6-25-2019

An Optimal Medium-Strength Regularity Algorithm for 3-uniform Hypergraphs

John Theado
University of South Florida, jmtx222@gmail.com

An Optimal Medium-Strength Regularity Algorithm for 3-uniform Hypergraphs

by

John Theado

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Mathematics & Statistics
College of Arts and Sciences
University of South Florida

Major Professor: Brendan Nagle, Ph.D.
Jay Ligatti, Ph.D.
Brian Curtin, Ph.D.
Natăsa Jonoska, Ph.D.
Theodore Molla, Ph.D.
Dmytro Savchuk, Ph.D.

Date of Approval:
May 28, 2019

Keywords: Extremal Combinatorics, Quasirandomness, Density, Links, Partition

Acknowledgments

First, I would like to thank my family for always believing that I was capable of making it this far.

I would like to thank my fiancèe Lisa, for her support and confidence in my ability to complete this work.

I would like to thank my advisor, Dr. Brendan Nagle, for all of his support and advice throughout my time at USF, and for first introducing me to combinatorics and graph theory.

I would like to thank my committee members for their support, encouragement, and advice throughout the years.

I would like to thank all of my professors at USF, especially those in the Department of Mathematics and Statistics, for teaching me so much during my time at USF.

Finally, I would like to thank you for your interest and time reading my dissertation.

Table of Contents

List of Figures

Abstract

Szemerédi's Regularity Lemma [32, 33] is an important tool in combinatorics, with numerous applications in combinatorial number theory, discrete geometry, extremal graph theory, and theoretical computer science.

The Regularity Lemma hinges on the following concepts. Let $G = (V, E)$ be a graph and let $\emptyset \neq X, Y \subset V$ be a pair of disjoint vertex subsets. We define the *density* of the pair (X, Y) by $d_G(X, Y) = |E[X, Y]|/(|X||Y|)$ where $E[X, Y]$ denotes the set of edges $\{x, y\} \in E$ with $x \in X$ and $y \in Y$. We say the pair (X, Y) is *ε-regular* if all subsets $X' \subseteq X$ and $Y' \subseteq Y$ satisfying $|X'| > \varepsilon|X|$ and $|Y'| > \varepsilon|Y|$ also satisfy $|d_G(X', Y') - d_G(X, Y)| < \varepsilon$.

The Regularity Lemma states that, for all $\varepsilon > 0$, all large n-vertex graphs $G = (V, E)$ admit a partition $V = V_1 \cup \cdots \cup V_t$, where $t = t(\varepsilon)$ depends on ε but not on n, so that all but εt^2 pairs (V_i, V_j), $1 \leq i < j \leq t$, are ε-regular. While Szemerédi's original proof demonstrates the *existence* of such a partition, it gave no method for (efficiently) *constructing* such partitions. Alon, Duke, Lefmann, Rödl, and Yuster [1, 2] showed that such partitions can be constructed in time $O(M(n))$, where $M(n)$ is the time needed to multiply two $n \times n$ $\{0, 1\}$-matrices over the integers. Kohayakawa, Rödl, and Thoma [17, 18] improved this time to $O(n^2)$.

The Regularity Lemma can be extended to k-uniform hypergraphs, as can algorithmic formulations thereof. The most straightforward of these extends the concepts above to k-uniform hypergraphs $\mathcal{H} = (V, E)$ in a nearly verbatim way. Let $\emptyset \neq X_1, \ldots, X_k \subset V$ be pairwise disjoint subsets, and let $E[X_1, \ldots, X_k]$ denote the set of k-tuples $\{x_1, \ldots, x_k\} \in E$ satisfying $x_1 \in X_1, \ldots, x_k \in X_k$. We define the *density* of $(X_1, \ldots, X_k)$ as

$$d_{\mathcal{H}}(X_1, \ldots, X_k) = \frac{|E[X_1, \ldots, X_k]|}{|X_1| \cdots |X_k|}.$$

We say that $(X_1, \ldots, X_k)$ is ε-*regular* if all subsets $X_i' \subseteq X_i$, $1 \leq i \leq k$, satisfying $|X_i'| > \varepsilon |X_i|$ also satisfy

$$|d_H(X_1', \ldots, X_k') - d_H(X_1, \ldots, X_k)| < \varepsilon.$$

With these concepts, Szemeredi's original proof can be applied to give that, for all integers $k \geq 2$ and for all $\varepsilon > 0$, all n-vertex k-uniform hypergraphs $\mathcal{H} = (V, E)$ admit a partition $V = V_1 \cup \cdots \cup V_t$, where $t = t(k, \varepsilon)$ is independent of n, so that all but εt^k many k-tuples $(V_{i_1}, \ldots, V_{i_k})$ are ε-regular, where $1 \leq i_1 < \cdots < i_k \leq t$. Czygrinow and Rödl [4] gave an algorithm for such a regularity lemma, which in the context above, runs in time $O(n^{2k-1} \log^5 n)$.

In this dissertation, we consider regularity lemmas for 3-uniform hypergraphs. In this setting, our first main result improves the algorithm of Czygrinow and Rödl to run in time $O(n^3)$, which is optimal in its order of magnitude. Our second main result shows that this algorithm gives a stronger notion of regularity than what is described above, where this stronger notion is described in the course of this dissertation. Finally, we discuss some ongoing applications of our constructive regularity lemmas to some classical algorithmic hypergraph problems.

Chapter 1
Introduction

Szemerédi's Regularity Lemma [32, 33] is an important tool in combinatorics, with numerous applications in combinatorial number theory, discrete geometry, extremal graph theory, and theoretical computer science. (While we describe a few such applications in this Introduction, we refer the Reader to the well-cited surveys [19, 20], or the graduate texts [3, 5, 34], for more applications.) The regularity lemma, as well as the main results of this dissertation, belong to the use and development of *quasirandom* techniques in combinatorics, pioneered by Endre Szemerédi, who devised his regularity lemma in the course of establishing his celebrated Density Theorem, which confirmed a long-standing conjecture of Erdős and Turán [7] from 1936.

Theorem 1.0.1 (Szemerédi Density Theorem)**.** *For all $\delta > 0$ and integers $k \in \mathbb{N}$, there exists an integer $n_0 = n_0(\delta, k) \in \mathbb{N}$ so that the following holds. Let $n \geq n_0$ and let $A = A_n \subseteq [n] = \{1, \dots, n\}$ be a subset with no arithmetic progression of length k. Then $|A| \leq \delta n$.*

Since this origin, the regularity lemma emerged to be an important tool for a wealth of combinatorial results, and for his work with the regularity lemma, Density Theorem, and many other career accomplishments, Szemerédi was awarded the *Abel Prize* in 2012.

Roughly speaking, the regularity lemma guarantees that all large graphs $G = (V, E)$ can be partitioned into a constant number of 'random-like' bipartite subgraphs. To make these notions precise, fix a graph $G = (V, E)$, and fix disjoint subsets $\emptyset \neq X, Y \subset V$. Let $E[X, Y] = E_G[X, Y]$ denote the set of edges $\{x, y\} \in E$ with $x \in X$ and $y \in Y$, and let $K[X, Y]$ denote all pairs $\{x, y\}$, not necessarily in E, where $x \in X$ and $y \in Y$. We define the *density* of (X, Y) as

$$d(X, Y) = d_G(X, Y) = \frac{|E[X, Y]|}{|K[X, Y]|} = \frac{|E[X, Y]|}{|X||Y|}$$

For $\varepsilon > 0$, the pair (X, Y) is said to be *ε-regular* whenever, for all $X' \subseteq X$ with $|X'| > \varepsilon|X|$ and for all $Y' \subseteq Y$ with $|Y'| > \varepsilon|Y|$, we have

$$|d(X', Y') - d(X, Y)| < \varepsilon. \tag{1.1}$$

Moreover, we say that (X, Y) is *(d, ε)-regular*, for some $d \geq 0$, if we can replace $d(X, Y)$ in (1.1) by d. When (1.1) fails to hold for some pair of subsets $X' \subseteq X$ with $|X'| > \varepsilon|X|$ and $Y' \subseteq Y$ with $|Y'| > \varepsilon|Y|$, then we say that (X, Y) is *ε-irregular*, and that (X', Y') is a *witness* to the ε-irregularity of (X, Y).

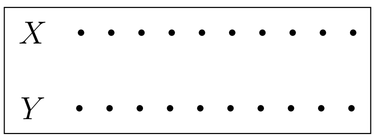

Figure 1.: $\bar{K}[X, Y]$ has density 0 and is ε-regular $\forall \varepsilon > 0$.

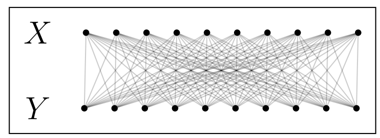

Figure 2.: $K[X, Y]$ has density 1 and is ε-regular $\forall \varepsilon > 0$.

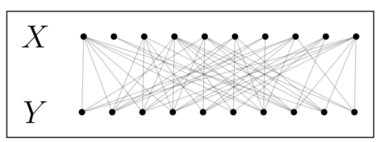

Figure 3.: $\forall \varepsilon > 0$, w.h.p., as $|X|, |Y| \to \infty$, $\mathbb{G}[X, Y; \frac{1}{2}]$ has density $\frac{1}{2} \pm \varepsilon$ and is ε-regular.

Note that (1.1) ensures that the edges $E[X, Y]$ are evenly distributed in portions of density near $d = d(X, Y)$ among all suitably large subsets $X' \subseteq X$ and $Y' \subseteq Y$. Uniform edge-distribution is a property shared by the binomial random bipartite graph $\mathbb{G}(X, Y; d)$ of edge-density d, which independently includes each pair $\{x, y\}$ as an edge, for each $x \in X$ and $y \in Y$, with probability d. Indeed, elementary probability establishes that

$$\mathbb{E}\big[\big|E_{\mathbb{G}}[X', Y']\big|\big] = d|X'||Y'|,$$

and so by the Chernoff inequality [16],

$$\mathbb{P}\big[\big|E_{\mathbb{G}}[X',Y']\big| \neq (1\pm\varepsilon)d|X'||Y'|\big] \leq 2\exp\big\{-\tfrac{1}{3}\varepsilon^2 d|X'||Y'|\big\}$$
$$\leq 2\exp\big\{-\tfrac{1}{3}\varepsilon^5 d|X||Y|\big\} \tag{1.2}$$

when $|X'| > \varepsilon|X|$ and $|Y'| > \varepsilon|Y|$. Since there are at most $2^{|X|+|Y|}$ subsets $X' \subseteq X$ and $Y' \subseteq Y$, we conclude from (1.2) that with probability $1 - o(1)$, where $o(1) \to 0$ as $|X|, |Y| \to \infty$, all of them satisfying $|X'| > \varepsilon|X|$ and $|Y'| > \varepsilon|Y|$ also satisfy $|d_{\mathbb{G}}(X', Y') - d| < \varepsilon$.

We now state the regularity lemma precisely.

Theorem 1.0.2 (Szemerédi's regularity lemma)**.** *For all $\varepsilon > 0$ and integers $t_0 \in \mathbb{N}$, there exist integers $T_0 = T_0(\varepsilon, t_0)$ and $N_0 = N_0(\varepsilon, t_0)$ so that every graph $G = (V, E)$ on $n = |V| > N_0$ many vertices admits a partition $V = V_1 \dot\cup \ldots \dot\cup V_t$, with $t_0 \leq t \leq T_0$, satisfying that*

(i) $V = V_1 \dot\cup \ldots \dot\cup V_t$ *is* t-equitable*:* $|V_1| \leq \cdots \leq |V_t| \leq |V_1| + 1$*;*

(ii) $V = V_1 \dot\cup \ldots \dot\cup V_t$ *is* ε-regular*: all but εt^2 of its pairs (V_i, V_j), $1 \leq i < j \leq t$, are ε-regular.*

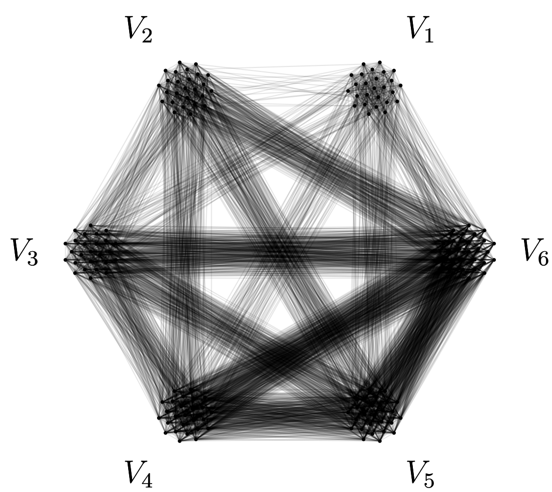

Figure 4.: A 'regular' partition.

1.1 Applications of Theorem 1.0.2 and the Counting Lemma

Many applications of Theorem 1.0.2 invoke a companion *Counting Lemma*, which in the context of Theorem 1.0.2 estimates the number of subgraphs of G which are partite-isomorphic to a given graph of a fixed isomorphism type. For this, it suffices to establish such a result for cliques of fixed size.

Lemma 1.1.1. *For all integers $\ell \geq 2$ and $d_0, \theta \in (0,1]$, there exist $\varepsilon = \varepsilon(\ell, d_0, \theta) \in (0,1]$ and $m_0 = m_0(\ell, d_0, \theta) \in \mathbb{N}$ so that the following holds. Let $F = (U, E(F))$ be an ℓ-partite graph with vertex ℓ-partition $U = U_1 \dot\cup \ldots \dot\cup U_\ell$ satisfying*

(i) $m \leq |U_i| \leq m+1$ for all $1 \leq i \leq \ell$, where $m \geq m_0$;

(ii) all pairs (U_i, U_j), $1 \leq i < j \leq \ell$, are (d_{ij}, ε)-regular with $d_{ij} \geq d_0$.

Then, the number of copies of K_ℓ in F is within the interval $(1 \pm \theta) m^\ell \prod_{1 \leq i < j \leq \ell} d_{ij}$.

In the context of Lemma 1.1.1, note that the number of cliques K_ℓ in F coincides with the expected such number in the corresponding random ℓ-partite environment. Indeed, with the partition $U_1 \dot\cup \ldots \dot\cup U_\ell$ of Lemma 1.1.1, consider the binomial random ℓ-partite graph $\mathbb{F} = \mathbb{G}[U_1, \ldots, U_\ell; \vec{d}\,]$, where $\vec{d} = (d_{ij} : \ 1 \leq i < j \leq \ell)$ is a sequence whereby for each $1 \leq i < j \leq \ell$, $\mathbb{F}$ independently includes each pair $\{u_i, u_j\}$ as an edge, for each $u_i \in U$ and $u_j \in U_j$, with probability $d_{ij} \in (0,1]$. Let $X_\ell = X_\ell(\mathbb{F})$ be the random variable counting the number of cliques K_ℓ in $\mathbb{F}$. Elementary probability establishes that

$$\mathbb{E}\big[X_\ell = \#\{K_\ell \subset \mathbb{F}\}\big] = m^\ell \prod_{1 \leq i < j \leq \ell} d_{ij},$$

and (more tediously) that $\mathrm{Var}[X_\ell] = O(m^{2\ell-1})$. The Chebyshev inequality [16] therefore ensures that

$$\mathbb{P}\big[X_\ell \neq (1 \pm \theta)\mathbb{E}[X_\ell]\big] = O\left(\frac{\mathrm{Var}[X_\ell]}{\mathbb{E}^2[X_\ell]}\right) = o(1) \to 0,$$

as $m \to \infty$, which Lemma 1.1.1 assimilates.

In the context of Theorem 1.0.2 and Lemma 1.1.1, the graph F would take the form of a subgraph $G[V_{i_1}, \ldots, V_{i_\ell}]$ induced on the (sub)partition $V_{i_1} \dot\cup \ldots \dot\cup V_{i_\ell}$, for some $1 \leq i_1 < \cdots < i_\ell \leq t$. Since Theorem 1.0.2 guarantees that at most εt^2 pairs (V_i, V_j), $1 \leq i < j \leq t$, can be ε-irregular, simple double-counting ensures that at most εt^ℓ many ℓ-tuples $(V_{i_1}, \ldots, V_{i_\ell})$, $1 \leq i_i < \cdots < i_\ell \leq t$, can host an ε-irregular pair (V_{i_j}, V_{i_k}), where $i_1 \leq i_j < i_k \leq i_\ell$. Consider the typical case when all

$\binom{\ell}{2}$ such pairs are ε-regular. When some pair (V_{i_j}, V_{i_k}), $i_1 \leq i_j < i_k \leq i_\ell$, has density below d_0, then $F = G[V_{i_1}, \ldots, V_{i_\ell}]$ hosts fewer than $d_0 m^\ell$ cliques K_ℓ, which is negligible when d_0 is small. Otherwise, when all pairs (V_{i_j}, V_{i_k}), $i_1 \leq i_j < i_k \leq i_\ell$, have density above d_0, Lemma 1.1.1 allows us to closely estimate the number of cliques K_ℓ in $F = G[V_{i_1}, \ldots, V_{i_\ell}]$, which from the discussion above assimilates the random case.

Applications of Theorem 1.0.2 with Lemma 1.1.1

While there are numerous applications of Theorem 1.0.2 with Lemma 1.1.1, we focus on the first such, and perhaps the most striking, which is due to Ruzsa and Szemerédi [31].

Theorem 1.1.2 (Triangle removal lemma)**.** *For all $\delta > 0$, there exist $c = c(\delta) > 0$ and integer $n_0 = n_0(\delta) \in \mathbb{N}$ so that every graph $G = (V, E)$ on $|V| = n \geq n_0$ many vertices which contains at most cn^3 triangles admits a triangle free subgraph $H \subseteq G$ where $|E \setminus E(H)| \leq \delta n^2$.*

Theorem 1.1.2 is easy to prove from Theorem 1.0.2 and Lemma 1.1.1.

Proof of Theorem 1.1.2 (Ruzsa and Szemerédi). Let $\delta > 0$ be given. To define the promised constant $c = c(\delta) > 0$, we define several auxiliary constants in terms of Lemma 1.1.1 and Theorem 1.0.2. To that end, set $d_0 = (1/6)\delta$ and $\theta = 1/2$. With $\ell = 3$, let

$$\varepsilon_{\text{Lem.1.1.1}} = \varepsilon_{\text{Lem.1.1.1}}(\ell = 3,\ d_0 = (1/6)\delta,\ \theta = 1/2)$$

be the constant guaranteed by Lemma 1.1.1. Let $t_0 \in \mathbb{N}$ be an integer satisfying

$$\varepsilon \stackrel{\text{def}}{=} \frac{1}{t_0} = \min\left\{\frac{1}{6}\delta,\ \varepsilon_{\text{Lem.1.1.1}}\right\}.$$

With $\varepsilon = 1/t_0$ above, let $T_0 = T_0(\varepsilon, t_0) \in \mathbb{N}$ be the integer constant guaranteed by Theorem 1.0.2. Finally, we define $c = d_0^3/(4T_0)^3$, which by the determinations above depends solely on the initial constant $\delta > 0$.

Let $G = (V, E)$ be a large n-vertex graph hosting at most cn^3 triangles K_3. With ε and t_0 above (where $\varepsilon = 1/t_0$), we apply Theorem 1.0.2 to G to obtain an ε-regular and t-equitable partition

$V = V_1 \cup \cdots \cup V_t$ of G, where $t_0 \leq t \leq T_0$. Now, form the subgraph $H \subseteq G$ by deleting all edges of G which

(i) reside within any single V_i, $1 \leq i \leq t$;

(ii) cross an ε-irregular pair (V_i, V_j), $1 \leq i < j \leq t$;

(iii) cross a pair (V_i, V_j) of density below d_0, $1 \leq i < j \leq t$.

Then we have deleted at most

$$t\binom{\lceil n/t \rceil}{2} + \left(\varepsilon t^2\right)\left\lceil \frac{n}{t} \right\rceil^2 + t_2\left(d_0 \left\lceil \frac{n}{t} \right\rceil^2\right) \leq \frac{1}{t}n^2 + 2\varepsilon n^2 + 2d_0 n^2$$
$$\leq \frac{1}{t_0}n^2 + 2\varepsilon n^2 + 2d_0 n^2 = 3\varepsilon n^2 + 2d_0 n^2 = \frac{1}{2}\delta n^2 + \frac{1}{3}\delta n^2 < \delta n^2$$

many edges to form the resulting graph H.

