2-\beta_2)$, we get the following upper bound for the right-hand side of the last equation

$$\left| \sum_{x_1,\dots,x_\ell\in(V)_\ell} d_{\mathcal{P}}(x_{\ell-1},x_\ell) \left(\prod_{\{i,j\}\in E(F^-)} \mathbb{1}_E(x_i,x_j) - \prod_{\{i,j\}\in E(F^-)} d_{\mathcal{P}}(x_i,x_j) \right) \right| \\ + \left| \sum_{x_1,\dots,x_\ell\in(V)_\ell} \left(\prod_{\{i,j\}\in E(F^-)} \mathbb{1}_E(x_i,x_j) \right) \Big(\mathbb{1}_E(x_{\ell-1},x_\ell) - d_{\mathcal{P}}(x_{\ell-1},x_\ell) \Big) \right| . \quad (15)$$

By the induction assumption we can bound the first term by $\gamma' n^\ell$, i.e., we have

$$\left| \sum_{x_1,\dots,x_\ell\in(V)_\ell} d_{\mathcal{P}}(x_{\ell-1},x_\ell) \left(\prod_{\{i,j\}\in E(F^-)} \mathbb{1}_E(x_i,x_j) - \prod_{\{i,j\}\in E(F^-)} d_{\mathcal{P}}(x_i,x_j) \right) \right| \le \gamma' n^\ell . \quad (16)$$

We will verify a similar bound for the second term in (15). For that we will split the second term of (15) into two parts and rewrite each of the parts (see (17) and (18) below).

We consider the induced subgraph F^* of F, which we obtain by removing the vertices labeled $\ell-1$ and ℓ from F. For a copy $\tilde{F}^*$ of F^* in G let $X_{\ell-1}(\tilde{F}^*)$ and $X_\ell(\tilde{F}^*)$ be those vertex sets such that for every pair $x_{\ell-1} \in X_{\ell-1}(\tilde{F}^*)$ and $x_\ell \in X_\ell(\tilde{F}^*)$ of distinct vertices, those two vertices extend $\tilde{F}^*$ in G to a copy of F^-. More precisely, if $x_1,\dots,x_{\ell-2}$ is the vertex set of $\tilde{F}^*$ then we set

$$X_{\ell-1}(\tilde{F}^*) = \bigcap_{i\colon \{i,\ell-1\}\in E(F)} \Gamma_G(x_i) \quad \text{and} \quad X_\ell(\tilde{F}^*) = \bigcap_{i\colon \{i,\ell\}\in E(F)} \Gamma_G(x_i) ,$$

where $\Gamma_G(x)$ denotes the set of neighbours of x in G. To simplify the notation, below we will write $X_{\ell-1}$ or X_ℓ instead of $X_{\ell-1}(\tilde{F}^*)$ or $X_\ell(\tilde{F}^*)$ as $\tilde{F}^*$ will be clear from the context. Since by definition edges contained in $X_{\ell-1}\cap X_\ell$ are counted twice in $e(X_{\ell-1},X_\ell)$ (cf. (2)) we observe for the first part of the seond term in (15) that

$$\sum_{x_1,\dots,x_\ell\in(V)_\ell} \left(\prod_{\{i,j\}\in E(F^-)} \mathbb{1}_E(x_i,x_j) \right) \mathbb{1}_E(x_{\ell-1},x_\ell) = \sum_{\tilde{F}^*} e(X_{\ell-1},X_\ell) , \quad (17)$$

Moreover, we have for the second part of the seond term in (15)

$$\sum_{x_1,\dots,x_\ell\in(V)_\ell}\left(\prod_{\{i,j\}\in E(F^-)}\mathbb{1}_E(x_i,x_j)\right)d_{\mathcal{P}}(x_{\ell-1},x_\ell)$$
$$=\sum_{\tilde{F}^*}\sum_{i\neq j\in[t]}d(V_i,V_j)|X_{\ell-1}\cap V_i||X_\ell\cap V_j| \quad (18)$$

and, consequently, we can bound the second term in (15) by

$$\sum_{\tilde{F}^*}\left|e(X_{\ell-1},X_\ell)-\sum_{i\neq j\in[t]}d(V_i,V_j)|X_{\ell-1}\cap V_i||X_\ell\cap V_j|\right| \quad (19)$$

Finally, we can apply the fact that $\mathcal{P}$ is an ε-FK-partition in form of (2) and the fact that $N_{F^*}(G)\leq n^{\ell-2}$ to bound (19) by $n^{\ell-2}\cdot 6\varepsilon n^2$. Hence, from (14)–(19) we infer

$$\left|N_F(G)-\sum_{F_R}\prod_{\{i_j,i_k\}\in E(F_R)}w_R(i_j,i_k)\prod_{i_j\in V(F_R)}|V_{i_j}|\right|\leq(\gamma'+6\varepsilon)n^\ell\leq\gamma n^\ell\,,$$

which concludes the proof of Theorem 18. □

A simple argument based on the principle of inclusion and exclusion yields an induced version of Theorem 16. Let $N_F^*(G)$ denote the number of labeled, induced copies of F in G.

Corollary 17. *Let F be a graph with vertex set $V(F)=[\ell]$. For every $\gamma>0$ there exists $\varepsilon>0$ such that for every $G=(V,E)$ with $|V|=n$ and every ε-FK-partition $\mathcal{P}=(V_i)_{i\in[t]}$ with reduced graph $R=R_G(\mathcal{P})$ we have*

$$N_F^*(G)=\sum_{F_R}\prod_{\{i_j,i_k\}\in E(F_R)}w_R(i_j,i_k)\prod_{\{i_j,i_k\}\in E(\overline{F_R})}\big(1-w_R(i_j,i_k)\big)\prod_{i_j\in V(F_R)}|V_{i_j}|\pm\gamma n^\ell,$$

where the sum runs over all labeled copies F_R of F in R and $\overline{F_R}$ denotes the complement graph of F_R on the same ℓ vertices $V(F_R)$.