We claim that H is triangle-free. Indeed, assume for contradiction that $\{x, y, z\}$ is a triangle remaining in H. By the construction of H, there exist $1 \leq i_x < i_y < i_z \leq t$ so that $x \in V_{i_x}$, $y \in V_{i_y}$, and $z \in V_{i_z}$, where each pair (V_{i_x}, V_{i_y}), (V_{i_y}, V_{i_z}), and (V_{i_x}, V_{i_z}) is ε-regular with respective density $d_{i_x i_y}, d_{i_y i_z}, d_{i_x i_z} \geq d_0 > 0$. By Lemma 1.1.1, the subgraph $F = H[V_{i_x}, V_{i_y}, V_{i_z}]$ admits at least

$$(1-\theta)d_{i_x i_y} d_{i_y i_z} d_{i_x i_z} \left\lfloor \frac{n}{t} \right\rfloor^3 \geq (1-\theta)d_0^3 \left\lfloor \frac{n}{T_0} \right\rfloor^3 > \frac{d_0^3}{4T_0^3} n^3 = cn^3$$

many triangles, contradicting our hypothesis on G. □

At first sight, Theorem 1.1.2 seems innocuous: If G contains few triangles, surely one could delete few edges from G to destroy them all. Whatever its appearance, there is nothing elementary about the removal lemma, and no proof of it is known which does not in some way employ regularity and counting lemmas, as sketched above. (A version of this argument using a slightly weaker form of a regularity lemma was recently given by Fox [8], which appeared in *Annals of Mathematics*.) Moreover, Ruzsa and Szemerédi [31] artfully observed that Theorem 1.1.2 (which itself is not elementary) provides an elementary proof of a famous result of Roth [30] (which is the $k = 3$ case of Theorem 1.0.1).

Theorem 1.1.3 (Roth, 1956)**.** *Every subset $A = A_n \subseteq [n] = \{1, \ldots, n\}$ containing no 3-term arithmetic progression (AP_3) satisfies $|A| = o(n)$.*

Proof of Theorem 1.1.3 (Ruzsa and Szemerédi). Indeed, let $G = G_A = (X \dot\cup Y \dot\cup Z, E)$ be the 3-partite graph whose vertex set V consists of the (formally) disjoint sets $X = [n]$, $Y = [2n]$, and $Z = [3n]$, and whose edge set E includes, for each $x \in X$ and $a \in A$, all three edges of the triangle $x \in X$, $x + a \in Y$, and $x + 2a \in Z$. Then these are the only triangles of G, for if $x \in X$, $y \in Y$, and $z \in Z$, span a triangle, then

(i) $\{x, y\} \in E$ since $y = x + a$ for some $a \in A$;

(ii) $\{x, z\} \in E$ since $z = x + 2a'$ for some $a' \in A$;

(iii) $\{y, z\} \in E$ since $y = x' + a''$ and $z = x' + 2a''$ for some $x' in X$ and $a'' \in A$

so that $a' = \frac{a+a''}{2}$, in which case $a = a' = a''$ and $x = x'$, lest (a, a', a'') form an AP_3 in A. Now, Theorem 1.1.2 eliminates all $n|A| \leq n^2 = o(|V|^3)$ triangles of G by removing $o(n^2)$ of its edges. Since each removed edge eliminates a single triangle, G has $n|A| = o(n^2)$ many triangles, and so $|A| = o(n)$. $\square$

1.2 Algorithms for Theorem 1.0.2

Theorem 1.0.2 guarantees, for every graph $G = (V, E)$, the *existence* of a regular partition $V = V_1 \dot\cup \ldots \dot\cup V_t$. It does not, however, say how one could go about efficiently *constructing* such a partition. Moreover, the original proof of the regularity lemma was non-constructive. Since many applications employing Theorem 1.0.2 would admit constructive counterparts if the regularity lemma were also constructive, it became of interest to find a constructive proof of Theorem 1.0.2. Some 20 years after the original proof of the regularity lemma, a constructive version of it was established by Alon, Duke, Lefmann, Rödl, and Yuster [1, 2]. They proved that a regular partition $V = V_1 \dot\cup \ldots \dot\cup V_t$ of an n-vertex graph $G = (V, E)$ can be constructed in time $O(M(n))$, where $M(n) = O(n^{2.3727})$ is the time needed to multiply two $n \times n$ matrices with $\{0, 1\}$-entries over the integers (see [35]). Kohayakawa, Rödl, and Thoma [17, 18] improved this time to the best possible order of magnitude $O(n^2)$.

Theorem 1.2.1 (Algorithmic Szemerédi Regularity Lemma, [17, 18]). *There exists an algorithm $\mathbb{A}_{\mathrm{reg}}$ which, for all $\varepsilon > 0$ and integers $t_0 \in \mathbb{N}$, determines an integer $T_0 = T_0(\varepsilon, t_0)$ and constructs, for every given n-vertex graph $G = (V, E)$, an ε-regular and t-equitable partition $V = V_1 \dot\cup \ldots \dot\cup V_t$, where $t_0 \leq t \leq T_0$, in time $O(n^2)$.*

We next consider a straightforward application of Theorem 1.2.1 (with Lemma 1.1.1), taken from [6].

An easy application of Theorem 1.2.1

Consider the problem of enumerating the cliques K_ℓ of a given n-vertex graph $G = (V, E)$. For example, take $\ell = 3$, where we wish to enumerate the triangles K_3 of a given n-vertex graph $G = (V, E)$. The naive algorithm would enumerate these in time $O(n^3)$, but fast-matrix multiplication would do so much more quickly. Indeed, set $A = [a_{uv}]_{u,v \in V}$ to be the adjacency matrix of G, and set $B = [b_{uv}]_{u,v \in V} = A^2$. Then G admits precisely $(1/3)\sum_{\{u,v\} \in E} b_{uv}$ many triangles, where B can be computed in time $O(M(n)) = O(n^{2.3727})$ above. Nešetřil and Poljak [27] extended this approach to enumerate copies of any fixed clique K_ℓ of a given n-vertex graph $G = (V, E)$, where for simplicity in what follows we consider $\ell = 99$. They used fast-matrix multiplication to enumerate the cliques K_{99} of an n-vertex graph $G = (V, E)$ in time $O((M(n)^{33})$. This running time is not worse than $O(n^{79})$, but not guaranteed to be much better. By using Theorem 1.2.1 and Lemma 1.1.1, one can estimate the number of cliques K_{99} of G, up to an additive error of $o(n^{99})$, in optimal time $O(n^2)$. We shall sketch this approach by following the ideas of the proof given for Theorem 1.1.2.

Let $\delta > 0$ be fixed. Using similar choices to those given for Theorem 1.1.2, we select suitably small constants

$$\delta \gg d_0, \theta \gg \varepsilon = \frac{1}{t_0} > 0, \tag{1.3}$$

where $t_0 \in \mathbb{N}$. Let $G = (V, E)$ be a large n-vertex graph. With $\varepsilon > 0$ and $t_0 \in \mathbb{N}$ fixed above, we apply Theorem 1.2.1 to construct, in time $O(n^2)$, an ε-regular and t-equitable partition $V = V_1 \dot\cup \ldots \dot\cup V_t$ of G, where $t_0 \leq t \leq T_0 = T_0(\varepsilon, t_0)$. Now, for each $1 \leq i < j \leq t$, we greedily compute the density $d_G(V_i, V_j) = |E[V_i, V_j]|/(|V_i||V_j|)$ in time $O((n/t)^2)$, and we do so over all

$\binom{t}{2}$ many such pairs. We return the estimate

$$\#\{K_{99} \subset G\} \approx \left(\frac{n}{t}\right)^{99} \sum_{1 \le i_1 < \cdots < i_{99} \le t} \prod_{1 \le a < b \le 99} d_G(V_{i_a}, V_{i_b}) \tag{1.4}$$

for the number of cliques K_{99} in G, and we argue that (1.4) is within δn^{99} of being correct.

First, the estimate in (1.4) ignores copies of K_{99} having an edge entirely within a single V_i, $1 \le i \le t$. However, the number of such copies is at most

$$t\binom{\lceil n/t \rceil}{2} \times n^{97} \le \frac{2}{t} n^{99} \le \frac{2}{t_0} n^{99} \overset{(1.3)}{\ll} \delta n^{99}, \tag{1.5}$$

and so these contribute negligibly to the tolerated error. Second, the estimate in (1.4) likely miscounts copies of K_{99} crossing a 99-tuple $(V_{1_1}, \ldots, V_{i_{99}})$, $1 \le i_1 < \cdots < i_{99} \le t$, where some pair (V_{i_a}, V_{i_b}), $1 \le a < b \le 99$, is ε-irregular. However, an ε-regular partition admits at most εt^{99} such 99-tuples $(V_{1_1}, \ldots, V_{i_{99}})$, $1 \le i_1 < \cdots < i_{99} \le t$, which in turn can host at most

$$\varepsilon t^{99} \times \left\lceil \frac{n}{t} \right\rceil^{99} \le \varepsilon (2n)^{99} \overset{(1.3)}{\ll} \delta n^{99} \tag{1.6}$$

many such copies of K_{99}, again contributing negligibly to the tolerated error. Third, the estimate in (1.4) likely miscounts copies of K_{99} crossing a 99-tuple $(V_{1_1}, \ldots, V_{i_{99}})$, $1 \le i_1 < \cdots < i_{99} \le t$, where some pair (V_{i_a}, V_{i_b}), $1 \le a < b \le 99$, has density $d_G(V_{i_a}, V_{i_b}) \le d_0$. By definition, each such 99-tuple can host at most $d_0 \lceil n/t \rceil^{99}$ many crossing copies of K_{99}, and so all such 99-tuples can host at most

$$t^{99} \times d_0 \left\lceil \frac{n}{t} \right\rceil^{99} \le d_0 (2n)^{99} \overset{(1.3)}{\ll} \delta n^{99} \tag{1.7}$$

many such copies of K_{99}, again contributing negligibly to the tolerated error. Altogether, (1.5)–(1.7) combine to say that all but some $(\delta/2) n^{99}$ many copies of K_{99} cross 99-tuples $(V_{1_1}, \ldots, V_{i_{99}})$, $1 \le i_1 < \cdots < i_{99} \le t$, where all pairs (V_{i_a}, V_{i_b}), $1 \le a < b \le 99$, are $(d_G(V_{i_a}, V_{i_b}), \varepsilon)$-regular. These environments precisely match the hypothesis of Lemma 1.1.1, which guarantees that each such 99-tuple $(V_{1_1}, \dots, V_{i_{99}})$, $1 \le i_1 < \cdots < i_{99} \le t$, hosts within

$$(1 \pm \theta) \left(\frac{n}{t}\right)^{99} \prod_{1 \le a < b \le 99} d_G(V_{i_a}, V_{i_b}),$$

many copies of K_{99}. These errors $\theta(n/t)^{99}$, over all such 99-tuples $(V_{1_1}, \dots, V_{i_{99}})$, $1 \le i_1 < \cdots < i_{99} \le t$, total only

$$t^{99} \times \theta \left(\frac{n}{t}\right)^{99} \le \theta n^{99} \overset{(1.3)}{\ll} \delta n^{99},$$

which completes the proof.

1.3 Hypergraphs and regularity

In light of the many applications of the Szemerédi Regularity Lemma, a natural question one may ask is whether or not it extends to hypergraphs. For this, one would wish to prove a hypergraph regularity lemma which would be strong enough to guarantee a corresponding counting lemma, but still weak enough to apply to all (large) hypergraphs. This problem proved to be challenging.

A *hypergraph* $\mathcal{H} = (V, E)$ is an ordered pair where V is a (finite) set, and where $E \subseteq 2^V$ is a collection of subsets from V. We say that $\mathcal{H}$ is a *k-uniform hypergraph*, or *k-graph* for short, when $E \subseteq \binom{V}{k}$ is a family of k-element subsets of V. (Thus, a graph $G = (V, E)$ is a 2-graph.) One may naturally extend the concepts of 2-graph density and regularity to the following notions for k-graphs. For disjoint subsets $\emptyset \neq X_1, \dots, X_k \subset V$, we define $E[X_1, \dots, X_k] = E_{\mathcal{H}}[X_1, \dots, X_k]$ to be the set of k-tuples $\{x_1, \dots, x_k\} \in E$ with $x_i \in X_i$, for $1 \le i \le k$ and further define $K[X_1, \dots, X_k]$ to be the set of all k-tuples $x_1, \dots, x_k$, not necessarily in E, where $x_i \in X_i$ for $1 \le i \le k$. We write

$$d_{\mathcal{H}}(X_1, \dots, X_k) = \frac{|E[X_1, \dots, X_k]|}{|K[X_1, \dots, X_k]|} = \frac{|E[X_1, \dots, X_k]|}{|X_1| \cdots |X_k|}$$

for the *density* of $(X_1, \dots, X_k)$. For $\varepsilon > 0$, we say that $(X_1, \dots, X_k)$ is *ε-regular* if, for all $X_i' \subseteq X_i$ with $|X_i'| > \varepsilon |X_i|$, $1 \le i \le k$, we have

$$\left| d_{\mathcal{H}}\big(X_1', \dots, X_k'\big) - d_{\mathcal{H}}(X_1, \dots, X_k) \right| < \varepsilon. \tag{1.8}$$

We say that $(X_1, \ldots, X_k)$ is (d, ε)-*regular* when, in (1.8), we can replace $d_{\mathcal{H}}(X_1, \ldots, X_k)$ by d. Using the original proof of Szemerédi for graphs nearly verbatim, one may prove an ε-regularity lemma for k-uniform hypergraphs (see, e.g. [10]). However, ε-regularity for k-uniform hypergraphs, when $k \geq 3$, is not strong enough to admit a corresponding counting lemma, as we now sketch when $k = 3$.

Example 1 (Dense and regular but cliqueless hypergraphs, B. Nagle)**.** Let $\varepsilon > 0$ be given, and let $\rho \approx 0.68\ldots$ be the real root of $f(x) = x^3 + x - 1$. Let $n = |V|$ be a large integer divisible by 4, and let $V = V_1 \dot\cup V_2 \dot\cup V_3 \dot\cup V_4$ be an equipartition of V. There exists a 4-partite 3-uniform hypergraph $\mathcal{H} = (V, E)$ with vertex partition $V_1 \dot\cup V_2 \dot\cup V_3 \dot\cup V_4$ which satisfies

(i) for each $1 \leq i < j < k \leq 4$, (V_i, V_j, V_k) is (ρ, ε)-regular;

(ii) $\mathcal{H}$ has no cliques $K_4^{(3)}$, i.e. complete 3-uniform sub-hypergraphs on four vertices.

Proof (sketch) of Example 1. Fix $\varepsilon > 0$ and let ρ, n, and $V = V_1 \dot\cup V_2 \dot\cup V_3 \dot\cup V_4$ be given as in the hypothesis of Example 1. The desired hypergraph $\mathcal{H}$ will be a suitable instance of the following randomly constructed hypergraph $\mathbb{H}$ constructed as follows. For each $2 \leq i < j \leq 4$, let $\mathbb{G}(V_i, V_j; \rho)$ be the binomial random subgraph with edge-density ρ. For each $(v_1, v_i, v_j) \in V_1 \times V_i \times V_j$, include $\{v_1, v_i, v_j\} \in E(\mathbb{H})$ if, and only if, $\{v_i, v_j\} \in E(\mathbb{G}(V_i, V_j; \rho))$. For each $(v_2, v_3, v_4) \in V_2 \times V_3 \times V_4$, include $\{v_2, v_3, v_4\} \in E(\mathbb{H})$ if, and only if, $\{v_2, v_3, v_4\}$ is *not* a triangle (K_3) of

$$\mathbb{G}(V_2, V_3; \rho) \dot\cup \mathbb{G}(V_3, V_4; \rho) \dot\cup \mathbb{G}(V_2, V_4; \rho). \tag{1.9}$$

With certainty, the hypergraph $\mathbb{H}$ so constructed admits no copies of the clique $K_4^{(3)}$ (see Figure 5).

Extending the details of (1.2), the Chernoff inequality guarantees that, with probability $1 - o(1)$, where $o(1) \to 0$ as $n \to \infty$, each of (V_1, V_2, V_3), (V_1, V_3, V_4), and (V_1, V_2, V_4) is $(\rho, o(1))$-regular, and hence, is (ρ, ε)-regular. Similar details with Janson's inequality show that with probability $1 - o(1)$, all sizable subsets $V_2' \subseteq V_2$, $V_3' \subseteq V_3$, $V_4' \subseteq V_4$, span

$$(1 \pm o(1))\rho^3 |V_2'||V_3'||V_4'|$$

many triangles in (1.9). Thus, (V_2, V_3, V_4) is $(1 - \rho^3, o(1))$-regular, where $1 - \rho^3 = \rho$. Since all the

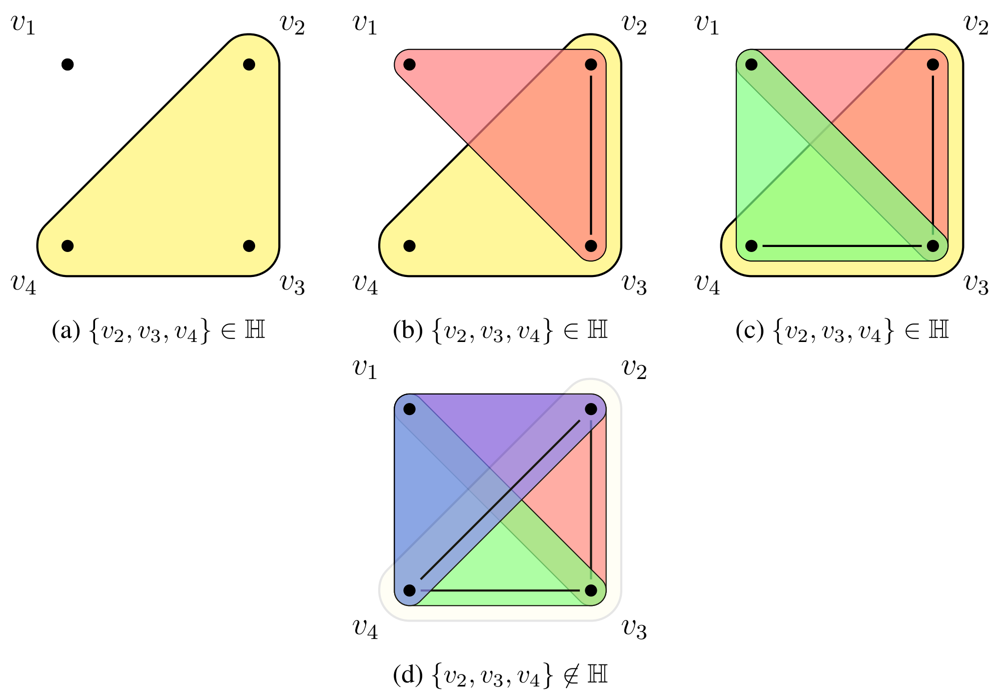

Figure 5.: $\mathbb{H}$ cannot have a $K_4^{(3)}$

desired properties of Example 1 hold for $\mathbb{H}$ with high (and hence positive) probability, there exists an instance $\mathcal{H}$ of $\mathbb{H}$ as desired in Example 1. $\square$

Strong hypergraph regularity

In light of Example 1, ε-regularity for k-graphs $\mathcal{H}$ is often referred to as *weak regularity*, because it does not admit a corresponding counting lemma. Over the years, *strong hypergraph regularity* and *counting lemmas* were, nonetheless, established, although they are quite technical in nature and we do not present them here. (Roughly speaking, these tools regularize, i.e. uniformly distribute, in a sparse way, a k-graph $\mathcal{H} = (V, E)$ with respect to the underlying set of $(k-1)$-tuples $\binom{V}{k-1}$, which are in turn regularized with respect to the underlying set of $(k-2)$-tuples $\binom{V}{k-2}$, and so on.) Frankl and Rödl [9] pioneered these efforts for $k = 3$ by proving a strong 3-uniform hypergraph regularity lemma, where Nagle and Rödl [23] proved a corresponding 3-uniform hypergraph counting lemma. These tools were later extended to k-uniform hypergraphs, for $k \geq 3$, by W.T. Gowers [11, 12] and by Nagle, Rödl, Schacht, and Skokan [24, 29]. These works prove a hypergraph removal lemma extending that of Theorem 1.1.2, and give quantitative proofs of the d-dimensional analogue of

Szemerédi's Density Theorem, among others. An algorithmic version of the *strong hypergraph regularity lemma* was recently established by Nagle, Rödl, and Schacht [25, 26], whereby a strong regular partition of an n-vertex k-graph $\mathcal{H} = (V, E)$ is constructed in time $O(n^{3k})$ (see also [14, 15]).

1.4 Main results of this dissertation

While an ε-regularity lemma for k-uniform hypergraphs is straightforward to prove, it is not easy to prove a constructive such version. Such an algorithm was established by Czygrinow and Rödl [4], and serves to inspire much of the work of this dissertation.