Proof. Let F be a graph with $V(F)=[\ell]$ and let K^ℓ be the complete graph on the same vertex set. Let ε be sufficiently small, so that we can apply Theorem 16 with $\gamma'=\gamma/2^{\binom{\ell}{2}-e(F)}$ for every graph $F'\subseteq K^\ell$ which contains F. Let G, an ε-FK-partition $\mathcal{P}$, and a reduced graph $R=R_G(\mathcal{P})$ be given.

Due to the principle of inclusion and exclusion we have

$$N_F^*(G)=\sum_{F\subseteq F'\subseteq K^\ell}(-1)^{e(F')-e(F)}N_{F'}(G)\,,$$

where we sum over all supergraphs F' of F contained in K^ℓ. Applying Theorem 16 for every such F' we obtain

$$N_F^*(G)=\sum_{F'}(-1)^{e(F')-e(F)}\left(\sum_{F'_R}\prod_{\{i_j,i_k\}\in E(F'_R)}w_R(i_j,i_k)\prod_{i_j\in V(F_R)}|V_{i_j}|\right)\pm\gamma n^\ell\,,$$

where the outer sum runs over all F' with $F\subseteq F'\subseteq K^\ell$ and the inner sum is indexed by all copies F'_R of F' in R. We can rewrite the main term by rearranging the sum in the following way: First we sum over all possible labeled copies F_R of F

in R. Note that this fixes a unique labeled copy $K^\ell(F_R)$ of K^ℓ as well, and in the inner sum we consider all graphs F'_R in R "sandwiched" between F_R and $K^\ell(F_R)$. This way we obtain

$$\begin{aligned}
N^*_F(G) \pm \gamma n^\ell &= \sum_{F_R} \sum_{F_R \subseteq F'_R \subseteq K^\ell(F_R)} (-1)^{e(F'_R)-e(F_R)} \prod_{\{i_j,i_k\}\in E(F'_R)} w_R(i_j,i_k) \prod_{i_j\in V(F_R)} |V_{i_j}| \\
&= \sum_{F_R} \prod_{\{i_j,i_k\}\in E(F_R)} w_R(i_j,i_k) \prod_{i_j\in V(F_R)} |V_{i_j}| \\
&\qquad \times \sum_{F_R \subseteq F'_R \subseteq K^\ell(F_R)} (-1)^{e(F'_R)-e(F_R)} \prod_{\{i_j,i_k\}\in E(F'_R)\setminus E(F_R)} w_R(i_j,i_k) \\
&= \sum_{F_R} \prod_{\{i_j,i_k\}\in E(F_R)} w_R(i_j,i_k) \prod_{i_j\in V(F_R)} |V_{i_j}| \\
&\qquad \times \prod_{\{i_j,i_k\}\in E(K^\ell(F_R))\setminus E(F_R)} \big(1 - w_R(i_j,i_k)\big)\,,
\end{aligned}$$

which concludes the proof. □

3.2. **The local counting lemma.** For graphs F and G, a partition $\mathcal{P} = (V_i)_{i\in[t]}$ of $V(G)$, and a labeled copy F_R of F in R with $V(F_R) = \{i_1,\dots,i_\ell\}$ we denote by $N_F(G[F_R])$ the number of *partite isomorphic-copies* of F_R (and hence of F) in G induced on $V_{i_1}\dot\cup\dots\dot\cup V_{i_\ell}$. In other words, $N_F(G[F_R])$ is the number of edge preserving mappings φ from $V(F_R)$ to $V_{i_1}\dot\cup\dots\dot\cup V_{i_\ell}$ such that $\varphi(i_j) \in V_{i_j}$ for every $j = 1,\dots,\ell$.

Roughly speaking, the global counting lemma from the last section asserts that if $\mathcal{P}$ is a sufficiently regular ε-FK-partition, then $N_G(F)$ can be estimated from the reduced graph $R_G(\mathcal{P})$. In fact, it follows that the average of $N_F(G[F_R])$ over all labeled copies F_R of F in R is "close" to its expectation. The local counting lemma (Theorem 18), states that if $\mathcal{P}$ is, in fact, a sufficiently regular Szemerédi-partition, then this is not only true on average, but indeed for every copy F_R of F in R.

Recall, that by definition the edge set $E(R)$ of a reduced graph of a Szemerédi-partition $\mathcal{P}$ corresponds to the regular pairs of $\mathcal{P}$. Consequently, for a copy F_R of F in R we require that all edges of F_R correspond to regular pairs.

Theorem 18. *Let F be a graph with ℓ vertices. For every $\gamma > 0$ there exists $\varepsilon > 0$ such that for every $G = (V,E)$ with $|V| = n$ and every ε-Szemerédi-partition $\mathcal{P} = (V_i)_{i\in[t]}$ with reduced graph $R = R_G(\mathcal{P},\varepsilon)$ we have for every labeled copy F_R of F in R with $V(F_R) = \{i_1,\dots,i_\ell\}$*

$$N_F(G[F_R]) = \prod_{\{i_j,i_k\}\in E(F_R)} w_R(i_j,i_k) \prod_{i_j\in V(F_R)} |V_{i_j}| \pm \gamma \prod_{i_j\in V(F_R)} |V_{i_j}|\,.$$

Theorem 18 concerns the number of copies of a fixed graph F and will only give interesting bounds if we can assert $\gamma \ll \prod_{\{i_j,i_k\}\in E(F_R)} w_R(i_j,i_k)$. Moreover, it was shown by Chvátal, Rödl, Szemerédi, and Trotter [7], that if H is a graph of bounded degree with cn/t vertices (for some appropriate $c > 0$ depending on $\min_{\{i_j,i_k\}\in E(F_R)} w_R(i_j,i_k)$ and $\Delta(H)$) and there exists a homomorphism from H into F_R, then, under the same assumptions as in Theorem 18, $G[F_R]$ contains a copy of H. A far reaching strengthening, the so-called *blow-up lemma*, was found by Komlós, Sárközy, and Szemerédi [24]. The blow-up lemma allows, under some slightly more restrictive assumptions, to embed spanning graphs H of bounded degree.

Proof. We prove Theorem 18 by induction on the number of edges of F. Since the theorem is trivial for graphs with no edges and it follows from the definition of ε-Szemerédi-partition for $\varepsilon = \gamma$ for graphs with precisely one edge.

Let F be a graph with at least two edges and ℓ vertices. For given $\gamma > 0$ let $\varepsilon \leq \gamma/2$ be sufficiently small, so that the theorem holds for F^- with $\gamma' = \gamma/2$. Let $G = (V, E)$ be given along with an ε-Szemerédi-partition $\mathcal{P} = (V_i)_{i\in[t]}$ and let F_R be a labeled copy of F in R. Without loss of generality we may assume that $V(F_R) = \{1, \ldots, \ell\}$ and that $\{\ell - 1, \ell\}$ is an edge of F_R. We denote by F_R^- the subgraph of F_R which we obtain after deleting the edge $\{\ell-1, \ell\}$ from F_R. We can express the number of partite isomorphic copies of F_R through

$$\begin{aligned} N_F(G[F_R]) &= \sum_{x_1\in V_1} \cdots \sum_{x_\ell \in V_\ell} \prod_{\{i,j\}\in E(F_R)} \mathbb{1}_E(x_i, x_j) \\ &= \sum_{x_1\in V_1} \cdots \sum_{x_\ell \in V_\ell} \prod_{\{i,j\}\in E(F_R^-)} \Big(\mathbb{1}_E(x_i, x_j) \times \\ &\qquad\qquad \times \Big(d(V_{\ell-1}, V_\ell) + \mathbb{1}_E(x_{\ell-1}, x_\ell) - d(V_{\ell-1}, V_\ell) \Big) \Big) . \end{aligned}$$

The last expression can be rewritten as

$$\begin{aligned} & d(V_{\ell-1}, V_\ell) \times N_{F^-}(G[F_R^-]) + \\ &\quad + \sum_{x_1\in V_1} \cdots \sum_{x_\ell\in V_\ell} \Big(\prod_{\{i,j\}\in E(F_R^-)} \mathbb{1}_E(x_i, x_j) \Big(\mathbb{1}_E(x_{\ell-1}, x_\ell) - d(V_{\ell-1}, V_\ell) \Big) \Big) . \end{aligned}$$

From the induction assumption we then infer

$$d(V_{\ell-1}, V_\ell) N_F(G[F_R^-]) = \prod_{\{i_j,i_k\}\in E(F_R)} w_R(i_j, i_k) \prod_{i_j\in V(F_R)} |V_{i_j}| \pm \frac{\gamma}{2} \prod_{i_j\in V(F_R)} |V_{i_j}|$$

and, therefore, it suffices to verify

$$\begin{aligned} \left| \sum_{x_1\in V_1} \cdots \sum_{x_\ell\in V_\ell} \Big(\prod_{\{i,j\}\in E(F_R^-)} \mathbb{1}_E(x_i, x_j) \Big(\mathbb{1}_E(x_{\ell-1}, x_\ell) - d(V_{\ell-1}, V_\ell) \Big) \Big) \right| & \\ \leq \frac{\gamma}{2} \prod_{i_j\in V(F_R)} |V_{i_j}| & \qquad (20) \end{aligned}$$

For that we will appeal to the regularity of $\mathcal{P}$. Let F_R^* be the induced subgraph of F_R which one obtains by removing the vertices $\ell - 1$ and ℓ. For a partite isomorphic copy $\tilde{F}^*$ of F_R^*, let $X_{\ell-1}(\tilde{F}^*) \subseteq V_{\ell-1}$ and $X_\ell(\tilde{F}^*) \subseteq V_\ell$ be those sets of vertices for which any choice of $x_{\ell-1} \in X_{\ell-1}(\tilde{F}^*)$ and $x_\ell \in X_\ell(\tilde{F}^*)$ complete $\tilde{F}^*$ to a partite isomorphic copy of F_R^-. (To simplify the notation, below we will write $X_{\ell-1}$ or X_ℓ instead of $X_{\ell-1}(\tilde{F}^*)$ or $X_\ell(\tilde{F}^*)$ as $\tilde{F}^*$ will be clear from the context.) Consequently, summing over all partite isomorphic copies $\tilde{F}^*$ of F_R^* in G we obtain

$$\left|\sum_{x_1\in V_1}\cdots\sum_{x_\ell\in V_\ell}\Big(\prod_{\{i,j\}\in E(F_R^-)}\mathbb{1}_E(x_i,x_j)\Big(\mathbb{1}_E(x_{\ell-1},x_\ell)-d(V_{\ell-1},V_\ell)\Big)\Big)\right|$$
$$=\left|\sum_{\tilde{F}^*}e\big(X_{\ell-1},X_\ell\big)-d(V_{\ell-1},V_\ell)|X_{\ell-1}||X_\ell|\right|$$
$$\leq\sum_{\tilde{F}^*}\Big|e\big(X_{\ell-1},X_\ell\big)-d(V_{\ell-1},V_\ell)|X_{\ell-1})|X_\ell|\Big|$$
$$\leq\prod_{i=1}^{\ell-2}|V_i|\times\varepsilon|V_{\ell-1}||V_\ell|\,,$$

where in the last estimate we used the ε-regularity of $(V_{\ell-1},V_\ell)$ and the obvious upper bound on the number of partite isomorphic copies $\tilde{F}^*$ of F_R^*. Since $\varepsilon\leq\gamma/2$ the assertion (20) follows and concludes the proof of Theorem 18. □

We close this section by noting that an induced version of Theorem 18 can be derived directly from Theorem 18 in a similar way as Corollary 17 (we omit the details).