Theorem 1.4.1 (algorithmic weak regularity lemma)**.** *There exists an algorithm $\mathbb{A}_{\text{weak}}$ which, for all $\varepsilon > 0$, for all integers $k \geq 2$, and for all integers $t_0 \geq 1$, determines an integer $T_0 = T_0(\varepsilon, k, t_0)$ and constructs, for every given n-vertex k-graph $\mathcal{H} = (V, E)$, in time $O(n^{2k-1} \log^2 n)$, a partition $V = V_1 \dot{\cup} \ldots \dot{\cup} V_t$, with $t_0 \leq t \leq T_0$, satisfying that*

1. *$V = V_1 \dot{\cup} \ldots \dot{\cup} V_t$ is t-*equitable*: $|V_1| \leq \cdots \leq |V_t| \leq |V_1| + 1$;*
2. *$V = V_1 \dot{\cup} \ldots \dot{\cup} V_t$ is ε-*regular*: all but εt^k many $(V_{i_1}, \ldots, V_{i_k})$, $1 \leq i_1 < \cdots < i_k \leq t$, are ε-regular.*

Our first main result optimizes the running time of Theorem 1.4.1 in the case when $k = 3$.

Theorem 1.4.2 (Main Result I)**.** *There exists an algorithm $\mathbb{A}^{(3)}_{\text{weak}}$ which, for all $\varepsilon > 0$ and for all integers $t_0 \geq 1$, determines an integer $T_0 = T_0(\varepsilon, t_0)$ and constructs, for every given n-vertex 3-graph $\mathcal{H} = (V, E)$, in time $O(n^3)$, an ε-regular and t-equitable partition $V = V_1 \dot{\cup} \ldots \dot{\cup} V_t$, with $t_0 \leq t \leq T_0$.*

To our knowledge, Theorem 1.4.2 is the first hypergraph regularity lemma which is guaranteed to have runtime the best possible order of magnitude $O(n^3)$.

To prove Theorem 1.4.2, we actually prove a stronger result in upcoming Theorem 5.0.5, but which is too technical to state in this introduction. To give a flavor of Theorem 5.0.5, we prove a regularity lemma for 3-graphs $\mathcal{H} = (V, E)$ which provides a stronger notion of regularity than

Theorems 1.4.1 and 1.4.2, and from which we derive Theorem 1.4.2 in a fairly standard way. Recall that the result of Nagle, Rödl, and Schacht [25, 26] would give a *strong regular partition* of a 3-graph $\mathcal{H} = (V, E)$ in time $O(n^9)$, which is too slow for our purposes. Our second main result proves an algorithmic regularity lemma for 3-graphs $\mathcal{H} = (V, E)$ which is weaker but much faster than that of Nagle, Rödl, and Schacht [25, 26], but stronger and faster than that of Czygrinow and Rödl [4]. We therefore call this result the *Medium Regularity Lemma*, which we informally state below.

Theorem 1.4.3 (Main Result II)**.** *There exists an algorithm $\mathbb{A}^{(3)}_{\mathrm{med}}$ which, for every given n-vertex 3-graph $\mathcal{H} = (V, E)$, constructs in time $O(n^3)$ a 'mediumly-regular' partition of $\mathcal{H}$.*

For a precise statement of this result, see Theorem 5.0.5. We next turn to applications of our work, which are in progress with T. Molla and B. Nagle.

Applications of our work

We consider two ongoing applications of the work in this dissertation, which are joint with T. Molla and B. Nagle [22] Our first result is an application of Theorem 1.4.2, and our second result is an application of Theorem 1.4.3.

Call a hypergraph $\mathcal{H} = (V, E)$ *2-colorable* if there exists a partition $V = A \dot\cup B$ where both A and B meet every edge of $\mathcal{H}$, and call such a partition $V = A \dot\cup B$ a *2-coloring* of H. Elementary graph theory establishes that one may determine whether or not a graph $G = (V, E)$ is 2-colorable in time $O(|V|)$, and when so, one may construct a 2-coloring $V = A \dot\cup B$ in this same time. However, as is well-known from the work of Lovász [21], the same problem for 3-graphs is NP-complete. Moreover, Guruswami, Håstad and Sudan [13] showed that it is NP-hard to properly color a 2-colorable 4-graph $\mathcal{H} = (V, E)$ with any constant number of colors. (A *proper coloring* of $\mathcal{H} = (V, E)$ is a coloring of V leaving no edge of $\mathcal{H}$ monochromatic.)

A recent work of Person and Schacht [28] shows that one may 2-color an n-vertex 2-colorable 3-graph $\mathcal{H} = (V, E)$ in average running time $O(n^5 \log^2 n)$. An important part of their proof employs Theorem 1.4.1. In our current work with T. Molla [22], we use Theorem 1.4.2 to optimize their running time.

Theorem 1.4.4 (Molla, Nagle, Theado (in progress))**.** *There exists an algorithm $\mathbb{A}_{\mathrm{bip}}$ which, with average running time $O(n^3)$, 2-colors an n-vertex 2-colorable 3-graph $\mathcal{H} = (V, E)$. In particular, if $\mathbb{H}_n$ is a 2-colorable n-vertex 3-graph chosen uniformly at random among all 2-colorable 3-graphs on vertex set $\{1, \ldots, n\}$, then with probability $1 - o(1)$, where $o(1) \to 0$ as $n \to \infty$, the algorithm $\mathbb{A}_{\mathrm{bip}}$ 2-colors $\mathbb{H}_n$ in time $O(n^3)$.*

Recall the earlier application of Theorem 1.0.2, whereby one estimates the frequency of a fixed clique K_ℓ in a given n-vertex graph $G = (V, E)$ up to an additive error of $o(n^\ell)$ and in time $O(n^2)$. Using the strong hypergraph regularity algorithm of Nagle, Rödl, and Schacht [25, 26], one may extend this algorithm to estimate the frequency of a fixed clique $K_\ell^{(k)}$ in a given n-vertex k-graph $\mathcal{H} = (V, E)$ up to an additive error of $o(n^\ell)$ and in time $O(n^{3k})$. It would be of interest to know whether such an algorithm could run in optimal time $O(n^k)$, but the work in [25, 26] doesn't give it. While Theorem 1.4.3 is not strong enough to make an improvement here, it does provide an improvement on a weaker problem.

Consider the 3-graph $K_4^{(3)} - e$ consisting of three triples on four points. We consider the problem of estimating the frequency of $K_4^{(3)} - e$ in a given n-vertex k-graph $\mathcal{H} = (V, E)$. If we could count *isomorphic* copies of $K_4^{(3)} - e$ in $\mathcal{H} = (V, E)$, or equivalently *induced* copies of $K_4^{(3)} - e$ in $\mathcal{H}$, then we could also count copies of $K_4^{(3)}$ in $\mathcal{H}$, which Theorem 1.4.3 won't give. However, if we only seek *not-necessarily induced* copies of $K_4^{(3)} - e$ in $\mathcal{H}$ (meaning we can't distinguish when such a copy corresponds to $K_4^{(3)} - e$ or $K_4^{(3)}$), then Theorem 1.4.3 can give this, which is ongoing work with T. Molla and B. Nagle [22].

Theorem 1.4.5 (Molla, Nagle, Theado (in progress))**.** *There exists an algorithm $\mathbb{A}_{\mathrm{freq}}$ which, for all $\delta > 0$, estimates the not-necessarily induced frequency of $K_4^{(3)} - e$ in a given n-vertex 3-graph $\mathcal{H} = (V, E)$ up to an additive error of δn^4 and in time $O(n^3)$.*

1.5 Itinerary of Dissertation

In Chapter 2, we prove an algorithm, $\mathbb{A}_{\mathrm{sparse}}$, for sparse bipartite graphs L, which efficiently confirms whether L is suitably regular in one sense, or efficiently constructs witnesses of a significant irregularity in another sense. These notations will be made precise in Chapter 2 where $\mathbb{A}_{\mathrm{sparse}}$

provides important underpinnings for future considerations.

In Chapter 3, we define several important graph and 3-graph concepts, including that of *triads*, $\mathcal{H}$-*triads*, and *links*, which provide the bedrock for all results in this dissertation. We also describe ways in which these objects can be regular and discuss witnesses of irregularity when they are not. These concepts form an analogue of the ε-regular pair (V_i, V_j) from Szemerédi's regularity lemma (Theorem 1.0.2).

In Chapter 4, we prove an algorithm $\mathbb{A}_{\text{link}}$ which efficiently confirms when an $\mathcal{H}$-triad is suitably regular in one sense, or efficiently constructs witnesses of significant irregularity in another sense. These notions will be made precise in Chapter 4, where $\mathbb{A}_{\text{link}}$ may be considered the center of this dissertation.

In Chapter 5, we use the results of our previous chapters to prove the algorithm $\mathbb{A}_{\text{med}}$ referenced in Theorem 1.4.3, which is one of our main results. We reference $\mathbb{A}_{\text{med}}$ more precisely therein as $\mathbb{A}_{\text{med}} = \mathbb{A}_{\text{linkreg}}$.

In Chapter 6, we prove a transference lemma which allows us to infer the notion of regularity in Theorem 1.4.2 from that of Theorem 1.4.3.

Finally, in Chapter 7, we use this lemma to prove Theorem 1.4.2.

Chapter 2
Sparse graphs

The goal of this chapter is to prove an algorithm $\mathbb{A}_{\text{sparse}}$ for sparse bipartite graphs (2-graphs) L which efficiently confirms whether L is suitably regular in one sense, or efficiently constructs witnesses of a significant irregularity in another sense. In particular, and in the context of Szemerédi's regularity lemma (Theorem 1.0.2), when (A', B') is a witness of irregularity for (A, B), it is a single such witness. Singular witnesses were sufficient for Szemerédi's proof of his regularity lemma. In this dissertation, we work with graphs and hypergraphs which are exceedingly sparse. As such, our witnesses will often need to be simultaneous systems of $r \geq 1$ many objects. Frequently, r will be a very large constant not depending on the size of the (hyper)graph. We now make these notions precise.

To that purpose, let $L = (A \dot\cup B, E)$ be a bipartite graph with vertex bipartition $V(L) = A \dot\cup B$. For an integer $r \in \mathbb{N}$, let $\vec{\mathcal{A}} = (A_1, \ldots, A_r)$ be a sequence of subsets $A_1, \ldots, A_r \subset A$. Then $\vec{\mathcal{A}}$ is an element of the r-fold Cartesian product $2^A \times \ldots \times 2^A = (2^A)^r$. Similarly, let $\vec{\mathcal{B}} = (B_1, \ldots, B_r) \in (2^B)^r$. We define

$$E[\vec{\mathcal{A}}, \vec{\mathcal{B}}] = E[A_1, B_1] \cup \ldots \cup E[A_r, B_r] = \bigcup_{i=1}^{r} E[A_i, B_i]$$

to be the set of edges $\{a, b\} \in E$ for which there exists $1 \leq i \leq r$ such that $a \in A_i$ and $b \in B_i$. Similarly, we define

$$K[\vec{\mathcal{A}}, \vec{\mathcal{B}}] = \bigcup_{i=1}^{r} K[A_i, B_i]$$

to be the set of all pairs $\{a, b\} \in K[A, B]$ for which there exists $1 \leq i \leq r$ such that $a \in A_i$ and $b \in B_i$. We say that the pair $(\vec{\mathcal{A}}, \vec{\mathcal{B}})$ is *pair-disjoint* if for $1 \leq i < j \leq r$, the edge sets $K[A_i, B_i]$, $K[A_j, B_j]$ are disjoint. Also, for $c > 0$ we say that $(\vec{\mathcal{A}}, \vec{\mathcal{B}})$ is *c-bounded* if for each $1 \leq i, j \leq r$

where $A_i, B_j \neq \emptyset$ we have $|A_i| \geq c|A|$ and $|B_j| \geq c|B|$. Finally, we define the r-*density* $d_L(\vec{\mathcal{A}}, \vec{\mathcal{B}})$ by

$$d_L(\vec{\mathcal{A}}, \vec{\mathcal{B}}) = \frac{|E[\vec{\mathcal{A}}, \vec{\mathcal{B}}]|}{|K[\vec{\mathcal{A}}, \vec{\mathcal{B}}]|} = \frac{|\bigcup_{i=1}^r E[A_i, B_i]|}{|\bigcup_{i=1}^r K[A_i, B_i]|}$$

Definition 2.0.1. Let $L = (A \dot\cup B, E)$ be a bipartite graph, as above, and let $d \in [0, 1]$, $\delta > 0$ and $r \in \mathbb{N}$ be given. We say that L is (d, δ, r)-*regular* if for any $\vec{\mathcal{A}} = (A_1, \ldots, A_r) \in (2^A)^r$ and $\vec{\mathcal{B}} = (B_1, \ldots, B_r) \in (2^B)^r$ satisfying

$$\left|K[\vec{\mathcal{A}}, \vec{\mathcal{B}}]\right| \geq \delta|K[A, B]| = \delta|A||B|, \tag{2.1}$$

we have

$$d(1 - \delta) \leq d_L(\vec{\mathcal{A}}, \vec{\mathcal{B}}) \leq d(1 + \delta). \tag{2.2}$$

Otherwise, we say that L is (d, δ, r)-*irregular* and any pair of sequences $(\vec{\mathcal{A}}, \vec{\mathcal{B}}) \in (2^A)^r \times (2^B)^r$ satisfying (2.1) but not (2.2) is called an r-*witness* of the (d, δ, r)-irregularity of L. Note that any s-witness to the (d, δ, s)-irregularity of L for $1 \leq s \leq r$ serves as an r-witness to the (d, δ, r)-irregularity of L as we may form an r-tuple from the s-tuple by adding empty sets and doing so does not affect the density.

Lemma 2.0.2 ($\mathbb{A}_{\text{sparse}}$)**.** *There exists an algorithm $\mathbb{A}_{\text{sparse}}$ so that for any $\delta > 0$, $\mathbb{A}_{\text{sparse}}$ determines $\delta_* > 0$ so that for any $D \in [0, 1]$, $\mathbb{A}_{\text{sparse}}$ determines $c > 0$ and $r = r(\delta, \delta_*, D) \in \mathbb{N}$ so that the following holds: Let $L = (A \dot\cup B, E)$ be a bipartite graph of density $d_L(A, B) = d \geq D$ on $n = |A| + |B|$ vertices where $|A|, |B| = \Theta(n)$. Then in time $O(n^2)$, $\mathbb{A}_{\text{sparse}}$ either confirms that L is $(d, \delta, 1)$-regular, or it constructs a c-bounded and pair-disjoint r-witness $(\vec{\mathcal{A}}, \vec{\mathcal{B}}) \in (2^A)^r \times (2^B)^r$ of the (d, δ_*, r)-irregularity of L.*

2.1 The Algorithm $\mathbb{A}_{\text{sparse}}$

Input: Let $\delta > 0$ be given. For convenience in upcoming considerations, define an auxiliary constant δ_0 by

$$\delta_0 = \frac{\delta^2}{22}, \text{ and then define } \delta_* = \frac{\delta_0}{4} = \frac{\delta^2}{88}. \tag{2.3}$$

Let $D \in [0,1]$ be given. To define the promised integer r, we first define an auxiliary constant

$$\varepsilon = D\delta_0. \tag{2.4}$$

With $\varepsilon > 0$ above and with $t_0 = 1$, let

$$T_0 = T_0(\varepsilon) \tag{2.5}$$

be the constant guaranteed by Theorem 1.2.1. Set

$$c = \frac{1}{2T_0} \text{ and } r = T_0^2 \tag{2.6}$$

Let L be a bipartite graph of density $d_L(A,B) = d \geq D$ on $n = |A| + |B|$ vertices, where $|A|, |B| = \Theta(n)$.

Procedure: The algorithm $\mathbb{A}_{\text{sparse}}$ proceeds along the following steps.

Step 1. (Apply $\mathbb{A}_{\text{Szem}}$) With $\varepsilon > 0$ given in (2.4), and $t_0 = 1$, apply the algorithm $\mathbb{A}_{\text{Szem}}$ (see Theorem 1.2.1) to the graph L to obtain in time $O(n^2)$, an ε-regular t-equitable partition

$$A = A_0 \,\dot\cup\, A_1 \,\dot\cup \ldots \dot\cup\, A_{t_1}, \quad B = B_0 \,\dot\cup\, B_1 \,\dot\cup \ldots \dot\cup\, B_{t_2},$$

where $t_0 \leq t = t_1 + t_2 \leq T_0$.

Step 2. (Compute densities) For each $(i,j) \in [t_1] \times [t_2]$, (greedily) compute the density $d_L(A_i, B_j)$ in time $O(|A_i||B_j|)$. Repeating over all $(i,j) \in [t_1] \times [t_2]$, the set of all densities $d_L(A_i, B_j)$ is recorded in time

$$\sum_{(i,j)\in[t_1]\times[t_2]} O(|A_i||B_j|) = O(|A||B|) = O(n^2).$$

Step 3. (Check uniformity of densities) Define

$$
\begin{aligned}
\Delta^- &= \{(i,j) \in [t_1] \times [t_2] : d_L(A_i, B_j) < d(1-\delta_0)\} \\
\Delta^+ &= \{(i,j) \in [t_1] \times [t_2] : d_L(A_i, B_j) > d(1+\delta_0)\} \\
\Delta \; &= ([t_1] \times [t_2]) \setminus (\Delta^- \cup \Delta^+)
\end{aligned}
$$

In time $O(t_1 t_2) = O(T_0) = O(1)$, decide among the following cases:

$$
\begin{array}{lcl}
\textbf{Case 1}: \quad |\Delta| & \geq & (1-2\delta_0)t_1 t_2, \\
\textbf{Case 2}: \quad |\Delta^-| & \geq & \delta_0 t_1 t_2, \\
\textbf{Case 3}: \quad |\Delta^+| & \geq & \delta_0 t_1 t_2,
\end{array}
$$

While these cases are not necessarily mutually disjoint, at least one of the above cases holds since

$$
|\Delta| + |\Delta^-| + |\Delta^+| = |\Delta \,\dot\cup\, \Delta^- \,\dot\cup\, \Delta^+| = t_1 t_2 = (1-2\delta_0)t_1 t_2 + \delta_0 t_1 t_2 + \delta_0 t_1 t_2.
$$

Output: If $\mathbb{A}_{\text{sparse}}$ determines that Case 1 holds, then the algorithm will return that L is $(d, \delta, 1)$-regular. Otherwise, $\mathbb{A}_{\text{sparse}}$ will report that L is not (d, δ_*, r)-regular and return an r-witness $(\vec{\mathcal{A}}, \vec{\mathcal{B}})$. Since either Case 2 or Case 3 hold, respectively either write Δ^- or Δ^+ as $\{(i_1, j_1), \ldots, (i_p, j_p)\}$ and then use $\vec{\mathcal{A}} = (A_{i_1}, \ldots, A_{i_p})$ and $\vec{\mathcal{B}} = (B_{i_1}, \ldots, B_{i_p})$ for the r-witness (recall (2.6)).

It is clear that $\mathbb{A}_{\text{sparse}}$ terminates in time $O(n^2)$, where this time is achieved in steps 1 and 2. It remains to verify that the output is indeed correct and we do so by examining each of the cases in Step 3.

2.2 Proof of Correctness

To prove the correctness of the algorithm $\mathbb{A}_{\text{sparse}}$, we consider the cases determined in the above procedure and verify that the corresponding output is indeed correct.

2.2.1 Case 1

Assume that $|\Delta| \geq (1-2\delta_0)t_1t_2$ and we show that L is $(d,\delta,1)$-regular. To that end, let $\vec{\mathcal{A}} = (A') \in 2^A$ and $\vec{\mathcal{B}} = (B') \in 2^B$ satisfy

$$|K(\vec{\mathcal{A}},\vec{\mathcal{B}})| = |K[A',B']| = |A'||B'| \geq \delta|A||B| = \delta|K[A,B]|.$$

We show

$$d(1-\delta) \leq d_L(\vec{\mathcal{A}},\vec{\mathcal{B}}) \leq d(1+\delta), \tag{2.7}$$

where

$$d_L(\vec{\mathcal{A}},\vec{\mathcal{B}}) = \frac{|E_L[\vec{\mathcal{A}},\vec{\mathcal{B}}]|}{|K[\vec{\mathcal{A}},\vec{\mathcal{B}}]|} = \frac{|E_L[A',B']|}{|K[A',B']|} = \frac{|E_L[A',B']|}{|A'||B'|}.$$

In particular, the proofs of the lower and upper bounds in (2.7) are very similar. For simplicity in what follows, we prove the lower bound only.

Let $A'_i = A' \cap A_i$ for $0 \leq i \leq t_1$ and $B'_j = B' \cap B_j$ for $0 \leq j \leq t_2$. Define

$$\begin{aligned}
\Delta_{\text{reg}} &= \left\{(i,j) \in \Delta : (A_i,B_j) \text{ is } \varepsilon\text{-regular}\right\}, \\
\Delta_{\text{big}} &= \left\{(i,j) \in \Delta : |A'_i| > \varepsilon|A_i| \text{ and } |B'_j| > \varepsilon|B_j|\right\}, \\
\Delta_{\text{good}} &= \Delta_{\text{reg}} \cap \Delta_{\text{big}}.
\end{aligned} \tag{2.8}$$

Since $\Delta_{\text{good}} \subset \Delta \subset [t_1] \times [t_2]$ we have

$$|E_L[A',B']| = \left| \dot{\bigcup_{\substack{0\leq i\leq t_1 \\ 0\leq j\leq t_2}}} E_L[A'_i,B'_j] \right| \geq \sum_{(i,j)\in\Delta_{\text{good}}} |E_L[A'_i,B'_j]|. \tag{2.9}$$

Fix $(i,j) \in \Delta_{\text{good}} = \Delta \cap \Delta_{\text{reg}} \cap \Delta_{\text{big}}$. Then the following hold:

(i) Since $(i,j) \in \Delta_{\text{big}}$, we have that $|A'_i| > \varepsilon|A_i|$ and $|B'_j| > \varepsilon|B_j|$;

(ii) Since $(i,j) \in \Delta_{\text{reg}}$, we have that (A_i,B_j) is ε-regular;

(iii) Since $(i,j) \in \Delta$, we have that $d(1-\delta_0) \le d_L(A_i, B_j) \le d(1+\delta_0)$.

From (ii) we get $|d_L(A'_i, B'_j) - d_L(A_i, B_j)| < \varepsilon$ so that by (iii),

$$\frac{|E_L[A'_i, B'_j]|}{|A'_i||B'_j|} = d_L(A'_i, B'_j) > d_L(A_i, B_j) - \varepsilon \ge d(1-\delta_0) - \varepsilon$$

and hence from this and (2.9),

$$|E_L[A', B']| \ge (d(1-\delta_0) - \varepsilon) \sum_{(i,j)\in\Delta_{\text{good}}} |A'_i||B'_j|. \tag{2.10}$$

Now we look at $|A'||B'|$, observing that

$$|A'||B'| = |K[A', B']| = \left| \dot{\bigcup_{\substack{0\le i\le t_1\\ 0\le j\le t_2}}} K[A'_i, B'_j] \right| \le T_0 n + \sum_{(i,j)\in[t_1]\times[t_2]} |A'_i||B'_j|$$

follows from $|A_0 \cup B_0| = t \le T_0 = O(1)$. Recalling the partition $\Delta \,\dot\cup\, \Delta^- \,\dot\cup\, \Delta^+$ (cf. Step 3), we see

$$\begin{aligned} |A'||B'| &\le O(n) + \sum_{(i,j)\in\Delta^-\cup\Delta^+} |A'_i||B'_i| + \sum_{(i,j)\in\Delta} |A'_i||B'_i| \\ &\le O(n) + \sum_{(i,j)\in\Delta^-\cup\Delta^+} |A_i||B_i| + \sum_{(i,j)\in\Delta} |A'_i||B'_i| \\ &\le O(n) + \left|\Delta^-\cup\Delta^+\right| \left\lfloor \frac{|A|}{t_1} \right\rfloor \left\lfloor \frac{|B|}{t_2} \right\rfloor + \sum_{(i,j)\in\Delta} |A'_i||B'_i|. \end{aligned}$$

Since $|\Delta| \ge (1-2\delta_0)t_1t_2$, we have $|\Delta^-\cup\Delta^+| \le 2\delta_0 t_1 t_2$, and so

$$|A'||B'| \le O(n) + 2\delta_0|A||B| + \sum_{(i,j)\in\Delta} |A'_i||B'_i|. \tag{2.11}$$

Now define

$$\begin{aligned}\Delta_{\text{irreg}} &= \Delta \setminus \Delta_{\text{reg}} = \{(i,j) \in \Delta : (A_i, B_j) \text{ is } \varepsilon\text{-irregular}\}, \\ \Delta_{\text{small}} &= \Delta \setminus \Delta_{\text{big}} = \left\{(i,j) \in \Delta : |A'_i| \le \varepsilon|A_i| \text{ and } |B'_j| \le \varepsilon|B_j|\right\}, \\ \Delta_{\text{bad}} &= \Delta \setminus \Delta_{\text{good}} = \Delta_{\text{irreg}} \cup \Delta_{\text{small}}.\end{aligned}$$

We have

$$\begin{aligned}\sum_{(i,j)\in\Delta} |A'_i||B'_j| &= \sum_{(i,j)\in\Delta_{\text{good}}} |A'_i||B'_j| + \sum_{(i,j)\in\Delta_{\text{bad}}} |A'_i||B'_j| \\ &\le \sum_{(i,j)\in\Delta_{\text{good}}} |A'_i||B'_j| + \sum_{(i,j)\in\Delta_{\text{irreg}}} |A'_i||B'_j| + \sum_{(i,j)\in\Delta_{\text{small}}} |A'_i||B'_j| \\ &\le |\Delta_{\text{irreg}}| \left\lfloor \frac{|A|}{t_1} \right\rfloor \left\lfloor \frac{|B|}{t_2} \right\rfloor + \sum_{(i,j)\in\Delta_{\text{small}}} |A'_i||B'_j| + \sum_{(i,j)\in\Delta_{\text{good}}} |A'_i||B'_j|\end{aligned}$$

By the application of $\mathbb{A}_{\text{Szem}}$, we have that $|\Delta_{\text{irreg}}| \le \varepsilon t_1 t_2$. Moreover, for each $(i,j) \in \Delta_{\text{small}}$ it follows (by definition of Δ_{small}) that $|A'_i||B'_j| \le \varepsilon|A_i||B_j|$ where $|\Delta_{\text{small}}| \le t_1 t_2$. Thus,

$$\sum_{(i,j)\in\Delta} |A'_i||B'_j| \le 2\varepsilon|A||B| + \sum_{(i,j)\in\Delta_{\text{good}}} |A'_i||B'_j|. \tag{2.12}$$

Combining (2.11) and (2.12), we see

$$\begin{aligned}|A'||B'| &\le O(n) + 2\delta_0|A||B| + 2\varepsilon|A||B| + \sum_{(i,j)\in\Delta_{\text{good}}} |A'_i||B'_j| \\ &\le 5\delta_0|A||B| + \sum_{(i,j)\in\Delta_{\text{good}}} |A'_i||B'_j| \end{aligned} \tag{2.13}$$

where we used $\varepsilon \le \delta_0$ in (2.4) and $O(n) = o(|A||B|)$ (since $|A|+|B| = n$, where $|A|, |B| = \Theta(n)$). Comparing (2.10) and (2.13), we see that

$$\frac{|E_L[A', B']|}{|A'||B'|} \ge \frac{(d(1-\delta_0)-\varepsilon)\sum_{(i,j)\in\Delta_{\text{good}}} |A'_i||B'_j|}{5\delta_0|A||B| + \sum_{(i,j)\in\Delta_{\text{good}}} |A'_i||B'_j|} = (d(1-\delta_0)-\varepsilon)\,\frac{1}{1+x},$$

where

$$x = 5\delta_0 \frac{|A||B|}{\sum_{(i,j)\in\Delta_{\text{good}}} |A'_i||B'_j|} \overset{(2.13)}{\leq} 5\delta_0 \frac{|A||B|}{|A'||B'| - 5\delta_0|A||B|} \leq 5\delta_0 \frac{|A||B|}{\delta|A||B| - 5\delta_0|A||B|}$$
$$\leq \frac{5\delta_0}{\delta - 5\delta_0},$$

where we used the hypothesis that $|A'||B'| \geq \delta|A||B|$. Therefore,

$$\begin{aligned}\frac{E_L[A', B']}{|A'||B'|} &\geq (d(1-\delta_0) - \varepsilon)\frac{1}{1 + \frac{5\delta_0}{\delta - 5\delta_0}} \\ &\geq d\left(1 - \delta_0 - \frac{\varepsilon}{D}\right)\left(1 - \frac{5\delta_0}{\delta - 5\delta_0}\right) \\ &\overset{(2.3)\,(2.4)}{\geq} d(1-2\delta_0)\left(1 - \frac{10\delta_0}{\delta}\right) \\ &\geq d\left(1 - 22\frac{\delta_0}{\delta}\right) \\ &\overset{(2.3)}{\geq} d(1-\delta),\end{aligned}$$

as desired.

2.2.2 Cases 2 and 3

The proofs of correctness for each of cases 2 and 3 are identical, so we focus only on Case 2. To that end, assume

$$p = |\Delta^-| \geq \delta_0 t_1 t_2, \tag{2.14}$$

where $\Delta^- = \{(i_1, j_1), \ldots, (i_p, j_p)\}$ and $\vec{\mathcal{A}} = (A_{i_1}, \ldots, A_{i_p})$ and $\vec{\mathcal{B}} = (B_{i_1}, \ldots, B_{i_p})$. Since $p \leq t_1 t_2 \leq T_0^2 = r$ (cf. (2.6)), the sequences $\vec{\mathcal{A}}$ and $\vec{\mathcal{B}}$ can be viewed as r-tuples with possibly empty coordinates. Clearly, $(\vec{\mathcal{A}}, \vec{\mathcal{B}})$ is pair-disjoint. Moreover, $(\vec{\mathcal{A}}, \vec{\mathcal{B}})$ is c-bounded as whenever A_i, B_j are non-empty coordinates, Step 1 gives

$$|A_i| \geq \frac{1}{t}(|A| - t) \geq \frac{1}{2T_0}|A| = c|A|.$$

We show, in fact, that under the conditions of Case 2, the pair $(\vec{\mathcal{A}}, \vec{\mathcal{B}})$ is an r-witness to the (d, δ_*, r)-irregularity of L. For that, we show

$$\left|K[\vec{\mathcal{A}}, \vec{\mathcal{B}}]\right| > \delta_* \left|K[A, B]\right| = \delta_* |A||B| \tag{2.15}$$

and

$$d_L(\vec{\mathcal{A}}, \vec{\mathcal{B}}) < d(1 - \delta_*). \tag{2.16}$$

Proof of (2.15). Recall that

$$K[\vec{\mathcal{A}}, \vec{\mathcal{B}}] = K[A_{i_1}, B_{j_1}] \dot\cup \ldots \dot\cup K[A_{i_p}, B_{j_p}] \tag{2.17}$$

Thus,

$$\begin{aligned} \left|K[\vec{\mathcal{A}}, \vec{\mathcal{B}}]\right| &= \sum_{k=1}^{p} |K[A_{i_k}, B_{j_k}]| = \sum_{k=1}^{p} |A_{i_k}||B_{j_k}| \\ &= p \left\lfloor \frac{|A|}{t_1} \right\rfloor \left\lfloor \frac{|B|}{t_2} \right\rfloor \\ &\overset{(2.14)}{\geq} \frac{1}{4}\delta_0 |A||B| \overset{(2.3)}{\geq} \delta_* |A||B|, \end{aligned} \tag{2.18}$$

as desired. □

Proof of (2.16). Recall that

$$d_L(\vec{\mathcal{A}}, \vec{\mathcal{B}}) = \frac{|E_L[\vec{\mathcal{A}}, \vec{\mathcal{B}}]|}{|K[\vec{\mathcal{A}}, \vec{\mathcal{B}}]|},$$

where similarly to (2.17),

$$E_L[\vec{\mathcal{A}}, \vec{\mathcal{B}}] = E_L[A_{i_1}, B_{j_1}] \dot\cup \ldots \dot\cup E_L[A_{i_p}, B_{j_p}]$$

is a pairwise disjoint union. As such,

$$d_L(\vec{\mathcal{A}}, \vec{\mathcal{B}}) = \frac{|E_L[\vec{\mathcal{A}}, \vec{\mathcal{B}}]|}{|K[\vec{\mathcal{A}}, \vec{\mathcal{B}}]|} = \frac{\sum_{k=1}^{p} |E_L[A_{i_k}, B_{j_k}]|}{|K[\vec{\mathcal{A}}, \vec{\mathcal{B}}]|}. \tag{2.19}$$

By the definition of Δ^-, each $(i_k, j_k) \in \Delta^-$ satisfies

$$d_L(A_{i_k}, B_{i_k}) = \frac{|E_L[A_{i_k}, B_{i_k}]|}{|A_{i_k}||B_{i_k}|} < d(1 - \delta_0).$$

Thus,

$$|E_L[A_{i_k}, B_{i_k}]| < d(1 - \delta_0)|A_{i_k}||B_{i_k}|.$$

Returning to (2.19), we see

$$\begin{aligned} d_L(\vec{\mathcal{A}}, \vec{\mathcal{B}}) &< d(1 - \delta_0)\frac{\sum_{k=1}^{p} |A_{i_k}||B_{j_k}|}{K[\vec{\mathcal{A}}, \vec{\mathcal{B}}]} \\ &\overset{(2.17)}{=} d(1 - \delta_0)\frac{|K[\vec{\mathcal{A}}, \vec{\mathcal{B}}]|}{|K[\vec{\mathcal{A}}, \vec{\mathcal{B}}]|} = d(1 - \delta_0) \\ &\overset{(2.3)}{\leq} d(1 - \delta_*), \end{aligned}$$

as promised. $\square$

Chapter 3

Triads, Links, and Regularity

The goal of this chapter is to precisely describe several important graph and 3-graph concepts which appear throughout this dissertation. In particular, we define *triads*, $\mathcal{H}$*-triads*, and *links* which provide the foundation of all results herein. We also describe ways in which these objects can be considered regular and discuss witnesses of irregularity when they are not. These concepts form an analogue of the ε-regular pair (V_i, V_j) from Szemerédi's regularity lemma (Theorem 1.0.2). We now proceed with introducing these concepts.

3.1 Triads and $\mathcal{H}$-triads

We begin by defining the concept of a triad.

Definition 3.1.1 (Triad)**.** Let $V = V_1 \dot\cup V_2 \dot\cup V_3$ be a set of vertices with the given 3-partition and let $P = P^{12} \dot\cup P^{13} \dot\cup P^{23}$ be a 3-partite graph with the 3-partition $V_1 \dot\cup V_2 \dot\cup V_3$, where for each $1 \leq i < j \leq 3$, we have $P^{ij} = P[V_i, V_j]$. We call (P, V) a *triad*. Moreover, for $d_{12}, d_{13}, d_{23}, \varepsilon \in (0, 1]$, set $\vec{d} = (d_{12}, d_{13}, d_{23})$. We say that (P, V) is $(\vec{d}, \varepsilon)$*-regular* if for each $1 \leq i < j \leq 3$, we have that P^{ij} is (d_{ij}, ε)-regular.

Triads (P, V), as above, will be the building blocks of the hypergraphs we consider in this section. To explain, we consider the following notation. Define

$$\mathcal{K}_3(P) = \left\{ \{v_1, v_2, v_3\} \in \binom{V}{3} : \{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\} \in P, \right\}$$

to be the *triangles* of the graph P. Now let $\mathcal{H} \subseteq \binom{V}{3}$ be a 3-uniform hypergraph also defined on V. We say that P *underlies* $\mathcal{H}$ if $\mathcal{H} \subseteq \mathcal{K}_3(P)$. The following definition extends the triad (P, V) to include $\mathcal{H}$, which we therefore call an $\mathcal{H}$*-triad*.

Definition 3.1.2 ($\mathcal{H}$-triad)**.** We call a triple $(\mathcal{H}, P, V)$ an *$\mathcal{H}$-triad* if

(i) (P, V) is a triad, where $\mathcal{K}_3(P) \neq \emptyset$;

(ii) $\mathcal{H} \subseteq \mathcal{K}_3(P)$ has underlying graph P.

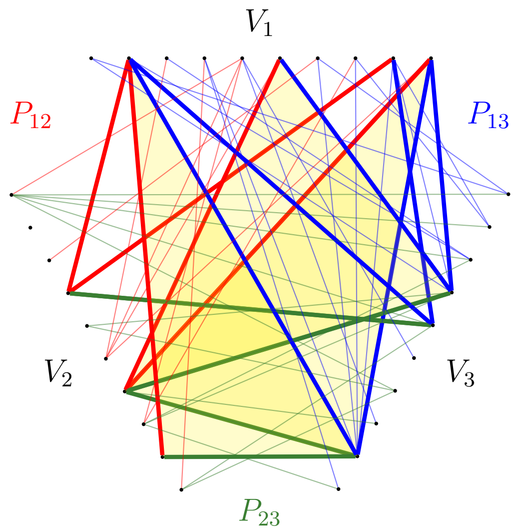

Figure 6.: An $\mathcal{H}$-triad. Members of $\mathcal{H}$ must be triangles of the triad (P, V).

Remark. Let $(\mathcal{H}, P, V)$ be an $\mathcal{H}$-triad, where $V = V_1 \dot\cup V_2 \dot\cup V_3$ and $P = P^{12} \dot\cup P^{13} \dot\cup P^{23}$ are as in Definition 3.1.1. Since P is a 3-partitite graph, with 3-partition $V_1 \dot\cup V_2 \dot\cup V_3$, and since $\mathcal{H} \subseteq \mathcal{K}_3(P)$, the hypergraph $\mathcal{H}$ is also 3-partitite, with the same 3-partition $V_1 \dot\cup V_2 \dot\cup V_3$.

3.2 Density and Regularity of $\mathcal{H}$-triads

Definition 3.2.1 (Density)**.** Let $(\mathcal{H}, P, V)$ be an $\mathcal{H}$-triad. We define the *density of $\mathcal{H}$ with respect to P* as

$$d_{\mathcal{H}}(P) = \frac{|\mathcal{H}|}{|\mathcal{K}_3(P)|}.$$

More generally, let $Q \subseteq P$ be a subgraph of P for which $\mathcal{K}_3(Q) \neq \emptyset$. We define the *density of $\mathcal{H}$ with respect to Q* as

$$d_{\mathcal{H}}(Q) = \frac{|\mathcal{H}(Q)|}{|\mathcal{K}_3(Q)|},$$

where $\mathcal{H}(Q) = \mathcal{H} \cap \mathcal{K}_3(Q)$ is the sub-hypergraph of $\mathcal{H}$ induced by $\mathcal{K}_3(Q)$.

We will also require an iterated version of the concept of density.

Definition 3.2.2 (r-density)**.** Let $(\mathcal{H}, P, V)$ be an $\mathcal{H}$-triad and let $r \in \mathbb{N}$ be a positive integer. Let $\vec{\mathcal{Q}} = (Q_1, \ldots, Q_r)$ be a sequence of subgraphs $Q_1, \ldots, Q_r \subseteq P$. We write

$$\mathcal{K}_3(\vec{\mathcal{Q}}) = \bigcup_{i=1}^{r} \mathcal{K}_3(Q_i)$$

for the familiy of triangles of $\vec{\mathcal{Q}}$. When $\mathcal{K}_3(\vec{\mathcal{Q}}) \neq \emptyset$ we define the *density of $\mathcal{H}$ with respect to $\vec{\mathcal{Q}}$* by

$$d_{\mathcal{H}}(\vec{\mathcal{Q}}) = \frac{|\mathcal{H}(\vec{\mathcal{Q}})|}{|\mathcal{K}_3(\vec{\mathcal{Q}})|},$$

where $\mathcal{H}(\vec{\mathcal{Q}}) = \mathcal{H} \cap \mathcal{K}_3(\vec{\mathcal{Q}})$ is the sub-hypergraph of $\mathcal{H}$ induced by $\mathcal{K}_3(\vec{\mathcal{Q}})$.

Remark. In Definition 3.2.2, the order of the sequence $\vec{\mathcal{Q}}$ is immaterial, and overlap or even repetition of coordinates is allowed.

We now define a concept of regularity for $\mathcal{H}$-triads $(\mathcal{H}, P, V)$.