Corollary 19. *Let F be a graph with ℓ vertices. For every $\gamma>0$ there exists $\varepsilon>0$ such that for every $G=(V,E)$ with $|V|=n$ and every ε-Szemerédi-partition $\mathcal{P}=(V_i)_{i\in[t]}$ with reduced graph $R=R_G(\mathcal{P},\varepsilon)$ the following is true.*

For every labeled copy F_R of F contained in a clique $K_R^\ell\subseteq R$ with $V(F_R)=V(K_R^\ell)=\{i_1,\dots,i_\ell\}$

$$\prod_{\{i_j,i_k\}\in E(F_R)}w_R(i_j,i_k)\prod_{\{i_j,i_k\}\in\binom{V(F_R)}{2}\setminus E(F_R)}\big(1-w_R(i_j,i_k)\big)\prod_{i_j\in V(F_R)}|V_{i_j}|$$
$$=N_F^*(G[F_R])\pm\gamma\prod_{i_j\in V(F_R)}|V_{i_j}|\,,$$

where $N_F^*(G[F_R])$ denotes the number of labeled, induced, partite isomorphic copies of F_R in $G[F_R]=G[V_{i_1}\dot\cup\dots\dot\cup V_{i_\ell}]$. □

Note that by assumption of Corollary 19 and the definition of the reduced graph for Szemerédi-partitions we require for $F_R\subseteq R$ with $V(F_R)=\{i_1,\dots,i_\ell\}$, that (V_{i_j},V_{i_k}) is ε-regular for every pair $\{i_j,i_k\}$ and not only for pairs corresponding to edges of F_R.

3.3. **The removal lemma and its generalizations.** A direct consequence of the local counting lemma is the so-called *removal lemma*. Answering a question of Brown, Sós, and Erdős [6, 45] Ruzsa and Szemerédi [43] established the *triangle removal lemma*. They proved that every graph which contains only $o(n^3)$ triangles can be made triangle free by removing at most $o(n^2)$ edges. This result was generalized by Erdős, Frankl, and Rödl [10] from triangles to arbitrary graphs.

Theorem 20 (Removal lemma for graphs)**.** *For every graph F with ℓ vertices and every $\eta>0$ there exists $c>0$ and n_0 such that every graph $G=(V,E)$ on*

$n \geq n_0$ vertices with $N_F(G) < cn^\ell$, there exists a subgraph $H = (V, E')$ such that $N_F(H) = 0$ and $|E \setminus E'| \leq \eta n^2$.

While the original proof of Ruzsa and Szemerédi was based on an iterated application of the early version of the regularity lemma, Theorem 12, the proof given in [10] is based on Szemerédi's regularity lemma, Theorem 7. We remark that even in the triangle case both proofs give essentially the same tower-type dependency between c and η, i.e., c is a polynomial in $1/T$, where T is a tower of 2's of height polynomial in $1/\eta$. It is an intriguing open problem to find a proof which gives a better dependency between c and η.

Proof. Suppose that $G = (V, E)$ is a graph which even after the deletion of any set of at most ηn^2 edges still contains a copy of F. We will show that such a graph G contains at least cn^ℓ copies of F. For that we apply Szemerédi's regularity lemma, Theorem 7, with

$$\varepsilon = \min\left\{\frac{\eta}{8\ell^2}, \frac{\varepsilon'}{3\ell^2}\right\} \quad \text{and} \quad t_0 = \frac{1}{\eta},$$

where ε' is given by the local counting lemma applied with F and

$$\gamma = \frac{1}{3}\left(\frac{\eta}{4}\right)^{e(F)},$$

and obtain an ε-Szemerédi-partition $\mathcal{P} = (V_i)_{i\in[t]}$ of V. Next we delete all edges $e \in E$ for which at least one of the following holds:

- $e \subseteq V_i$ for some $i \in [t]$,
- $e \in E(V_i, V_j)$ for some $1 \leq i < j \leq t$ such that (V_i, V_j) is not ε-regular,
- $e \in E(V_i, V_j)$ for some $1 \leq i < j \leq t$ such that $d(V_i, V_j) \leq \eta/2$.

Simple calculations show that we delete at most ηn^2 edges in total. Let G' be the graph, which we obtain after the deletion of those edges. Due to the assumption on G, the graph G' must still contain a copy F_0 of F. Therefore the reduced graph $R = R_{G'}(\mathcal{P}, \varepsilon)$ must contain a copy of a homomorphic image F'_R of F for which $w_R(i_j, j_k) \geq \eta/2$ for all $\{i_j, j_k\} \in E(F'_R)$.

If F'_R is a copy of F, then the local counting lemma, Theorem 18, implies that G' contains, for sufficiently large n at least

$$\left(\frac{\eta}{2}\right)^{e(F)}\left\lfloor\frac{n}{t}\right\rfloor^\ell - \gamma\left\lceil\frac{n}{t}\right\rceil^\ell \geq \frac{1}{2}\left(\frac{\eta}{2}\right)^{e(F)}\left(\frac{n}{t}\right)^\ell$$

copies of F. Consequently, $N_F(G) \geq N_F(G') \geq cn^\ell$, for some c only depending on η and $T_{\mathrm{Sz}}(\min\{\eta/(8\ell^2), \varepsilon'/(3\ell^2)\}, 1/\eta)$, where ε' only depends on F and η. In other words, there exists such a c which only depends on the graph F and η as claimed.

The case when F'_R is not isomorphic to F is very similar. For example, we may subdivide every vertex class V_i into ℓ classes, $V_{i,1}\dot\cup\ldots\dot\cup V_{i,\ell} = V_i$, and obtain a refinement $\mathcal{Q}$. It follows from the definition of ε-regular pair, that if (V_i, V_j) is ε-regular, then $(V_{i,a}, V_{j,b})$ is $(3\ell^2\varepsilon)$-regular for any $a, b \in [\ell]$ and $d(V_{i,a}, V_{j,b}) \geq d(V_i, V_j) - 2\ell^2\varepsilon$. Since F'_R was contained in R, the reduced graph $S = S_{G'}(\mathcal{Q}, 3\ell^2\varepsilon)$ must contain a full copy F_R of F for which $w_R(i_j, j_k) \geq \eta/2 - 2\ell^2\varepsilon \geq \eta/4$ for all $\{i_j, j_k\} \in E(F_R)$ and the local counting lemma yields $N_F(G) \geq cn^\ell$ for

$$c = \frac{1}{2\ell^\ell}\left(\frac{\eta}{4}\right)^{e(F)} T_{\mathrm{Sz}}\left(\min\left\{\frac{\eta}{8\ell^2}, \frac{\varepsilon'}{3\ell^2}\right\}, \frac{1}{\eta}\right)^{-\ell}.$$

□

It was shown by Ruzsa and Szemerédi [43] that the removal lemma for triangles can be used to deduce Szemerédi's theorem on arithmetic progressions for progressions of length 3, which was earlier (and with better quantitative bounds) proved by Roth [42]. This connection was generalized by Frankl and Rödl [11, 35], who showed that the removal lemma for the complete k-uniform hypergraph with $k+1$ vertices implies Szemerédi's theorem for arithmetic progressions of length $k+1$. Moreover, Frankl and Rödl [11] verified such a removal lemma for $k=3$ (see also [32] for the general removal lemma for 3-uniform hypergraphs) and Rödl and Skokan [40] for $k=4$. The general result for k-uniform hypergraphs, based on generalizations of the regularity lemma and the local counting lemma for hypergraphs, was obtained independently by Gowers [18] and by Nagle, Skokan, and authors [33, 39, 41]. Moreover Solymosi [44] and Tengan, Tokushige, and authors [38] showed that this result also implies multidimensional versions of Szemerédi's theorem first obtained by Furstenberg and Katznelson [13, 14].

Besides those extensions to hypergraphs, generalizations of Theorem 20 for graphs were proved by several authors. In particular, the regularity lemma of Alon, Fischer, Krivelevich, and M. Szegedy, Theorem 10, was introduced to prove the natural analog of the removal lemma for induced copies of F. In fact, the proof of this statement is already considerably more involved. Later, Alon and Shapira [4, 3] generalized those results by replacing the fixed graph F by a possibly infinite family of graphs $\mathcal{F}$. All those proofs relied on Theorem 10. The most general version, due to Alon and Shapira [3], states the following.

Theorem 21. *For every (possibly infinite) family of graphs $\mathcal{F}$ and every $\eta > 0$ there exist constants $c > 0$, $C > 0$, and n_0 such that the following holds. Suppose $G = (V, E)$ is a graph on $n \geq n_0$ vertices. If for every $\ell = 1, \ldots, C$ and every $F \in \mathcal{F}$ on ℓ vertices we have $N_F^*(G) < cn^\ell$, then there exists a graph $H = (V, E')$ on the same vertex set as G such that $|E \triangle E'| \leq \eta n^2$ and $N_F^*(H) = 0$ for every $F \in \mathcal{F}$.* □

The proof of Theorem 21 is more involved and we will not present it here. The hypergraph extensions of Theorem 21 were obtained in [36].

Theorem 21 has interesting consequences in the area of *property testing*. Roughly speaking, it asserts that every graph G which is "far" (more than ηn^2 edges must be deleted or added) from some given hereditary property $\mathcal{A}$ (a property of graphs closed under isomorphism and vertex removal) must contain "many" ($cn^{|V(F)|}$) induced copies of some graph $F \notin \mathcal{A}$ of fixed size ($|V(F)| \leq C$). Consequently, a randomized algorithm can easily distinguish between graphs having $\mathcal{A}$ and those which are far from $\mathcal{A}$, provided $\mathcal{A}$ is decidable. One of the main questions in property testing, posed by Goldreich, Goldwasser, and Ron [16], concerns a natural characterization of properties allowing such a randomized algorithm. With respect to this question, the result of Alon and Shapira shows that all decidable, hereditary properties belong to that class (see [3] and [27] for more details).

An alternative proof of Theorem 21 was found by Lovász and B. Szegedy [28]. This new proof was based on the *limit approach* for sequences of dense graphs of those authors [29], which can be viewed as an infinitary iteration of Theorem 1. We will briefly explain this approach in the next section.

4. Graph limits

In Section 2 we first introduced the (weak) regularity lemma of Frieze and Kannan and from an iterated version we deduced Szemerédi's regularity lemma and the

(ε, r)-regularity lemma. Iterating Szemerédi's regularity lemma then resulted in the (strong) regularity lemma of Alon, Fischer, Krivelevich, and M. Szegedy, which was the key ingredient for the proof of Theorem 21.

It seems natural to further iterate any of those regularity lemmas. In fact, this was studied by Lovász and B. Szegedy [29]. Roughly speaking, those authors iterated the regularity lemma of Frieze and Kannan infinitely often. Below we will briefly outline some of their ideas. Note that due to the discussion above it does not matter which regularity lemma we iterate infinitely often, since we "pick up the other ones along the way".