Definition 3.2.3 ((δ, r)-regular)**.** Let $(\mathcal{H}, P, V)$ be an $\mathcal{H}$-triad and let $\delta > 0$ and $r \in \mathbb{N}$ be given. We say that $\mathcal{H}$ is (δ, r)*-regular with respect to* P if for any sequence $\vec{\mathcal{Q}} = (Q_1, \ldots, Q_r)$ of subgraphs of P for which

$$|\mathcal{K}_3(\vec{\mathcal{Q}})| > \delta |\mathcal{K}_3(P)|, \tag{3.1}$$

we have that $\vec{\mathcal{Q}}$ satisfies

$$|d_{\mathcal{H}}(\vec{\mathcal{Q}}) - d_{\mathcal{H}}(P)| < \delta. \tag{3.2}$$

When $\mathcal{H}$ is not (δ, r)-regular with respect to P, then we say $\mathcal{H}$ is (δ, r)*-irregular with respect to* P and any sequence $\vec{\mathcal{Q}} = (Q_1, \ldots, Q_r)$ satisfying 3.1 but failing 3.2 is said to be an r*-witness* to the (δ, r)-irregularity of $\mathcal{H}$ with respect to P.

3.3 Links and Link-regularity

We now describe when an $\mathcal{H}$-triad $(\mathcal{H}, P, V)$ enjoys a different regularity condition, which will be framed in terms of the following so-called link graphs L_v of $\mathcal{H}$, where $v \in V$.

Definition 3.3.1 (Link Graph)**.** Let $(\mathcal{H}, P, V)$ be an $\mathcal{H}$-triad. For a vertex $x \in V$, we define

$$L_x = \Big\{ \{y, z\} \in P : \{x, y, z\} \in \mathcal{H} \Big\}$$

to be the *link graph* of x.

Remark. In the definition above, suppose that $P = P^{12} \dot\cup P^{13} \dot\cup P^{23}$ is a 3 partite graph with vertex partition $V = V_1 \dot\cup V_2 \dot\cup V_3$ as in Definition 3.1.1. If $x \in V_1$, then L_x is the bipartite subgraph of P^{23} whose vertex partition is given by $N_{P^{12}}(x) \dot\cup N_{P^{13}}(x)$ where $N_{P^{12}}(x) = N_P(x) \cap V_2$ and $N_{P^{13}}(x) = N_P(x) \cap V_3$ are the P_{12} and P_{13} neighborhoods of x in V_2 and V_3 respectively.

For an $\mathcal{H}$-triad $(\mathcal{H}, P, V)$, we would like the link graphs L_x of most vertices $x \in V$ to be $(d_x, \delta, 1)$-regular in the sense of Definition 2.0.1 where d_x is an appropriate average given in terms of $\mathcal{H}$ and P. To describe this average, we impose the following further hypotheses on $(\mathcal{H}, P, V)$.

Definition 3.3.2 ($(\alpha, \vec{d}, \varepsilon)$-triad)**.** Let $\alpha, d_{12}, d_{13}, d_{23}, \varepsilon > 0$ be given, and write $\vec{d} = (d_{12}, d_{13}, d_{23})$. We say that an $\mathcal{H}$-triad $(\mathcal{H}, P, V)$ on n vertices is an $(\alpha, \vec{d}, \varepsilon)$-triad if it satisfies the following properties:

(i) $V = V_1 \dot\cup V_2 \dot\cup V_3$ is a partition satisfying

$$|V_1|, |V_2|, |V_3| = \Theta(n)$$

in which case all cardinalities $|V_1|$, $|V_2|$, $|V_3|$ are comparable;

(ii) (P, V) is a $(\vec{d}, \varepsilon)$-*triad*, i.e. $P = P^{12} \dot\cup P^{13} \dot\cup P^{23}$ has each bipartite graph P^{ij}, $1 \le i < j \le 3$, being (d_{ij}, ε)-regular;

(iii) $\mathcal{H} \subseteq \mathcal{K}_3(P)$ satisfies

$$d_{\mathcal{H}}(P) = \frac{|\mathcal{H}|}{|\mathcal{K}_3(P)|} = \frac{|\mathcal{H}(P)|}{|\mathcal{K}_3(P)|} = \alpha.$$

Now, for an $(\alpha, \vec{d}, \varepsilon)$-triad $(\mathcal{H}, P, V)$, we would like the link graphs L_x of most $x \in V_1$ to be $(d_x, \delta, 1)$-regular, and we would like, moreover, for d_x to be typically given by

$$d_x = \alpha d_{23}. \tag{3.3}$$

To justify this expression, we invoke two well-known facts from the literature on ε-regularity.

Fact 3.3.3 (Neighborhood lemma, [19]). *Let (P, V) be a $(\vec{d}, \varepsilon)$-triad (as in Definition 3.3.2 (ii)). Then all but $2\varepsilon|V_1|$ vertices $v \in V_1$ satisfy*

$$(d_{1j} - \varepsilon)|V_j| \leq |N_{P^{1j}}(v)| \leq (d_{1j} + \varepsilon)|V_j|$$

for both $j = 2$ and $j = 3$.

Fact 3.3.4 (Triangle counting lemma, [19]). *For all $d_0, \gamma > 0$, there exists $\varepsilon > 0$ so that the following holds. Let (P, V) be a $(\vec{d}, \varepsilon)$-triad (as in Definition 3.3.2), where $\vec{d} = (d_{12}, d_{13}, d_{23})$ satisfies $d_{12}, d_{13}, d_{23} > d_0$, and $V = V_1 \dot\cup V_2 \dot\cup V_3$ has each $|V_1|$, $|V_2|$, $|V_3|$ sufficiently large. Then*

$$|\mathcal{K}_3(P)| = (1 \pm \gamma) d_{12} d_{13} d_{23} |V_1||V_2||V_3|.$$

We now informally justify (3.3). On one hand, for an appropriately given $(\alpha, \vec{d}, \varepsilon)$-triad $(\mathcal{H}, P, V)$, we have

$$\sum_{x \in V_1} |L_x| = |\mathcal{H}| = \alpha |\mathcal{K}_3(P)| \overset{\text{Fact 3.3.4}}{\approx} \alpha d_{12} d_{13} d_{23} |V_1||V_2||V_3|.$$

On the other hand, for each $x \in V$ we have $|L_x| = d_x |N_{P^{12}}(x)||N_{P^{13}}(x)|$, where $d_x = d_{L_x}(N_{P^{12}}(x), N_{P^{13}}(x))$ is the density of L_x. Typically, these neighborhoods are governed by Fact 3.3.3, and so

$$d_{12} d_{13} |V_2||V_3| \sum_{x \in V_1} d_x \approx \sum_{x \in V_1} |L_x| \approx \alpha d_{12} d_{13} d_{23} |V_1||V_2||V_3|,$$

so that the average density satisfies

$$\mathbb{E}_{x \in V_1}[d_x] = \frac{1}{|V_1|} \sum_{x \in V_1} d_x \approx \alpha d_{23},$$

as sought in (3.3).

To make the concepts above precise, we introduce the following definition.

Definition 3.3.5 (δ-link regular)**.** Let $(\mathcal{H}, P, V)$ be an $(\alpha, \vec{d}, \varepsilon)$-triad. For $\delta > 0$, we say that $(\mathcal{H}, P, V)$ is δ-*link regular* if for all but $\delta|V_1|$ vertices $x \in V_1$, the link graph L_x is $(\alpha d_{23}, \delta, 1)$-regular.

Chapter 4
The Link Algorithm

The goal of this chapter is to demonstrate an algorithm $\mathbb{A}_{\text{link}}$ which efficiently confirms whether an $(\alpha, \vec{d}, \varepsilon)$-triad $(\mathcal{H}, P, V)$ is δ-link regular (see Definition 3.3.5), or efficiently constructs an r-witness of the $(\delta_{\#}, r)$-irregularity of $(\mathcal{H}, P, V)$ (see Definition 3.2.3), where here, $0 < \delta_{\#} \ll \delta$ is suitably smaller than δ. In practice, (P, V) will be class of a partition (see Definition 5.0.3), where $\mathcal{H}$ will be a sub-hypergraph of a larger hypergraph $\mathcal{G}$, induced by the triangles $\mathcal{K}_3(P)$ of P. The job of $\mathbb{A}_{\text{link}}$ will be to determine if these classes are regular or provide a witness so that the partition may be refined in the case that too many classes are irregular. We now proceed to the promised algorithm $\mathbb{A}_{\text{link}}$.

4.1 The Algorithm $\mathbb{A}_{\text{link}}$

Lemma 4.1.1 (Algorithm $\mathbb{A}_{\text{link}}$)**.** *There exists an algorithm $\mathbb{A}_{\text{link}}$ so that, for all $\alpha_0, \delta > 0$, $\mathbb{A}_{\text{link}}$ determines $\delta_{\#} = \delta_{\#}(\alpha_0, \delta)$ so that, for all $d_0 > 0$, $\mathbb{A}_{\text{link}}$ determines $\varepsilon = \varepsilon(\alpha_0, \delta, \delta_{\#}, d_0) > 0$ and an integer $r = r(\alpha_0, \delta, \delta_{\#}, d_0) \in \mathbb{N}$ so that the following holds:*

Let $(\mathcal{H}, P, V)$ be a given $(\alpha, \vec{d}, \varepsilon)$-triad on n vertices, where α and $\vec{d} = (d_{12}, d_{13}, d_{23})$ satisfy $\alpha \geq \alpha_0$ and $d_{12}, d_{13}, d_{23} \geq d_0$. Then, in time $O(n^3)$, $\mathbb{A}_{\text{link}}$ either confirms that $(\mathcal{H}, P, V)$ is δ-link regular, or confirms that $(\mathcal{H}, P, V)$ is $(\delta_{\#}, r)$-irregular with respect to P, and constructs a corresponding r-witness $\vec{\mathcal{Q}}$ to this effect.

To describe the algorithm $\mathbb{A}_{\text{link}}$, we begin by discussing its input.

Input: Let $\alpha_0, \delta > 0$ be given. To define the promised constant $\delta_{\#} > 0$, we consider some auxiliary constants. First, set

$$\delta_1 = \frac{\delta}{3}. \tag{4.1}$$

Second, with $\delta_1 > 0$ given above, Lemma 2.0.2 determines the constant

$$\delta_* = \delta_*(\delta_1) > 0 \tag{4.2}$$

to be appropriate for an application of $\mathbb{A}_{\text{sparse}}$. We set

$$\delta_\# = \frac{1}{16}\delta_1\delta_*^6. \tag{4.3}$$

Let δ_0 be given. To determine the promised constants $\varepsilon = \varepsilon(\alpha_0, \delta, \delta_\#, d_0) > 0$ and $r = r(\alpha_0, \delta, \delta_\#, d_0) \in \mathbb{N}$, we again consider several auxiliary constants. First, set

$$D = \frac{1}{2}\alpha_0 d_0. \tag{4.4}$$

Now, Lemma 2.0.2 determines constants

$$c = c(\delta_1, \delta_*, D) \;\text{ and }\; r = r(\delta_1, \delta_*, D) \in \mathbb{N} \tag{4.5}$$

to be appropriate for an application of $\mathbb{A}_{\text{sparse}}$. We define the promised constant $r = r(\alpha_0, \delta, \delta_\#, d_0)$ to be

$$r = r(\delta_1, \delta_*, D). \tag{4.6}$$

Second, for defining the promised constant $\varepsilon = \varepsilon(\alpha_0, \delta, \delta_\#, d_0)$, let

$$\varepsilon_1 = \varepsilon_{\text{Fact 3.3.4}}(d_0, \gamma = 1) \tag{4.7}$$

be the constant guaranteed by the Triangle Counting Lemma (Fact 3.3.4). Set

$$\varepsilon = \frac{1}{2}\delta_*^3 \cdot \min\{\varepsilon_1, c, d_0\}. \tag{4.8}$$

This concludes our definitions of the promised constants.

Let $(\mathcal{H}, P, V)$ be a given $(\alpha, \vec{d}, \varepsilon)$-triad on n vertices, where α and $\vec{d} = (d_{12}, d_{13}, d_{23})$ satisfy $\alpha \geq \alpha_0$ and $d_{12}, d_{13}, d_{23} \geq d_0$, for $\alpha_0 > 0$ and $d_0 > 0$ given above. We describe the algorithm $\mathbb{A}_{\text{link}}$, which either confirms that $(\mathcal{H}, P, V)$ is δ-link regular, with $\delta > 0$ given above, or confirms that $\mathcal{H}$ is $(\delta_{\#}, r)$-irregular with respect to P, where $\delta_{\#}$ and r are given in (4.3) and (4.6) respectively. Moreover, in this case, $\mathbb{A}_{\text{link}}$ will construct an r-witness to this effect.

Procedure: The algorithm $\mathbb{A}_{\text{link}}$ proceeds along the following steps.

Step 1. (Identify good vertices) Fix $v \in V_1$, and $j \in \{2, 3\}$. Compute $|N_{P^{1j}}(v)|$ in time $O(|V_j|) = O(n)$. By Fact 3.3.3, all but $4\varepsilon|V_1|$ vertices $v \in V_1$ satisfy that for both $j = 2$ and $j = 3$,

$$(d_{1j} - \varepsilon)|V_j| \leq |N_{P^{1j}}(v)| \leq (d_{1j} + \varepsilon)|V_j|.$$

We call such vertices *good*, and denote the set of good vertices $v \in V_1$ by V_1^{good}. Repeating over all $v \in V_1$, we identify V_1^{good} in time $O(|V_1||V_2| + |V_1||V_3|) = O(n^2)$.

Step 2. (Compute link densities) Fix $x \in V_1^{\text{good}}$. Construct L_x in time $O(|P^{23}|) = O(|V_1||V_2|) = O(n^2)$. Equivalently, we have computed $d_x = d_{L_x}(N_{P^{12}}(x), N_{P^{13}}(x))$. We say a vertex $v \in V_1^{\text{good}}$ is *nice* if

$$\alpha d_{23}(1 - \delta_*) \leq d_x \leq \alpha d_{23}(1 + \delta_*),$$

and we denote the set of nice vertices $x \in V_1$ by $V_1^{\text{nice}} \subseteq V_1^{\text{good}} \subseteq V_1$. Repeating over all $x \in V_1^{\text{good}}$, we identify V_1^{nice} in time $O(|V_1|n^2) = O(n^3)$.

Now, using V_1^{nice} and V_1^{good}, the principal part of our final step will invoke the algorithm $\mathbb{A}_{\text{sparse}}$, and make other related constructions.

Step 3. ($\mathbb{A}_{\text{link}}$ and other constructions) In all that follows, set $X = V_1^{\text{good}} \setminus V_1^{\text{nice}}$. Prior to applying $\mathbb{A}_{\text{sparse}}$, we consider the following sets

$$X^+ = \{x \in X : d_x > \alpha d_{23}(1 + \delta_*)\}$$
$$X^- = \{x \in X : d_x < \alpha d_{23}(1 + \delta_*)\}$$

Note that $X = X^+ \,\dot\cup\, X^-$ is a partition. If either X^+ or X^- is large, we would be able to avoid invoking $\mathbb{A}_{\text{sparse}}$, and it would suffice to consider the following easy constructions. For X^+, we construct the following 1-tuple $\vec{\mathcal{Q}}_+ = (Q_+)$ of subgraphs $Q_+ \subseteq P = P_{12} \dot{\cup} P_{13} \dot{\cup} P_{23}$:

$$\begin{aligned} Q_+^{12} &= P^{12}[X^+, V_2], & Q_+^{13} &= P^{13}[X^+, V_3], \\ Q_+^{23} &= P^{23}, \qquad \text{and} & Q_+ &= Q_+^{12} \dot{\cup} Q_+^{13} \dot{\cup} Q_+^{23}. \end{aligned}$$

Clearly, Q_+ is constructed in time $O(|P^{12}| + |P^{13}| + |P^{23}|) = O(n^2)$. Similarly, in time $O(n^2)$ we construct the 1-tuple $\vec{\mathcal{Q}}_- = (Q_-)$ of subgraphs $Q_- \subseteq P = P_{12} \dot{\cup} P_{13} \dot{\cup} P_{23}$:

$$\begin{aligned} Q_-^{12} &= P^{12}[X^-, V_2], & Q_-^{13} &= P^{13}[X^-, V_3], \\ Q_-^{23} &= P^{23}, \qquad \text{and} & Q_- &= Q_-^{12} \dot{\cup} Q_-^{13} \dot{\cup} Q_-^{23}. \end{aligned}$$

Now set $Y = V_1^{\text{nice}}$. For each $y \in Y$, we run the algorithm $\mathbb{A}_{\text{sparse}}$ (see Lemma 2.0.2) on the link graph L_y. (For a proof that $\mathbb{A}_{\text{sparse}}$ applies to L_y, see the next subsection.) With $y \in Y = V_1^{\text{nice}}$ fixed, $\mathbb{A}_{\text{sparse}}$ yields in time $O(n^2)$ one of the following two outputs:

(I) $\mathbb{A}_{\text{sparse}}$ confirms that L_y is $(d_y, \delta_1, 1)$-regular;

(II) $\mathbb{A}_{\text{sparse}}$ detects that L_y is (d_y, δ_*, r)-irregular and constructs pair-disjoint, c-bounded r-witness $(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)$ to that effect, where $\vec{\mathcal{A}}_y = (A_1^y, \dots, A_r^y)$ and $\vec{\mathcal{B}}_y = (B_1^y, \dots, B_r^y)$ with $A_1^y, \dots, A_r^y \subseteq N_{P^{12}}(y)$ and $B_1^y, \dots, B_r^y \subseteq N_{P^{13}}(y)$. In particular, either

   (a) $d_{L_y}(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y) < d_y(1 - \delta_*)$, or

   (b) $d_{L_y}(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y) > d_y(1 + \delta_*)$.

Repeating over all $y \in Y = V_1^{\text{nice}}$, we distinguish, in time $O(|Y|n^2) = O(n^3)$, between (I) and (II) for every $y \in Y = V_1^{\text{nice}}$. Continuing, define the following sets:

$$\begin{aligned} Y_{\text{reg}} &= \{y \in Y = V_1^{\text{nice}} : \text{(I) holds}\}; \\ Y_{\text{irreg}} &= \{y \in Y = V_1^{\text{nice}} : \text{(II) holds}\}; \\ Y_{\text{irreg}}^- &= \{y \in Y = V_1^{\text{nice}} : \text{(II) and (a) hold}\}; \\ Y_{\text{irreg}}^+ &= \{y \in Y = V_1^{\text{nice}} : \text{(II) and (b) hold}\}. \end{aligned}$$

Clearly, $Y = Y_{\text{reg}} \dot\cup Y_{\text{irreg}}$ and $Y_{\text{irreg}} = Y^-_{\text{irreg}} \dot\cup Y^+_{\text{irreg}}$ are partitions. For Y^-_{irreg}, we construct the following r-tuple $\vec{\mathcal{Q}}_- = (Q^-_1, \ldots, Q^-_r)$ of subgraphs $Q^-_i \subseteq P$:

$$Q^{12}_{i,-} = P^{12}\left[Y^-_{\text{irreg}}, \bigcup_{y \in Y^-_{\text{irreg}}} A^y_i\right], \qquad Q^{13}_{i,-} = P^{13}\left[Y^-_{\text{irreg}}, \bigcup_{y \in Y^-_{\text{irreg}}} B^y_i\right],$$
$$Q^{23}_{i,-} = P^{23}, \qquad \text{and} \quad Q^-_i = Q^{12}_{i,-} \dot\cup Q^{13}_{i,-} \dot\cup Q^{23}_{i,-}.$$

Since $r = o(1)$ is contstant, we have that $\vec{\mathcal{Q}}_-$ is constructed in time $O\left(|P^{12}| + |P^{13}| + |P^{23}|\right) = O(n^2)$. Similarly, for Y^+_{irreg}, we construct in time $O(n^2)$ the following r-tuple $\vec{\mathcal{Q}}_+ = (Q^+_1, \ldots, Q^+_r)$ of subgraphs $Q^+_i \subseteq P$:

$$Q^{12}_{i,+} = P^{12}\left[Y^+_{\text{irreg}}, \bigcup_{y \in Y^+_{\text{irreg}}} A^y_i\right], \qquad Q^{13}_{i,+} = P^{13}\left[Y^+_{\text{irreg}}, \bigcup_{y \in Y^+_{\text{irreg}}} B^y_i\right],$$
$$Q^{23}_{i,+} = P^{23}, \qquad \text{and} \quad Q^+_i = Q^{12}_{i,+} \dot\cup Q^{13}_{i,+} \dot\cup Q^{23}_{i,+}.$$

This concludes Step 3.

Output:

1) For $X = V_1^{\text{good}} \setminus V_1^{\text{nice}} = X^+ \dot\cup X^-$,
   `if` $|X^+| \geq \delta_1|V_1|$ `or` $|X^-| \geq \delta_1|V_1|$,
   `return` $\mathcal{H}$ is $(\delta_\#, 1)$-irregular and return either qualifying 1-tuple $\vec{\mathcal{Q}}_+$ or $\vec{\mathcal{Q}}_-$ respectively.