Suppose $(G_i)_{i\in\mathbb{N}}$ is an infinite sequence of graphs with $|V(G_i)| \to \infty$ and $(\varepsilon_i)_{i\in\mathbb{N}}$ is a sequence of positive reals which tend to 0. Now we may apply Theorem 1 with ε_1 and $t_0 = 1$ to every sufficiently large graph G_i of the sequence. This way we obtain for every such graph G_i an ε_1-FK-partition $\mathcal{P}_{i,1}$ and a reduced graph $R_{i,1} = R_{G_i}(\mathcal{P}_{i,1})$. Note that all those partitions have at most $T_{\mathrm{FK}}(\varepsilon_1)$ parts. Hence, if we discretize the weights of the reduced graphs $R_{i,1}$ by quantities of up to ε_1, we note that there are only $\lceil 1/\varepsilon_1\rceil^{\binom{T_{\mathrm{FK}}(\varepsilon_1)}{2}}$ different possible reduced graphs. Consequently, there exists a weighted graph R_1 with at most $T_{\mathrm{FK}}(\varepsilon_1)$ vertices such that $R_{i,1} = R_1$ for infinitely many choices $i \in \mathbb{N}$. In other words, there exists an infinite subsequence $(G_{i_j})_{j\in\mathbb{N}}$ such that for every member there exists an ε_1-FK-partition, which yields R_1 as the reduced graph. We rename this sequence to $(G_i^1)_{i\in\mathbb{N}}$ and let $(\mathcal{P}_i^1)_{i\in\mathbb{N}}$ be the corresponding sequence of ε_1-FK-partitions.

We then repeat the above procedure with ε_2 for the infinite subsequence $(G_i^1)_{i\in\mathbb{N}}$, where the ε_2-FK-partitions should refine the ε_1-FK-partitions. This way we obtain a reduced graph R_2, an infinite subsequence $(G_i^2)_{i\in\mathbb{N}}$ of $(G_i^1)_{i\in\mathbb{N}}$, and a corresponding sequence of $(\mathcal{P}_i^2)_{i\in\mathbb{N}}$ of ε_2-FK-partitions. Repeating this step for every ε_j with $j \in \mathbb{N}$, we obtain a sequence of subsequences $(G_i^j)_{i\in\mathbb{N}}$ of graphs and a sequence of reduced graphs $(R_j)_{j\in\mathbb{N}}$. To avoid sequences of sequences of graphs we may pass to the diagonal sequence and let $(H_j)_{j\in\mathbb{N}} = (G_j^j)_{j\in\mathbb{N}}$ which is a subsequence of the original sequence of graphs $(G_i)_{i\in\mathbb{N}}$.

Summarizing the above, we have argued that for every infinite sequence of graphs $(G_i)_{i\in\mathbb{N}}$ with $|V(G_i)| \to \infty$ and every sequence of positive reals $(\varepsilon_i)_{i\in\mathbb{N}}$ there exists a subsequence $(H_j)_{j\in\mathbb{N}}$ of $(G_i)_{i\in\mathbb{N}}$, and a sequence of reduced graphs $(R_j)_{j\in\mathbb{N}}$ such that for every $j \in \mathbb{N}$ and every $k \in [j]$ the following holds:

(*a*) There exists an ε_k-FK-partition $\mathcal{P}_j^k$ of H_j such that $R_k = R_{H_j}(\mathcal{P}_j^k)$ and
(*b*) if $k < j$, then $\mathcal{P}_j^{k+1}$ refines $\mathcal{P}_j^k$.

In some sense the graphs in the sequence $(H_j)_{j\in\mathbb{N}}$ become more and more similar, since they have almost identical FK-partitions for smaller and smaller ε. On the other hand, they may have very different sizes, which makes it hard to compare them directly. In order to circumvent that we may scale them all to the same size, by viewing them as functions on $[0,1]^2$. We will now make this more precise.

Let R_j be a reduced graph with t_j vertices. We split $[0,1]$ into t_j intervals $I_{j,1}\dot\cup\ldots\dot\cup I_{j,t_j} = [0,1]$ each of size $1/t_j$. We then define the symmetric, step-function $\hat{R}_j\colon [0,1]^2 \to [0,1]$ by setting

$$\hat{R}_j(x,y) = \begin{cases} w_R(k,\ell), & \text{if } (x,y) \text{ belongs to the interior of } I_{j,k}\times I_{j,\ell},\\ 0, & \text{otherwise.}\end{cases}$$

Recall that those reduced graphs came from refining partitions (see (*b*) above) and it will be important for us to assume that the partitions $I_{j,1}\dot\cup\ldots\dot\cup I_{j,t_j}$ and $I_{j+1,1}\dot\cup\ldots\dot\cup I_{j+1,t_{j+1}}$ refine each other in the "same" way. More precisely, we assume that the first t_{j+1}/t_j vertices of R_{j+1} correspond in the ε_{j+1}-FK-partitions to those classes which were all contained in the first class of the ε_j-FK-partitions, while the second set of t_{j+1}/t_j vertices of R_{j+1} correspond in the ε_{j+1}-FK-partitions to those classes which were all contained in the second class of the ε_j-FK-partitions and so on. This way we embedded the sequence of reduced graphs $(R_j)_{j\in\mathbb{N}}$ into the family of symmetric step-functions from $[0,1]^2 \to [0,1]$. Similarly, we may embed the graphs from the sequence $(H_j)_{j\in\mathbb{N}}$. Here for a graph H_j on n_j vertices we split $[0,1]$ into n_j intervals $J_{j,1}\dot\cup\ldots\dot\cup J_{j,n_j} = [0,1]$ (identified by the vertices of H_j) and we set

$$\hat{H}_j(x,y) = \begin{cases} 1, & \text{if } (x,y) \text{ belongs to the interior of } J_{j,u}\times J_{j,v} \text{ and } \{u,v\}\in E(H_j),\\ 0, & \text{otherwise.}\end{cases}$$

Again we suppose that the labeling of the vertices of H_j is "consistent", i.e., if u is a vertex contained in the k-th vertex class of the fixed ε_j-FK-partition of H_j, then we impose that $J_{j,u}\subseteq I_{j,k}$.

After this embedding we can rewrite the property that R_j is the reduced graph of an ε_j-FK-partition of H_j, by

$$\sup_{U\subseteq[0,1]} \left| \int_{U\times U} \hat{H}_j(x,y) - \hat{R}_j(x,y)\,\mathrm{d}x\mathrm{d}y \right| \le \hat{\varepsilon}_j\,, \tag{21}$$

for some $\hat{\varepsilon}_j$ which tends to 0 as ε_j tends to 0. (Note that $\hat{H}_j$ and $\hat{R}_j$ are piecewise linear and, hence, (Lebesgue) measurable on $[0,1]^2$.) Moreover, we can rephrase the global counting lemma, Theorem 16: Let F be a graph with $V(F)=[\ell]$ and let j be sufficiently large (so that ε_j is sufficiently small). Then

$$\frac{N_F(H_j)}{n_j^\ell} = \int_{(x_1,\ldots,x_\ell)\in[0,1]^\ell} \prod_{\{p,q\}\in E(F)} \hat{R}_j(x_p,x_q)\,\mathrm{d}x_1\ldots\mathrm{d}x_\ell \pm \hat{\gamma}_j\,, \tag{22}$$

where for fixed F we have $\hat{\gamma}_j\to 0$ as $\varepsilon_j\to 0$.

It was proved by Lovász and B. Szegedy in [29] that, due to property (*b*) above, the sequence $(\hat{R}_j)_{j\in\mathbb{N}}$ converges almost everywhere to a measurable, symmetric function $\hat{R}\colon [0,1]^2\to[0,1]$ and that (21) and (22) stay valid in the limit. The function $\hat{R}$ is called the limit of the sequence $(H_j)_{j\in\mathbb{N}}$.

Theorem 22. *For every sequence of graphs $(G_i)_{i\in\mathbb{N}}$ with $|V(G_i)|\to\infty$ there exists a subsequence $(H_j)_{j\in\mathbb{N}}$ and a sequence of reduced graphs $(R_j)_{j\in\mathbb{N}}$, and a measurable, symmetric function $\hat{R}\colon [0,1]^2\to[0,1]$ such that*

(*i*) *$\hat{R}_j$ converges pointwise almost everywhere to $\hat{R}$,*

(*ii*)

$$\lim_{j\to\infty}\sup_{U\subseteq[0,1]} \left| \int_{U\times U} \hat{H}_j(x,y) - \hat{R}(x,y)\,\mathrm{d}x\mathrm{d}y \right| = 0\,,$$

and

(*iii*) *for every $\ell\in\mathbb{N}$ and every graph F with $V(F)=[\ell]$*

$$\lim_{j\to\infty} \frac{N_F(H_j)}{n_j^\ell} = \int_{(x_1,\dots,x_\ell)\in[0,1]^\ell} \prod_{\{p,q\}\in E(F)} \hat{R}(x_p, x_q)\,\mathrm{d}x_1 \dots \mathrm{d}x_\ell\,. \qquad \square$$

The proof of Theorem 22 indicated above, essentially follows the lines of the proof of the implication $(a) \Rightarrow (b)$ of Theorem 2.2 in [29] (see Lemma 5.1 and 5.2 in [29]). Based on Theorem 22 Lovász and B. Szegedy [28] gave a different and conceptually simpler proof of Theorem 21. The proof of the generalization of Theorem 21 to k-uniform hypergraphs in [36] followed similar ideas (see also [5]).

References

[1] N. Alon, A. Coja-Oghlan, H. Hàn, M. Kang, V. Rödl, and M. Schacht, *Quasi-randomness and algorithmic regularity for graphs with general degree distributions*, SIAM J. Comput., to appear. 1

[2] N. Alon, E. Fischer, M. Krivelevich, and M. Szegedy, *Efficient testing of large graphs*, Combinatorica **20** (2000), no. 4, 451–476. 1, 2, 2.2, 2.4

[3] N. Alon and A. Shapira, *A characterization of the (natural) graph properties testable with one-sided error*, SIAM J. Comput. **37** (2008), no. 6, 1703–1727. 3.3, 3.3

[4] ———, *Every monotone graph property is testable*, SIAM J. Comput. **38** (2008), no. 2, 505–522. 3.3

[5] T. Austin and T. Tao, *On the testability and repair of hereditary hypergraph properties*, submitted. 4

[6] W. G. Brown, P. Erdős, and V. T. Sós, *Some extremal problems on r-graphs*, New directions in the theory of graphs (Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich, 1971), Academic Press, New York, 1973, pp. 53–63. 2.6, 3.3

[7] V. Chvátal, V. Rödl, E. Szemerédi, and W. T. Trotter, Jr., *The Ramsey number of a graph with bounded maximum degree*, J. Combin. Theory Ser. B **34** (1983), no. 3, 239–243. 3.2

[8] R. A. Duke, H. Lefmann, and V. Rödl, *A fast approximation algorithm for computing the frequencies of subgraphs in a given graph*, SIAM J. Comput. **24** (1995), no. 3, 598–620. 2, 2.6, 2.6

[9] P. Erdős and P. Turán, *On some sequences of integers*, J. London. Math. Soc. **11** (1936), 261–264. 1

[10] P. Erdős, P. Frankl, and V. Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs Combin. **2** (1986), no. 2, 113–121. 3.3, 3.3

[11] P. Frankl and V. Rödl, *Extremal problems on set systems*, Random Structures Algorithms **20** (2002), no. 2, 131–164. 2, 2.3, 3.3

[12] A. Frieze and R. Kannan, *Quick approximation to matrices and applications*, Combinatorica **19** (1999), no. 2, 175–220. 1, 2, 2.1

[13] H. Furstenberg and Y. Katznelson, *An ergodic Szemerédi theorem for commuting transformations*, J. Analyse Math. **34** (1978), 275–291 (1979). 3.3

[14] ———, *An ergodic Szemerédi theorem for IP-systems and combinatorial theory*, J. Analyse Math. **45** (1985), 117–168. 3.3