2) `else` $|X| = |V_1^{\text{good}} \setminus V_1^{\text{nice}}| < 2\delta_1|V_1|$,
   and so

$$\begin{aligned}|Y| = |V_1^{\text{nice}}| &= |V_1| - |V_1 \setminus V_1^{\text{good}}| - |V_1^{\text{good}}| \setminus V_1^{\text{nice}}| \\ &\geq (1 - 4\varepsilon - 2\delta_1)|V_1|,\end{aligned}$$

   where $Y = Y_{\text{reg}} \dot\cup Y_{\text{irreg}} = Y_{\text{reg}} \dot\cup Y^+_{\text{irreg}} \dot\cup Y^-_{\text{irreg}}$.
   `if` $|Y^+_{\text{irreg}}| \geq \delta_1|V_1|$ `or` $|Y^-_{\text{irreg}}| \geq \delta_1|V_1|$
   `return` $\mathcal{H}$ is $(\delta_\#, r)$-irregular and return the r-tuple $\vec{\mathcal{Q}}_+$ or $\vec{\mathcal{Q}}_-$ respectively.

3) `else` $|Y_{\text{reg}}| = |Y| - |Y_{\text{irreg}^+} \dot\cup Y^-_{\text{irreg}}| \geq (1 - 4\varepsilon - 2\delta_1)|V_1|$
   and so
   `return` $(\mathcal{H}, P, V)$ is δ-link regular.

This concludes the descriptions of the algorithm $\mathbb{A}_{\text{link}}$. It now remains to prove its correctness.

4.2 Proof of Lemma 4.1.1

It is clear that the algorithm $\mathbb{A}_{\text{link}}$ runs in time $O(n^2)$, which is achieved in steps 2 and 3. We must show the correctness of its output. We examine cases 1)-3) in the output of $\mathbb{A}_{\text{link}}$ in the order of increasing technicality.

4.2.1 Output Case 3

The algorithm is correct by definition. Indeed,

$$|Y_{\text{reg}}| \geq (1 - 4\varepsilon - 2\delta_1)|V_1| \overset{(4.8)}{\geq} (1 - 3\delta_1)|V_1| \geq (1 - \delta)|V_1|.$$

The link graph L_y of every vertex $y \in Y_{\text{reg}} \subseteq Y = V_1^{\text{nice}}$ is $(d_y, \delta_1, 1)$-regular, where $\alpha d_{23}(1-\delta_*) \leq d_y \leq \alpha d_{23}(1 + \delta_*)$. From Definition 2.0.1, this means that for any sequence $\vec{\mathcal{A}}_y = (A_1, \ldots, A_r)$ of subsets from $N_{P^{12}}(y)$ and for any sequence $\vec{\mathcal{B}}_y = (B_1, \ldots, B_r)$ of subsets from $N_{P^{13}}(y)$, we have that $\left|K(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)\right| > \delta_1 |N_{P^{12}}(y)||N_{P^{13}}(y)|$ implies

$$\begin{aligned} &\alpha d_{23}(1 - \delta_*)(1 - \delta_1) \leq d_y(1 - \delta_1) \leq d_{L_y}(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y) \leq d_y(1 + \delta_1) \leq \alpha d_{23}(1 + \delta_*)(1 + \delta_1) \\ \Longrightarrow\ &\alpha d_{23}(1 - \delta_1)^2 \leq d_{L_y}(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y) \leq \alpha d_{23}(1 + \delta_1)^2 \end{aligned}$$

where we use that $\delta_* \leq \delta_1$ from (4.2). Moreover, since $(1 + \delta_1)^2 \leq 1 + 3\delta_1 \overset{(4.1)}{\leq} 1 + \delta$ and $(1 - \delta_1)^2 \geq 1 - 2\delta_1 \overset{(4.1)}{\geq} 1 - \delta$, the set Y_{reg} verifies that Definition 3.3.5 is satisfied.

4.2.2 Output Case 1

Whether $\mathbb{A}_{\text{link}}$ constructs the 1-tuple $\vec{\mathcal{Q}}_+$ or $\vec{\mathcal{Q}}_-$ in its output, the arguments for its correctness are entirely symmetric. We verify only the outcome $\vec{\mathcal{Q}}_-$. To that end, recall that $\vec{\mathcal{Q}}_- = (Q_-)$ is the 1-tuple constructed with the set

$$X^- = \left\{ x \in X = V_1^{\text{good}} \setminus V_1^{\text{nice}} : d_x < \alpha d_{23}(1 - \delta_*) \right\}$$

where $|X^-| \geq \delta_1|V_1|$, given by

$$\begin{aligned} Q_-^{12} &= P^{12}[X^-, V_2], & Q_-^{13} &= P^{13}[X^-, V_3], \\ Q_-^{23} &= P^{23}, \qquad \text{and} & Q_- &= Q_-^{12} \dot\cup Q_-^{13} \dot\cup Q_-^{23}. \end{aligned}$$

To see that $\vec{\mathcal{Q}}_-$ is a 1-witness of the $(\delta_\#, 1)$-irregularity of $\mathcal{H}$ with respect to P, we note the following easy identities:

$$\begin{aligned} |\mathcal{K}_3(Q_-)| &= \sum_{x \in X^-} \left| P^{23}[N_{P^{12}}(x), N_{P^{13}}(x)] \right|, \\ |\mathcal{H} \cap \mathcal{K}_3(Q_-)| &= \sum_{x \in X^-} |L_x|. \end{aligned}$$

To bound the quantities, fix $x \in X^- \subset V_1^{\text{good}}$. Then, with $\varepsilon < \frac{1}{2}\min\{d_{12}, d_{13}\}$ we have

$$|N_{P^{12}}(x)| \geq (d_{12} - \varepsilon)|V_2| \geq \varepsilon|V_2| \text{ and } |N_{P^{13}}(x)| \geq (d_{13} - \varepsilon)|V_3| \geq \varepsilon|V_3|. \tag{4.9}$$

By the (d_{23}, ε)-regularity of P^{23}, we infer

$$\begin{aligned} |\mathcal{K}_3(Q_-)| &= \sum_{x \in X^-} \left| P^{23}[N_{P^{12}}(x), N_{P^{13}}(x)] \right| \\ &\geq (d_{23} - \varepsilon) \sum_{x \in X^-} |N_{P^{12}}(x)||N_{P^{13}}(x)| \\ &\geq (d_{22} - \varepsilon)(d_{12} - \varepsilon)|V_2|(d_{13} - \varepsilon)|V_3||X^-| \\ &\geq \frac{1}{8}\delta_1 d_{12} d_{13} d_{23} |V_1||V_2||V_3|, \end{aligned} \tag{4.10}$$

where we used $\varepsilon < \frac{1}{2}\min\{d_{12}, d_{13}, d_{23}\}$, $|X^-| \geq \delta_1|V_1|$, and that every vertex $x \in X^- \subset V_1^{\text{good}} \setminus V_1^{\text{nice}}$ is a good vertex. By our choice of ε in 4.7 and 4.8, the triangle counting lemma applies to P to say

$$|\mathcal{K}_3(P)| \leq 2 d_{12} d_{13} d_{23} |V_1||V_2||V_3|, \tag{4.11}$$

and so comparing (4.10) and (4.11) renders

$$|\mathcal{K}_3(Q_-)| \geq \frac{1}{16}\delta_1|\mathcal{K}_3(P)| \overset{(4.3)}{>} \delta_\#|\mathcal{K}_3(P)|,$$

as desired. By definition of X^-,

$$\begin{aligned}
|\mathcal{H} \cap \mathcal{K}_3(Q_-)| &= \sum_{x \in X^-} |L_x| \\
&< (1-\delta_*)\alpha \sum_{x \in X^-} d_{23}|N_{P^{12}}(x)||N_{P^{13}}(x)| \\
&= (1-\delta_*)\alpha \frac{\delta_{23}}{\delta_{23}-\varepsilon} \sum_{x \in X^-} (d_{23}-\varepsilon)|N_{P^{12}}(x)||N_{P^{13}}(x)| \\
&\overset{(4.10)}{\leq} (1-\delta_*)\alpha \frac{1}{1-\varepsilon d_{23}^{-1}}|\mathcal{K}_3(Q_-)| \\
&\leq (1-\delta_*)\alpha(1+\varepsilon d_{23}^{-1})|\mathcal{K}_3(Q_-)| \\
&\leq (1-\delta_*)(1+\delta_*)\alpha|\mathcal{K}_3(Q_-)|,
\end{aligned}$$

where the last inequalities hold with $\varepsilon < \delta_* d_{23}/2$. As such,

$$\begin{aligned}
d_{\mathcal{H}}(Q_-) = \frac{|\mathcal{H} \cap \mathcal{K}_3(Q_-)|}{|\mathcal{K}_3(Q_-)|} &\leq (1-\delta_*)(1+\delta_*)\alpha \\
&= (1-\delta_*^2)\alpha \leq \alpha - \delta_*^2 \overset{(4.3)}{\leq} \alpha - \delta_\#,
\end{aligned}$$

as desired.

4.2.3 Output Case 2

To prove the correctness of output Case 2, we should first verify that its key contributor, the algorithm $\mathbb{A}_{\text{sparse}}$ of Lemma 2.0.2, applies to the setting of the link graph L_y of a fixed vertex $y \in Y = V_1^{\text{nice}}$. For that, recall that we chose $\delta_1 > 0$ in (4.1), and determined $\delta_* = \delta_*(\delta_1)$ in (4.2) to be appropriate for an application of $\mathbb{A}_{\text{sparse}}$. We set $D = \alpha_0 d_0/2$ in (4.4), which is a lower bound on

$$d_{L_y}\left(N_{P^{12}}(y), N_{P^{13}}(y)\right) \geq \alpha d_{23}(1-\delta_*) \geq \frac{\alpha_0 d_0}{2} = D,$$

as required by an application of $\mathbb{A}_{\text{sparse}}$. Moreover, we chose $r = r(D)$ in (4.6) to also be appropriate for an application of $\mathbb{A}_{\text{sparse}}$. Finally, since $y \in Y = V_1^{\text{nice}} \subseteq V_1^{\text{good}}$ is a good vertex, the vertex bipartition $N_{P^{12}}(y) \,\dot\cup\, N_{P^{13}}(y)$ of L_y satisfies

$$|N_{P^{12}}(y)| = (d_{12} \pm \varepsilon)|V_2| = \Omega(n) \ \text{ and } \ |N_{P^{13}}(y)| = (d_{13} \pm \varepsilon)|V_3| = \Omega(n)$$

as required by $\mathbb{A}_{\text{sparse}}$. Thus, $\mathbb{A}_{\text{sparse}}$ may be applied to the link graph L_y whenever $y \in V_1^{\text{nice}}$.

Whether $\mathbb{A}_{\text{link}}$ constructs the r-tuple $\vec{\mathcal{Q}}_+$ or $\vec{\mathcal{Q}}_-$ in its output, the arguments for its correctness are entirely symmetric. Moreover, these arguments are essentially similar to output Case 1, just with added symbolic technicality. We verify only the outcome $\vec{\mathcal{Q}}_-$. To that end, recall that $\vec{\mathcal{Q}}_- = (Q_1^-, \ldots, Q_r^-)$ is the r-tuple of subgraphs $Q_i^- \subseteq P$ constructed with the set $Y = V_1^{\text{nice}}$ and with the r-witnesses $(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)$ over $y \in Y_{\text{irreg}}^-$, as follows: For each $1 \le i \le r$,

$$\begin{aligned} Q_{i,-}^{12} &= P^{12}\left[Y_{\text{irreg}}^-, \textstyle\bigcup_{y \in Y_{\text{irreg}}^-} A_i^y\right], & Q_{i,-}^{13} &= P^{13}\left[Y_{\text{irreg}}^-, \textstyle\bigcup_{y \in Y_{\text{irreg}}^-} B_i^y\right], \\ Q_{i,-}^{23} &= P^{23}, & \text{and} \quad Q_i^- &= Q_{i,-}^{12} \,\dot\cup\, Q_{i,-}^{13} \,\dot\cup\, Q_{i,-}^{23}. \end{aligned}$$

To see that $\vec{\mathcal{Q}}_- = (Q_1^-, \ldots, Q_r^-)$ is an r-witness of the $(\delta_\#, r)$-irregularity of $\mathcal{H}$ with respect to P, we first note the following identities. First,

$$\begin{aligned} |\mathcal{K}_3(\vec{\mathcal{Q}}_-)| &= \sum_{y \in Y_{\text{irreg}}^-} \left| E_{P^{23}}\left[\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y\right] \right| \\ &= \sum_{y \in Y_{\text{irreg}}^-} \sum_{i=1}^{r} \left| P^{23}\left[A_i^y, B_i^y\right] \right| \end{aligned} \tag{4.12}$$

follows from the pair-disjointness of each $(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)$, $y \in Y_{\text{irreg}}^-$. Second, we have

$$|\mathcal{H} \cap \mathcal{K}_3(\vec{\mathcal{Q}}_-)| = \sum_{y \in Y_{\text{irreg}}^-} \left| E_{L_y}\left[\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y\right] \right|. \tag{4.13}$$

To bound these quantities, fix $y \in Y_{\text{irreg}}^-$. Since $(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)$ is a c-bounded r-witness, we have $|A_i^y| \ge |N_{P^{12}}(y)|$ and $|B_i^y| \ge |N_{P^{13}}(y)|$ for all $1 \le i \le r$. Since $\varepsilon < c$ in (4.8), we thus have $|P^{23}[A_i^y, B_i^y]| = (d_{23} \pm \varepsilon)|A_i^y||B_i^y|$ for all $1 \le i \le r$, and therefore in (4.12),

$$\begin{aligned}
|\mathcal{K}_3(\vec{\mathcal{Q}}_-)| &= \sum_{y \in Y^-_{\text{irreg}}} \sum_{i=1}^{r} \left| P^{23}\left[A_i^y, B_i^y\right]\right| \\
&\ge (d_{23} - \varepsilon) \sum_{y \in Y^-_{\text{irreg}}} \sum_{i=1}^{r} |A_i^y||B_i^y| \\
&= (d_{23} - \varepsilon) \sum_{y \in Y^-_{\text{irreg}}} |K(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)|, \qquad (4.14)
\end{aligned}$$

where the last identity follows from the pair-disjointness of $(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)$ for $y \in Y^-_{\text{irreg}}$. Since $(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)$ is an r-witness of the (d_y, δ_*, r)-irregularity of L_y, we have that (see Definition 2.0.1) that for each $y \in Y^-_{\text{irreg}}$,

$$\begin{aligned}
|K(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)| &> \delta_* |N_{P^{12}}(y)||N_{P^{13}}(y)| \\
&\ge \delta_*(d_{12} - \varepsilon)|V_2|(d_{13} - \varepsilon)|V_3|,
\end{aligned}$$

where the last inequality follows from the fact that $y \in Y^-_{\text{irreg}} \subseteq V_1^{\text{nice}} \subseteq V_1^{\text{good}}$ is a good vertex. Thus, returning to (4.14), we have

$$\begin{aligned}
|\mathcal{K}_3(\vec{\mathcal{Q}}_-)| &= (d_{23} - \varepsilon) \sum_{y \in Y^-_{\text{irreg}}} |K(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)| \\
&\ge \delta_*(d_{12} - \varepsilon)(d_{13} - \varepsilon)(d_{23} - \varepsilon)|Y^-_{\text{irreg}}||V_2||V_3| \\
&\ge \delta_*\delta_1(d_{12} - \varepsilon)(d_{13} - \varepsilon)(d_{23} - \varepsilon)|V_1||V_2||V_3| \\
&\overset{(4.8)}{\ge} \frac{1}{8}\delta_*\delta_1 d_{12} d_{13} d_{23}|V_1||V_2||V_3|, \qquad (4.15)
\end{aligned}$$

where we used the assumption that $|Y^-_{\text{irreg}}| \ge \delta_1|V_1|$ from Output Case 2. Recalling (4.11), we therefore conclude

$$|\mathcal{K}_3(\vec{\mathcal{Q}}_-)| \ge \frac{1}{16}\delta_*\delta_1|\mathcal{K}_3(P)| \overset{(4.3)}{>} \delta_\#|\mathcal{K}_3(P)|,$$

as desired.

Returning to (4.13), the r-witness $(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y)$ for $y \in Y^-_{\text{irreg}}$, ensures (recall (4.10)) that $d_{L_y}(\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y) < d_y(1-\delta_*)$, or equivalently

$$\begin{aligned}\frac{|E_{L_y}[\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y]|}{|K[\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y]|} &< dy(1-\delta_*) \leq \alpha d_{23}(1+\delta_*)(1-\delta_*) \\ &= \alpha d_{23}(1-\delta_*^2),\end{aligned}$$

where we used that $y \in Y^-_{\text{irreg}} \subseteq V_1^{\text{nice}}$ is a nice vertex. Thus, from (4.13), we have

$$\begin{aligned}|\mathcal{H} \cap \mathcal{K}_3(\vec{\mathcal{Q}}_-)| &= \sum_{y \in Y^-_{\text{irreg}}} \left| E_{L_y}\left[\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y\right]\right| \\ &< \alpha d_{23}(1-\delta_*^2) \sum_{y \in Y^-_{\text{irreg}}} \left|K[\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y]\right| \\ &= \alpha \frac{d_{23}}{d_{23}-\varepsilon}(1-\delta_*^2) \sum_{y \in Y^-_{\text{irreg}}} (d_{23}-\varepsilon)\left|K[\vec{\mathcal{A}}_y, \vec{\mathcal{B}}_y\right| \\ &\overset{(4.15)}{\leq} \alpha \frac{1-\delta_*^2}{1-\varepsilon d_{23}^{-1}} |\mathcal{K}_3(\vec{\mathcal{Q}}_-)| \\ &\leq \alpha(1-\delta_*^2)(1+2\varepsilon d_{23}^{-1})|\mathcal{K}_3(\vec{\mathcal{Q}}_-)| \\ &\overset{(4.8)}{\leq} \alpha(1-\delta_*^3)(1+\delta_*^3)|\mathcal{K}_3(\vec{\mathcal{Q}}_-)|,\end{aligned}$$

where the last inequalities hold with $\varepsilon < \delta_*^3 d_{23}/2$. As such,

$$\begin{aligned}d_{\mathcal{H}}(\vec{\mathcal{Q}}_-) = \frac{|\mathcal{H} \cap \mathcal{K}_3(\vec{\mathcal{Q}}_-)|}{|\mathcal{K}_3(\vec{\mathcal{Q}}_-)|} &\leq \alpha(1-\delta_*^3)(1+\delta_*^3) \\ &= \alpha(1-\delta_*^6) \leq \alpha - \delta_*^6 \overset{(4.3)}{\leq} \alpha - \delta_{\#},\end{aligned}$$

as desired. This concludes our proof of correctness for the algorithm $\mathbb{A}_{\text{link}}$.

Chapter 5

The Algorithmic Link-Regularity Lemma

In this chapter, we prove Theorem 1.4.3 in a precise sense. However, this requires some investment in nomenclature and auxiliary concepts that we will introduce first.

For a 3-graph $\mathcal{G}$, the algorithmic link-regularity lemma will partition both the vertices $V = V(\mathcal{G})$ and the pairs $\binom{V}{2}$. The basic structure of these partitions is described in the following definition.

Definition 5.0.1 ((ℓ, t)-partition)**.** Let $\ell, t \in \mathbb{N}$ be positive integers. For a set of vertices V, an (ℓ, t)-*partition* of V is a pair of partitions $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ so that

(i) $\Pi^{(1)} : V = V_0 \dot\cup V_1 \dot\cup \ldots \dot\cup V_t$ is a partition of V,

(ii) $\Pi^{(2)}$ is an edge partition of $K[V_1, \ldots, V_t]$ with classes given by, for each $1 \le i < j \le t$, bipartite subgraphs
$$K[V_i, V_j] = P_0^{ij} \dot\cup P_1^{ij} \dot\cup \ldots \dot\cup P_{\ell_{ij}}^{ij}$$
where $0 \le \ell_{ij} \le \ell$ is an integer depending on i, j.

Remark. Pairs $\{u, v\} \in \binom{V}{2}$ not belonging to any class P_a^{ij}, where $1 \le i < j \le t$ and $0 \le a \le \ell_{ij}$, constitute their own class which we disregard in application.

Observe that an (ℓ, t)-partition $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ of V defines a family of triads

$$P_{abc}^{ijk} := P_a^{ij} \cup P_b^{jk} \cup P_c^{ik},$$

where $P_a^{ij}, P_b^{jk}, P_c^{ik} \in \Pi^{(2)}$ are classes of $\Pi^{(2)}$ for some $1 \le i < j < k \le t$. We define the following notation.

Definition 5.0.2 (Triad($\mathbf{\Pi}$))**.** Let $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ be an (ℓ, t)-partition of V. Define

$$\mathrm{Triad}(\mathbf{\Pi}) = \left\{ P = P^{ijk}_{abc} : \{i,j,k\} \in \binom{[t]}{3}, (a,b,c) \in [0, \ell_{ij}] \times [0, \ell_{jk}] \times [0, \ell_{ik}] \right\}$$

to be the family of all

$$\sum_{\{i,j,k\} \in \binom{[t]}{3}} (\ell_{ij}+1)(\ell_{jk}+1)(\ell_{ik}+1) \le \binom{t}{3}(1+\ell)^3 \le 2\ell^3 t^3 \tag{5.1}$$

many triads of $\mathbf{\Pi}$.

We establish some related notation. Since

$$\bigcup_{P \in \mathrm{Triad}(\mathbf{\Pi})} \mathcal{K}_3(P) \subset \binom{V}{3}$$

is a disjoint union, every $\{x,y,z\} \in \binom{V}{3}$ admits at most one triad $P = P^{ijk}_{abc} \in \mathrm{Triad}(\mathbf{\Pi})$ so that

$$\{x,y,z\} \in \mathcal{K}_3(P) = \mathcal{K}_3(P^{ij}_a \cup P^{jk}_b \cup P^{ik}_c).$$

For simplicity of notation, we denote this triad by $P = P_{xyz}$ and say that x, y, z *belongs* to the triad P_{xyz}.

In context, we will want (ℓ, t)-partitions $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ to be equitable and regular, in the following sense (see Figure 7).

Definition 5.0.3 ($(\ell, t, \gamma, \varepsilon)$-partition)**.** Let V have (ℓ, t)-partition $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ and let $\varepsilon, \gamma > 0$ be given. We say that $\mathbf{\Pi}$ is an $(\ell, t, \gamma, \varepsilon)$-*partition* if the following conditions are satisfied:

(i) $\Pi^{(1)} : V = V_0 \,\dot\cup\, V_1 \,\dot\cup \ldots \dot\cup\, V_t$ is t-equitable, i.e., $|V_0| < t$ and $|V_1| = \ldots = |V_t|$;

(ii) $\Pi^{(2)}$ is (ℓ, γ)-equitable, meaning that for all but $\gamma\binom{t}{2}$ pairs $1 \le i < j \le t$, we have $|P^{ij}_0| < \gamma |V_i||V_j|$ and for all $1 \le a \le \ell_{ij} \le \ell$, we have

$$\left(\frac{1}{\ell} - \varepsilon\right)|V_i||V_j| \le |P^{ij}_a| \le \left(\frac{1}{\ell} + \varepsilon\right)|V_i||V_j|;$$

(iii) All but $\gamma|V|^2$ pairs $\{x,y\} \in \binom{V}{2}$ belong to a class $P_a^{ij} \in \Pi^{(2)}$, where P_a^{ij} is ε-regular.

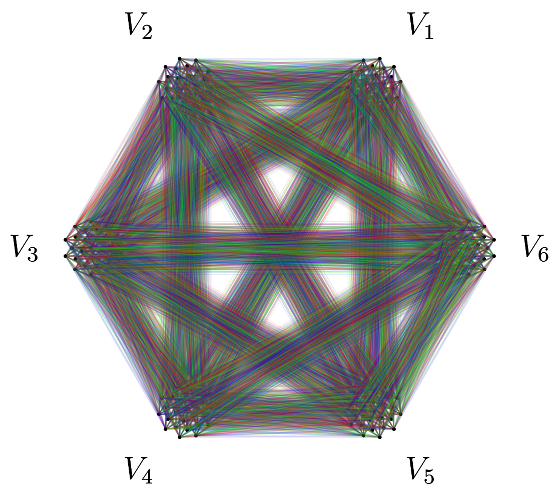

Figure 7.: An $(\ell, t, \gamma, \varepsilon)$-partition where $t = 6$ and $\ell = 3$ is the number of colors (red, blue, green), and where $\gamma, \varepsilon > 0$ are some positive constants.

We continue with some related terminology. For an $(\ell, t, \gamma, \varepsilon)$-partition $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ of V, we shall say a triad $P = P_{abc}^{ijk} \in \mathrm{Triad}(\mathbf{\Pi})$ is $(\frac{1}{\ell}, \varepsilon)$-*typical* if $0 \notin \{a, b, c\}$, each of P_a^{ij}, P_b^{jk}, P_c^{ik} is ε-regular and if

$$\frac{1}{\ell} - \varepsilon \leq d_{P_a^{ij}}(V_i, V_j), d_{P_b^{jk}}(V_j, V_k), d_{P_c^{ik}}(V_i, V_k) \leq \frac{1}{\ell} + \varepsilon.$$

In other words, $P_{abc}^{ijk} = P_a^{ij} \cup P_b^{jk} \cup P_c^{ik}$ is $(\frac{1}{\ell}, \varepsilon)$-typical if and only if $0 \notin \{a, b, c\}$ and the pair $(P_{abc}^{ijk}, V_i \cup V_j \cup V_k)$ is $(\vec{d}, \varepsilon)$-regular where $\vec{d} = (d_{ij}, d_{jk}, d_{ik})$ satisfies $d_{ij} = d_{jk} = d_{ik} = \frac{1}{\ell}$. For an $(\ell, t, \gamma, \varepsilon)$-partition $\mathbf{\Pi}$ of V, we write

$$\mathrm{Triad}_{\mathrm{typ}}(\mathbf{\Pi}) = \left\{ P \in \mathrm{Triad}(\mathbf{\Pi}) : P \text{ is } \left(\frac{1}{\ell}, \varepsilon\right)\text{-typical} \right\} \tag{5.2}$$

for the family of all $(\frac{1}{\ell}, \varepsilon)$-typical triads $P \in \mathrm{Triad}(\mathbf{\Pi})$.

The set V above will be the vertex set of a 3-uniform hypergraph $\mathcal{G}$. In context, we will want the $(\ell, t, \gamma, \varepsilon)$-partition $\mathbf{\Pi}$ of V to be 'link-regular' with respect to $\mathcal{G}$, in the sense of Definition 3.3.5. In the following definition, we describe this condition precisely.

Definition 5.0.4 (δ-link regular partition)**.** Let $\mathcal{G}$ be a 3-uniform hypergraph with vertex set $V = V(\mathcal{G})$, and let $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ be an $(\ell, t, \gamma, \varepsilon)$-partition of V. For $\alpha_0, \delta > 0$, we say that $\mathcal{G}$ is (α_0, δ)-*link regular with respect to* $\mathbf{\Pi}$ if all but $\delta|V|^3$ many triples $\{x, y, z\} \in \mathcal{G}$ satisfy the following property: If $\{x, y, z\}$ belongs to a triad $P_{xyz} = P^{ijk}_{abc} \in \mathrm{Triad}(\mathbf{\Pi})$ where

(i) $P_{xyz} \in \mathrm{Triad}_{\mathrm{typ}}(\mathbf{\Pi})$ is $(\frac{1}{\ell}, \varepsilon)$-typical, and

(ii) $\mathcal{H}_{xyz} = \mathcal{G} \cap \mathcal{K}_3(P_{xyz})$ has density $\alpha_{xyz} = d_{\mathcal{H}_{xyz}}(P_{xyz}) \geq \alpha_0$,

then $(\mathcal{H}_{xyz}, P_{xyz}, V_i \cup V_j \cup V_k)$ is δ-link regular. In other words, whenever the $\mathcal{H}_{xyz}$-triad $(\mathcal{H}_{xyz}, P_{xyz}, V_i \cup V_j \cup V_k)$ is an $(\alpha_{xyz}, \vec{d}, \varepsilon)$-triad where $\vec{d} = (\frac{1}{\ell}, \frac{1}{\ell}, \frac{1}{\ell})$, then $(\mathcal{H}_{xyz}, P_{xyz}, V_i \cup V_j \cup V_k)$ is δ-link regular.

We now proceed to the promised theorem.

Theorem 5.0.5 (ALGORITHM $\mathbb{A}_{\mathrm{linkreg}}$: LINK-REGULARITY LEMMA)**.** *There exists an algorithm $\mathbb{A}_{\mathrm{linkreg}}$ which, for all $\alpha_0, \delta, \gamma > 0$, for all integers $\ell_0, t_0 \geq 1$, and for all functions $\boldsymbol{\varepsilon} : \mathbb{N} \to (0, 1)$, determines positive integers $L_0 = L_0(\alpha_0, \delta, \gamma, \ell_0, t_0, \boldsymbol{\varepsilon})$, $T_0 = T_0(\alpha_0, \delta, \gamma, \ell_0, t_0, \boldsymbol{\varepsilon})$, and $N_0 = N_0(\alpha_0, \delta, \gamma, \ell_0, t_0, \boldsymbol{\varepsilon})$ so that the following holds.*

Let $\mathcal{G}$ be a 3-uniform hypergraph with vertex set V, where $|V| = N \geq N_0$. Then, in time $O(N^3)$, the algorithm $\mathbb{A}_{\mathrm{linkreg}}$ constructs an $(\ell, t, \gamma, \boldsymbol{\varepsilon}(\ell))$-partition $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ of V, where $\ell_0 \leq \ell \leq L_0$, $t_0 \leq t \leq T_0$, and where $\mathcal{G}$ is (α_0, δ)-link regular with respect to $\mathbf{\Pi}$.

5.1 Proof of Theorem 5.0.5

We first collect some more notation and terminology that we use in our proof.

5.1.1 Notation and terminology

For a fixed $P \in \mathrm{Triad}(\mathbf{\Pi})$, recall that $\mathcal{G}_P = \mathcal{G} \cap \mathcal{K}_3(P)$ has density $d_{\mathcal{G}_P}(P) = |\mathcal{G}_P|/|\mathcal{K}_3(P)|$ (which could be zero). For $\alpha_0 > 0$, we define

$$\mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi}) = \{P \in \mathrm{Triad}_{\mathrm{typ}}(\mathbf{\Pi}) : \, d_{\mathcal{G}_P}(P) \geq \alpha_0\} \tag{5.3}$$

to denote those triads $P \in \mathrm{Triad}_{\mathrm{typ}}(\mathbf{\Pi})$ which are $(1/\ell, \varepsilon)$-typical and which also have 'non-trivial' density $d_{\mathcal{G}_P}(P) \geq \alpha_0$.

Continuing, we define the *index* of $\mathbf{\Pi}$ with respect to $\mathcal{G}$ by

$$\mathrm{ind}(\mathbf{\Pi}) = \frac{1}{N^3} \sum_{P \in \mathrm{Triad}(\mathbf{\Pi})} d^2_{\mathcal{G}_P}(P) \big|\mathcal{K}_3(P)\big|. \tag{5.4}$$

It is easy to see that $\mathrm{ind}(\mathbf{\Pi}) \leq 1$ never exceeds one. Indeed, the densities $d_{\mathcal{G}_P}(P) \leq 1$ never exceed one for all $P \in \mathrm{Triad}(\mathbf{\Pi})$, and $\sum_{P \in \mathrm{Triad}(\mathbf{\Pi})} |\mathcal{K}_3(P)| \leq \binom{|V|}{3} < N^3$.

Finally, we now define some important classes of triads $P \in \mathrm{Triad}(\mathbf{\Pi})$ of an $(\ell, t, \gamma, \varepsilon)$-partition $\mathbf{\Pi}$ of V. To begin, for $\delta > 0$ and an integer $r \geq 1$, define

$$\mathrm{Triad}_{(\delta,r)\text{-irr}}(\mathbf{\Pi}) = \Big\{ P \in \mathrm{Triad}(\mathbf{\Pi}) : \ \mathcal{G}_P = \mathcal{G} \cap \mathcal{K}_3^{(2)}(P) \text{ is } (\delta, r)\text{-irregular w.r.t. } P \Big\}. \tag{5.5}$$

In the definition above, the triads $P \in \mathrm{Triad}_{(\delta,r)\text{-irr}}(\mathbf{\Pi})$ need not be $(1/\ell, \varepsilon)$-typical.

5.1.2 An outline of the proof of Theorem 5.0.5

The proof of Theorem 5.0.5 follows the well-known lines of Szemerédi [33] and Alon et al. [1]. The proof uses Lemma 4.1.1 above, and the upcoming Lemma 5.1.1 below, which was originally due to Frankl and Rödl [9] (but which had roots in Szemerédi [33], and which is taken here from Haxell, Nagle, and Rödl [15]). Since this lemma is somewhat technical in appearance, we wish to motivate it by showing how we will use it within a sketch of the proof of Theorem 5.0.5. In this sketch, we will assume (for convenience) that all constants have been chosen suitably. Later, when we give all real details of the argument, we will describe precisely how such constants are determined.

For an N-vertex 3-uniform hypergraph $\mathcal{G}$, let $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ be an $(\ell, t, \gamma, \varepsilon)$-partition of $V = V(\mathcal{G})$, where $n = \lfloor N/t \rfloor = |V_1| = \cdots = |V_t|$. For a fixed triad $P = P^{ijk}_{abc} \in \mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi})$, we apply Lemma 4.1.1 to $(\mathcal{G}_P, P, V_i \cup V_j \cup V_k)$, where in time $O(n^3)$,

(i) either $\mathbb{A}_{\mathrm{linkreg}}$ confirms that $(\mathcal{G}_P, P, V_i \cup V_j \cup V_k)$ is δ-link regular,

(ii) or $\mathbb{A}_{\mathrm{linkreg}}$ constructs a witness $\vec{\mathcal{Q}}_P = \vec{\mathcal{Q}}_{r,P}$ of the $(\delta_{\#}, r)$-irregularity of $\mathcal{G}_P$ w.r.t. P.

In the former case, we have confirmed that $P \in \mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi})$ is at least one triad which would be desirable in the conclusion of Theorem 5.0.5. In the latter case, we learn that $\mathcal{G}_P$ is not $(\delta_{\#}, r)$-regular with respect to P, and $\mathbb{A}_{\mathrm{linkreg}}$ builds a witness $\vec{\mathcal{Q}}_P$ to this effect. In the latter case, we record this witness $\vec{\mathcal{Q}}_P$, and we maintain $\vec{\mathcal{Q}}_P$ with P. We now repeat the procedure above for every $P \in \mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi})$, of which there are at most $2\ell^3 t^3$ many (cf. (5.1)). Since ℓ and t will be constants not depending on N, Lemma 4.1.1 determines the outcome of (i) or (ii) above for every $P \in \mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi})$ in time $O(N^3)$.

Continuing, we count how many triples $\{x, y, z\} \in \binom{V}{3}$ belong to triads $P \in \mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi})$ where (i) occurred. If (i) occurred for nearly all $\{x, y, z\} \in \binom{V}{3}$, then $\mathbf{\Pi}$ will be the partition sought in Theorem 5.0.5. However, if some non-negligible portion of triples $\{x, y, z\} \in \binom{V}{3}$ belong to triads $P \in \mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi})$ where (ii) occurred, then $\mathbf{\Pi}$ will not be the partition sought in Theorem 5.0.5. In this case, we will replace $\mathbf{\Pi}$ with a new partition $\mathbf{\Pi}'$ of V, which we build from $\mathbf{\Pi}$ and the collection of witnesses $\vec{\mathcal{Q}}_P$ constructed in (ii). The mechanism for building $\mathbf{\Pi}'$ is precisely due to Lemma 5.1.1 below.

Lemma 5.1.1 (ALGORITHM $\mathbb{A}_{\mathrm{index}}$: INDEX-PUMPING)**.** *There exists an algorithm $\mathbb{A}_{\mathrm{index}}$ which, for all constants $\delta_{\#}$ and γ, integers ℓ_{old} and t_{old}, and functions $\boldsymbol{\varepsilon} : \mathbb{N}^+ \to (0,1)$ and $\boldsymbol{r} : \mathbb{N}^+ \to \mathbb{N}^+$, determines integer constants $L_0 = L_0(\delta_{\#}, \gamma, \boldsymbol{\varepsilon}, \boldsymbol{r}, \ell_{\mathrm{old}}, t_{\mathrm{old}})$, $T_0 = T_0(\delta_{\#}, \gamma, \boldsymbol{\varepsilon}, \boldsymbol{r}, \ell_{\mathrm{old}}, t_{\mathrm{old}})$, and $N_0 = N_0(\delta_{\#}, \gamma, \boldsymbol{\varepsilon}, \boldsymbol{r}, \ell_{\mathrm{old}}, t_{\mathrm{old}})$ so that the following holds:*

Let $\mathcal{G}$ be a given 3-graph on $N > N_0$ vertices with given $(\ell_{\mathrm{old}}, t_{\mathrm{old}}, \gamma, \boldsymbol{\varepsilon}(\ell_{\mathrm{old}}))$-partition $\mathbf{\Pi}_{\mathrm{old}}$ of $V = V(\mathcal{G})$. Let $\mathcal{T}_{\#} \subseteq \mathrm{Triad}_{(\delta_{\#}, \boldsymbol{r}(\ell_{\mathrm{old}}))\text{-irr}}(\mathbf{\Pi}_{\mathrm{old}})$ be a **given** *subfamily of the collection of all $(\delta_{\#}, \boldsymbol{r}(\ell_{\mathrm{old}}))$-irregular triads, which satisfies the following properties:*

(i) each triad $P \in \mathcal{T}_{\#}$ is equipped with a **given** *witness $\vec{\mathcal{Q}}_{\boldsymbol{r}(\ell_{\mathrm{old}}),P}$ of the $(\delta_{\#}, \boldsymbol{r}(\ell_{\mathrm{old}}))$-irregularity of $\mathcal{G}_P = \mathcal{G} \cap \mathcal{K}_3(P)$ with respect to P;*

(ii)
$$\sum_{P \in \mathcal{T}_{\#}} |\mathcal{K}_3(P)| \geq \delta_{\#} N^3.$$

Then, in time $O(N^2)$, the algorithm $\mathbb{A}_{\mathrm{index}}$ constructs (from $\mathbf{\Pi}_{\mathrm{old}}$ and the given collection of wit-

nesses $\{\vec{\mathcal{Q}}_{\boldsymbol{r}(\ell_{\text{old}}),P} : P \in \mathcal{T}_{\#}\}$*) an equitable* $(\ell_{\text{new}}, t_{\text{new}}, \gamma, \boldsymbol{\varepsilon}(\ell_{\text{new}}))$*-partition* $\boldsymbol{\Pi}_{\text{new}}$ *of* V *for which*

$$\operatorname{ind}(\boldsymbol{\Pi}_{\text{new}}) \geq \operatorname{ind}(\boldsymbol{\Pi}_{\text{old}}) + (\delta_{\#}^4/2),$$

for some integers $\ell_{\text{old}} \leq \ell_{\text{new}} \leq L_0$ *and* $t_{\text{old}} \leq t_{\text{new}} \leq T_0$.

We now conclude our outline of the proof of Theorem 5.0.5. Set $\mathcal{T}_{\#}$ to be the collection of triads $P \in \operatorname{Triad}_{\text{typ}}(\alpha_0, \boldsymbol{\Pi})$ for which the algorithm $\mathbb{A}_{\text{linkreg}}$ of Lemma 4.1.1 constructed a witness $\vec{\mathcal{Q}}_{r,P}$ of the $(\delta_{\#}, r)$-irregularity of $\mathcal{G}_P$ with respect to P. If $\mathcal{T}_{\#}$ claims many triples $\{x, y, z\} \in \binom{V}{3}$, meaning $\sum_{P \in \mathcal{T}_{\#}} |\mathcal{K}_3(P)| \geq \delta_{\#} N^3$, then we submit $\boldsymbol{\Pi}$ (as $\boldsymbol{\Pi}_{\text{old}}$), $\mathcal{T}_{\#}$, and $\{\vec{\mathcal{Q}}_{r,P} : P \in \mathcal{T}_{\#}\}$ to Lemma 5.1.1, all of which we have in hand. Algorithm $\mathbb{A}_{\text{index}}$ constructs, in time $O(N^2)$, the new partition $\boldsymbol{\Pi}_{\text{new}}$, where $\operatorname{ind}(\boldsymbol{\Pi}_{\text{new}}) \geq \operatorname{ind}(\boldsymbol{\Pi}_{\text{old}}) + (\delta_{\#}^4/2)$. We now repeat all the steps above from this subsection on $\boldsymbol{\Pi}_{\text{new}}$. However, this procedure can be repeated at most $2/\delta_{\#}^4$ many (constantly many) times, since the index function never exceeds one. Thus, some iteration of this procedure arrives at a partition $\boldsymbol{\Pi}$ as desired in Theorem 5.0.5.