[15] S. Gerke and A. Steger, *The sparse regularity lemma and its applications*, Surveys in combinatorics 2005, London Math. Soc. Lecture Note Ser., vol. 327, Cambridge Univ. Press, Cambridge, 2005, pp. 227–258. 1

[16] O. Goldreich, S. Goldwasser, and D. Ron, *Property testing and its connection to learning and approximation*, J. ACM **45** (1998), no. 4, 653–750. 3.3

[17] W. T. Gowers, *Lower bounds of tower type for Szemerédi's uniformity lemma*, Geom. Funct. Anal. **7** (1997), no. 2, 322–337. 2.2, 2.5

[18] ———, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. of Math. (2) **166** (2007), no. 3, 897–946. 3.3

[19] P. E. Haxell, Y. Kohayakawa, and T. Łuczak, *The induced size-Ramsey number of cycles*, Combin. Probab. Comput. **4** (1995), no. 3, 217–239. 1

[20] ———, *Turán's extremal problem in random graphs: forbidding even cycles*, J. Combin. Theory Ser. B **64** (1995), no. 2, 273–287. 1

[21] ———, *Turán's extremal problem in random graphs: forbidding odd cycles*, Combinatorica **16** (1996), no. 1, 107–122. 1
[22] Y. Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, Foundations of computational mathematics (Rio de Janeiro, 1997), pp. 216–230. 1
[23] Y. Kohayakawa and V. Rödl, *Szemerédi's regularity lemma and quasi-randomness*, Recent advances in algorithms and combinatorics, CMS Books Math./Ouvrages Math. SMC, vol. 11, Springer, New York, 2003, pp. 289–351. 1
[24] J. Komlós, G. N. Sárközy, and E. Szemerédi, *Blow-up lemma*, Combinatorica **17** (1997), no. 1, 109–123. 3.2
[25] J. Komlós, A. Shokoufandeh, M. Simonovits, and E. Szemerédi, *The regularity lemma and its applications in graph theory*, Theoretical aspects of computer science (Tehran, 2000), Lecture Notes in Comput. Sci., vol. 2292, Springer, Berlin, 2002, pp. 84–112. 1
[26] 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. 1
[27] L. Lovász and B. Szegedy, *Testing properties of graphs and functions*, Israel J. Math., to appear. 3.3
[28] ———, *Graph limits and testing hereditary graph properties*, Tech. Report MSR-TR-2005-110, Microsoft Research, 2005. 3.3, 4
[29] ———, *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), no. 6, 933–957. 3.1, 3.3, 4, 4, 4
[30] ———, *Szemerédi's lemma for the analyst*, Geom. Funct. Anal. **17** (2007), no. 1, 252–270. 1, 2.1, 2.5, 2.5
[31] T. Łuczak, *Randomness and regularity*, International Congress of Mathematicians. Vol. III, Eur. Math. Soc., Zürich, 2006, pp. 899–909. 1
[32] B. Nagle and V. Rödl, *Regularity properties for triple systems*, Random Structures Algorithms **23** (2003), no. 3, 264–332. 3.3
[33] B. Nagle, V. Rödl, and M. Schacht, *The counting lemma for regular k-uniform hypergraphs*, Random Structures Algorithms **28** (2006), no. 2, 113–179. 3.3
[34] V. Rödl, unpublished. 1
[35] ———, *Some developments in Ramsey theory*, Proceedings of the International Congress of Mathematicians, Vol. I, II (Kyoto, 1990) (Tokyo), Math. Soc. Japan, 1991, pp. 1455–1466. 3.3
[36] V. Rödl and M. Schacht, *Generalizations of the removal lemma*, Combinatorica, to appear. 3.3, 4
[37] ———, *Regular partitions of hypergraphs: regularity lemmas*, Combin. Probab. Comput. **16** (2007), no. 6, 833–885. 2, 2.5
[38] V. Rödl, M. Schacht, E. Tengan, and N. Tokushige, *Density theorems and extremal hypergraph problems*, Israel J. Math. **152** (2006), 371–380. 3.3
[39] V. Rödl and J. Skokan, *Regularity lemma for k-uniform hypergraphs*, Random Structures Algorithms **25** (2004), no. 1, 1–42. 3.3
[40] ———, *Counting subgraphs in quasi-random 4-uniform hypergraphs*, Random Structures Algorithms **26** (2005), no. 1-2, 160–203. 3.3
[41] ———, *Applications of the regularity lemma for uniform hypergraphs*, Random Structures Algorithms **28** (2006), no. 2, 180–194. 3.3
[42] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 104–109. 3.3
[43] I. Z. Ruzsa and E. Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II, Colloq. Math. Soc. János Bolyai, vol. 18, North-Holland, Amsterdam, 1978, pp. 939–945. 2.6, 2.6, 3.3, 3.3
[44] J. Solymosi, *A note on a question of Erdős and Graham*, Combin. Probab. Comput. **13** (2004), no. 2, 263–267. 3.3
[45] V. T. Sós, P. Erdős, and W. G. Brown, *On the existence of triangulated spheres in 3-graphs, and related problems*, Period. Math. Hungar. **3** (1973), no. 3-4, 221–228. 2.6, 3.3
[46] E. Szemerédi, *On graphs containing no complete subgraph with 4 vertices*, Mat. Lapok **23** (1972), 113–116 (1973). 2.6
[47] ———, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245, Collection of articles in memory of Juriĭ Vladimirovič Linnik. 1, 2, 2.6

[48] ______, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401. 2, 2.2, 2.2, 2.2, 2.3

[49] T. Tao, *Szemerédi's regularity lemma revisited*, Contrib. Discrete Math. **1** (2006), no. 1, 8–28 (electronic). 2.2

Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30322, USA

E-mail address: `rodl@mathcs.emory.edu`

Institut für Informatik, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany

E-mail address: `schacht@informatik.hu-berlin.de`