5.1.3 Formal proof of Theorem 5.0.5

It remains only to fill in a few formal details of the outline above. We begin with a precise description of all parameters used above in the argument.

Constants of Theorem 5.0.5

Let $\alpha_0, \delta, \gamma > 0$ be given as well as function $\boldsymbol{\varepsilon} : \mathbb{N} \to (0, 1)$. Let integers $\ell_0, t_0 \geq 1$ be given, where for simplicity we take $\ell_0 = 1$. We now consider several auxiliary constants. For α_0 and δ given above, let

$$\delta_{\#} = \delta_{\#,\text{Lem.4.1.1}}(\alpha_0, \delta) \tag{5.6}$$

be the constant guaranteed by Lemma 4.1.1. It follows from the proof of Lemma 4.1.1 that

$$\delta_{\#} \leq \delta. \tag{5.7}$$

Let $\ell \in \mathbb{N}$ be an integer *variable*. In the context of Lemma 4.1.1, set $d_0 = d_0(\ell) = 1/\ell$ and let

$$\varepsilon_{\text{Lem.4.1.1}}(1/\ell) = \varepsilon_{\text{Lem. 4.1.1}}(\alpha_0, \delta, \delta_\#, 1/\ell) \quad \text{and} \quad \boldsymbol{r}(\ell) = r_{\text{Lem.4.1.1}}(\alpha_0, \delta, \delta_\#, 1/\ell) \tag{5.8}$$

be the *functions* (of the variable $d_0 = 1/\ell$) guaranteed by Lemma 4.1.1. We may assume, without loss of generality, that the given function $\boldsymbol{\varepsilon}$ satisfies, for every integer $\ell \in \mathbb{N}$,

$$\boldsymbol{\varepsilon}(\ell) \leq \varepsilon_{\text{Lem.4.1.1}}(1/\ell). \tag{5.9}$$

Theorem 5.0.5 promises integer constants L_0, T_0 and N_0, which we now formally describe (but which will be more easily understood in context). With $\gamma > 0$ given above, $\delta_\#$ given in (5.6), functions $\boldsymbol{\varepsilon}(\ell)$ and $\boldsymbol{r}(\ell)$ given in (5.8) and (5.9), and for arbitrary integer *variables* $\ell_{\text{old}}, t_{\text{old}} \geq 1$, Lemma 5.1.1 guarantees integer *functions* (of the variables $\ell_{\text{old}}, t_{\text{old}}$)

$$L_0(\ell_{\text{old}}, t_{\text{old}}) = L_0(\gamma, \delta_\#, \boldsymbol{\varepsilon}, \boldsymbol{r}, \ell_{\text{old}}, t_{\text{old}}), \quad T_0(\ell_{\text{old}}, t_{\text{old}}) = T_0(\gamma, \delta_\#, \boldsymbol{\varepsilon}, \boldsymbol{r}, \ell_{\text{old}}, t_{\text{old}}),$$
$$\text{and} \quad N_0(\ell_{\text{old}}, t_{\text{old}}) = N_0(\gamma, \delta_\#, \boldsymbol{\varepsilon}, \boldsymbol{r}, \ell_{\text{old}}, t_{\text{old}}). \tag{5.10}$$

We successively define constants $L_0^{(i)}$, $T_0^{(i)}$ and $N_0^{(i)}$, for each $0 \leq i \leq 2/\delta_\#^4$, as follows: with given integers $\ell_0 = 1$ and t_0, let

$$L_0^{(0)} = L_0(\ell_0 = 1, t_0), \quad T_0^{(0)} = T_0(\ell_0 = 1, t_0), \quad N_0^{(0)} = N_0(\ell_0 = 1, t_0)$$

be given by the functions in (5.10). For $1 \leq i \leq 2/\delta_\#^4$, let

$$L_0^{(i)} = L_0(L_0^{(i-1)}, T_0^{(i-1)}), \quad T_0^{(i)} = T_0(L_0^{(i-1)}, T_0^{(i-1)}), \quad N_0^{(i)} = N_0(L_0^{(i-1)}, T_0^{(i-1)}) \tag{5.11}$$

be given by the functions in (5.10). Then, the constants L_0, T_0 and N_0 of Theorem 5.0.5 are given by

$$L_0 = L_0^{(i_*)}, \quad T_0 = T_0^{(i_*)}, \quad N_0 = N_0^{(i_*)}, \quad \text{where} \quad i_* = \left\lfloor 2/\delta_\#^4 \right\rfloor. \tag{5.12}$$

This concludes our discussion of the constants.

The argument

With the constants chosen above, we continue to fill in a few further details of the earlier outline. Let $\mathcal{G}$ be an N-vertex 3-graph with vertex set $V = V(\mathcal{G})$, where in all that follows, we assume that $N \geq N_0$ (from (5.12)), and more generally that N is sufficiently large whenever needed. (Thus, L_0 and T_0 from (5.12) are constants which are independent of N.) We will construct, in time $O(N^3)$, an $(\ell, t, \gamma, \boldsymbol{\varepsilon}(\ell))$ partition $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ of V with respect to which $\mathcal{G}$ is (α_0, δ)-link regular, where $1 = \ell_0 \leq \ell \leq L_0$ and $t_0 \leq t \leq T_0$ (cf. (5.12)).

We start by taking $\Pi_1^{(1)}$ to be any vertex partition $V = V_0 \cup V_1 \cup \cdots \cup V_{t_0}$ satisfying $|V_0| < t_0$ and $|V_1| = \cdots = |V_{t_0}|$. Let $\Pi_1^{(2)}$ be the partition of $K[V_1, \ldots, V_{t_0}]$ where, for each $1 \leq i < j \leq t_0$, $K[V_i, V_j]$ is its own class, i.e., $\ell_{ij} = 1$ and $P_1^{ij} = K[V_i, V_j]$. Then, $\mathbf{\Pi}_1 = (\Pi_1^{(1)}, \Pi_1^{(2)})$ is an equitable $(\ell_0 = 1, t_0, \gamma, \boldsymbol{\varepsilon}(\ell_0))$-partition because complete bipartite graphs $K[X, Y]$ are always $o(1)$-regular with density 1.

For an integer $1 \leq s < 2/\delta_{\#}^4$, assume $\mathbf{\Pi}_1, \ldots, \mathbf{\Pi}_s$ are inductively constructed, where $\mathbf{\Pi}_s$ is an equitable $(\ell_s, t_s, \gamma, \boldsymbol{\varepsilon}(\ell_s))$-partition of V, and where

$$1 = \ell_0 \leq \ell_s \leq L_0^{(s-1)} \quad \text{and} \quad t_0 \leq t_s \leq T_0^{(s-1)} \tag{5.13}$$

for the constants $L_0^{(s-1)}$ and $T_0^{(s-1)}$ defined in (5.11). Now, for each $P_{abc}^{ijk} \in \mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi}_s)$, we apply algorithm $\mathbb{A}_{\mathrm{linkreg}}$ of Lemma 4.1.1 to $(\mathcal{G}_P, P, V_i \cup V_j \cup V_k)$, where $n_s = \lfloor N/t_s \rfloor = |V_i| = |V_j| = |V_k|$. In time $O(n_s^3)$,

(i) either $\mathbb{A}_{\mathrm{linkreg}}$ confirms that $(\mathcal{G}_P, P, V_i \cup V_j \cup V_k)$ is δ-link regular,

(ii) or $\mathbb{A}_{\mathrm{linkreg}}$ constructs a witness $\vec{\mathcal{Q}}_{\boldsymbol{r}(\ell_s),P}$ of the $(\delta_{\#}, \boldsymbol{r}(\ell_s))$-irregularity of $\mathcal{G}_P$ w.r.t. P.

We repeat algorithm $\mathbb{A}_{\mathrm{linkreg}}$ over all triads $P \in \mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi}_s)$, which takes time $O(N^3)$, since by (5.1), (5.11) and (5.12) there are $\leq 2\ell_s^3 t_s^3 \leq 2(L_0^{(s-1)} T_0^{(s-1)})^3 \leq 2(L_0 T_0)^3 = O(1)$ many such triads. Now, let $\mathcal{T}_{\#} = \mathcal{T}_{\#}(s)$ be the collection of those triads $P \in \mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi}_s)$ for which $\mathbb{A}_{\mathrm{linkreg}}$ reports outcome (ii), where we also collect the corresponding system of constructed witnesses $\{\vec{\mathcal{Q}}_{\boldsymbol{r}(\ell_s),P} : P \in \mathcal{T}_\#\}$. Finally, in time $O(N^3)$, we greedily compute

$$\sum_{P \in \mathcal{T}_\#} |\mathcal{K}_3(P)|.$$

We now consider two cases. On the one hand, if

$$\sum_{P \in \mathcal{T}_\#} |\mathcal{K}_3(P)| \le \delta N^3,$$

then $\mathbf{\Pi}_s$ is the desired partition. Indeed, the inequality above says that all but δN^3 triples $\{x, y, z\} \in \binom{V}{3}$ belong to triads $P_{xyz} = P^{ijk}_{abc} \in \mathrm{Triad}_{\mathrm{typ}}(\alpha_0, \mathbf{\Pi}_s)$ which are $(1/\ell_s, \boldsymbol{\varepsilon}(\ell_s))$-typical, which have density $d_{\mathcal{G}_{P_{xyz}}}(P_{xyz}) \ge \alpha_0$, and for which $(\mathcal{G}_{P_{xyz}}, P_{xyz}, V_i \cup V_j \cup V_k)$ is δ-link regular. By Definition 5.0.4, this precisely means that $\mathcal{G}$ is (α_0, δ)-link regular with respect to $\mathbf{\Pi}_s$, as desired. On the other hand, if

$$\sum_{P \in \mathcal{T}_\#} |\mathcal{K}_3(P)| > \delta N^3, \quad \text{then by (5.7) we also have} \quad \sum_{P \in \mathcal{T}_\#} |\mathcal{K}_3(P)| > \delta_\# N^3. \tag{5.14}$$

Now, the constructed collections $\mathcal{T}_\#$ and $\{\vec{\mathcal{Q}}_{\boldsymbol{r}(\ell_s),P} : P \in \mathcal{T}_\#\}$ and the condition in (5.14) precisely meet the hypothesis of Lemma 5.1.1. As such, $\mathbb{A}_{\mathrm{index}}$ constructs, in time $O(N^2)$, an equitable $(\ell_{s+1}, t_{s+1}, \gamma, \boldsymbol{\varepsilon}(\ell_{s+1}))$-partition $\mathbf{\Pi}_{s+1}$ of V for which

$$\mathrm{ind}(\mathbf{\Pi}_{s+1}) \ge \mathrm{ind}(\mathbf{\Pi}_s) + (\delta_\#^4/2), \tag{5.15}$$

and for which

$$\ell_{s+1} \le L_0(\ell_s, t_s) \overset{(5.13)}{\le} L_0(L_0^{(s-1)}, T_0^{(s-1)}) \overset{(5.11)}{=} L_0^{(s)}$$
$$\text{and} \quad t_0 \le t_{s+1} \le T_0(\ell_s, t_s) \overset{(5.13)}{\le} T_0(L_0^{(s-1)}, T_0^{(s-1)}) \overset{(5.11)}{=} T_0^{(s)}.$$

Thus, we have completed an inductive construction for partition $\mathbf{\Pi}_{s+1}$, which by (5.15), can only be done at most $2/\delta_\#^4$ many times. Thus, on some iteration, we must construct an equitable

$(\ell, t, \gamma, \boldsymbol{\varepsilon}(\ell))$-partition $\mathbf{\Pi} = (\Pi^{(1)}, \Pi^{(2)})$ with respect to which $\mathcal{G}$ is (α_0, δ)-link regular, where $\ell_0 \leq \ell \leq L_0$ and $t_0 \leq t \leq T_0$ (cf. (5.12)).

Chapter 6

A Transference of Regularity

The goal of this chapter is to prove a transference lemma which allows us to infer the notion of regularity of Theorem 1.4.2 from that of Theorem 1.4.3.

Suppose $(\mathcal{H}, P, V)$ is a δ-link regular $(\alpha, \vec{d}, \varepsilon)$-triad, where $\vec{d} = (d_{12}, d_{13}, d_{23})$, $V = V_1 \dot\cup V_2 \dot\cup V_3$, and $P = P_{12} \dot\cup P_{13} \dot\cup P_{23}$. Our transference lemma says that when 'large' subsets $U_i \subseteq V_i$, $1 \le i \le 3$, are given, they must induce close to the 'expected number' of triples from $\mathcal{H}$.

Lemma 6.0.1 (Transference Lemma)**.** *For all $\omega_0 > 0$, there exists $\delta > 0$ so that, for all $d_0 > 0$, there exists $\varepsilon > 0$ so that the following holds. Let $(\mathcal{H}, P, V)$ be a δ-link regular, $(\alpha, \vec{d}, \varepsilon)$-triad, where $\alpha \ge 0$ is arbitrary, $\vec{d} = (d_{12}, d_{13}, d_{23})$ has $d_{12}, d_{13}, d_{23} \ge d_0$, $V = V_1 \dot\cup V_2 \dot\cup V_3$ has each $|V_1|$, $|V_2|$, $|V_3|$ sufficiently large, and $P = P_{12} \dot\cup P_{13} \dot\cup P_{23}$. Let subsets $U_i \subseteq V_i$, $1 \le i \le 3$, be given satisfying $|U_i| > \omega_0 |V_i|$, $1 \le i \le 3$. Then $U_1 \dot\cup U_2 \dot\cup U_3$ induce*

$$|\mathcal{H}[U_1, U_2, U_3]| = (1 \pm \omega_0)\alpha d_{12} d_{13} d_{23} |U_1||U_2||U_3|$$

many triples from $\mathcal{H}$.

Remark. Note that the lemma above is trivial when $\alpha = 0$. Moreover, when $\alpha = 0$, the conclusion holds with no hypothesis on $P = P_{12} \dot\cup P_{13} \dot\cup P_{23}$.

Proof. Let $\omega_0 > 0$ be given. Set

$$\delta = \frac{\omega_0^2}{144}. \tag{6.1}$$

Let $d_0 > 0$ be given. Set

$$\varepsilon = \min\left\{\frac{1}{2}d_0, \frac{1}{8}\delta\right\}. \tag{6.2}$$

Now, let $(\mathcal{H}, P, V)$ be a δ-link regular, $(\alpha, \vec{d}, \varepsilon)$-triad, where $\alpha \geq 0$ is arbitrary, $\vec{d} = (d_{12}, d_{13}, d_{23})$ has $d_{12}, d_{13}, d_{23} \geq d_0$, $V = V_1 \dot{\cup} V_2 \dot{\cup} V_3$ has each $|V_1|$, $|V_2|$, $|V_3|$ sufficiently large, and $P = P_{12} \dot{\cup} P_{13} \dot{\cup} P_{23}$. Let $U_i \subseteq V_i$ be subsets, $1 \leq i \leq 3$, satisfying $|U_i| > \omega_0 |V_i|$. For simplicity, we show the lower bound only:

$$|\mathcal{H}[U_1, U_2, U_3]| \geq (1 - \omega_0)\alpha d_{12} d_{13} d_{23} |U_1||U_2||U_3|. \tag{6.3}$$

(In most applications, only the lower bound is needed.) The proof of the corresponding upper bound follows from symmetric arguments. Now, the main idea for proving (6.3) is not difficult. First, we use the identity

$$|\mathcal{H}[U_1, U_2, U_3]| = \sum_{u \in U_1} \left| L_u \left[N_{P^{12}}(u) \cap U_2, N_{P^{13}}(u) \cap U_3 \right] \right|. \tag{6.4}$$

Second, we use the hypothesis that all but $\delta |V_1|$ many vertices $u \in U_1$ satisfy that L_u is $(\alpha d_{23}, \delta, 1)$-regular. We denote these vertices by U_1^L, where

$$|U_1^L| \geq |U_1| - \delta |V_1|. \tag{6.5}$$

Then,

$$|\mathcal{H}[U_1, U_2, U_3]| \geq \sum_{u \in U_1^L} \left| L_u \left[N_{P^{12}}(u) \cap U_2, N_{P^{13}}(u) \cap U_3 \right] \right|.$$

Third, to use the $(\alpha d_{23}, \delta, 1)$-regularity of each link L_u, where $u \in U_1^L$, we use the following standard properties of the $(\vec{d}, \varepsilon)$-triad P: All but

- $2\varepsilon |V_1|$ vertices $u \in U_1$ satisfy $|N_{P^{12}}(u) \cap U_2| = (d_{12} \pm \varepsilon)|U_2|$;
- $2\varepsilon |V_1|$ vertices $u \in U_1$ satisfy $|N_{P^{12}}(u)| = (d_{12} \pm \varepsilon)|V_2|$;
- $2\varepsilon |V_1|$ vertices $u \in U_1$ satisfy $|N_{P^{13}}(u) \cap U_3| = (d_{13} \pm \varepsilon)|U_3|$;
- $2\varepsilon |V_1|$ vertices $u \in U_1$ satisfy $|N_{P^{13}}(u)| = (d_{13} \pm \varepsilon)|V_3|$.

Thus, all but $8\varepsilon|V_1|$ vertices $u \in U_1$ satisfy

$$\frac{|N_{P^{12}}(u) \cap U_2|}{|N_{P^{12}}(u)|} \geq \frac{d_{12} - \varepsilon}{d_{12} + \varepsilon} \cdot \frac{|U_2|}{|V_2|} \geq \frac{1 - \varepsilon/d_0}{1 + \varepsilon/d_0} \cdot \omega_0 \geq \frac{1 - 1/2}{1 + 1/2} \cdot \omega_0 \geq \frac{\omega_0}{4} \tag{6.6}$$

and similarly, $|N_{P^{13}}(u) \cap U_3| \geq (\omega_0/4)|N_{P^{13}}(u)|$. (In the calculations above, we used $d_{12} \geq d_0 \geq 2\varepsilon$ from (6.2) and $|U_2| \geq \omega_0|V_2|$.) We denote these vertices $u \in U_1$ by U_1^P, where we have

$$|U_1^P| \geq |U_1| - 8\varepsilon|V_1|. \tag{6.7}$$

and we set $U_1^{L,P} = U_1^L \cap U_1^P$, where (6.5) and (6.7) give

$$|U_1^{L,P}| \geq |U_1| - \delta|V_1| - 8\varepsilon|V_1|. \tag{6.8}$$

Now,

$$|\mathcal{H}[U_1, U_2, U_3]| \geq \sum_{u \in U_1^{L,P}} \left| L_u \left[N_{P^{12}}(u) \cap U_2, N_{P^{13}}(u) \cap U_3 \right] \right|. \tag{6.9}$$

where the properties of $U_1^{L,P}$ will allow us to bound each term above. Fix $u \in U_1^{L,P}$. In (6.6) we saw

$$|N_{P^{12}}(u) \cap U_2| \geq \frac{\omega_0}{4}|N_{P^{12}}(u)|$$

and

$$|N_{P^{13}}(u) \cap U_3| \geq \frac{\omega_0}{4}|N_{P^{13}}(u)|$$

In the setting of Definition 2.0.1, set

$$\begin{aligned} A' &= N_{P^{12}}(u) \cap U_2, \qquad & A &= N_{P^{12}}(u), \\ B' &= N_{P^{13}}(u) \cap U_3, \qquad & B &= N_{P^{13}}(u). \end{aligned}$$

Then, the properties above imply that

$$|K(A', B') \geq \frac{\omega_0^2}{16}|A||B|,$$

so the $(\alpha d_{23}, \delta, 1)$-regularity of L_u implies that

$$\alpha d_{23}(1-\delta) \leq d_{L_u}(A', B') \leq \alpha d_{23}(1+\delta),$$

or equivalently

$$|L_u[N_{P^{12}}(u) \cap U_2, N_{P^{13}}(u) \cap U_3]| = \alpha d_{23}(1 \pm \delta)|N_{P^{12}}(u) \cap U_2||N_{P^{13}}(u) \cap U_3|.$$

Thus, we conclude from (6.9) that

$$\begin{aligned} |\mathcal{H}[U_1, U_2, U_3]| &\geq \alpha d_{23}(1-\delta) \sum_{u \in U_1^{L,P}} |N_{P^{12}}(u) \cap U_2||N_{P^{13}}(u) \cap U_3| \\ &\geq \alpha d_{12} d_{13} d_{23}(1-\delta)(1-\varepsilon)^3 \left|U_1^{L,P}\right| |U_2||U_3|, \end{aligned}$$

where we used the properties defined by U_1^P. Recalling (6.8), we further conclude

$$|\mathcal{H}[U_1, U_2, U_3]| \geq \alpha d_{12} d_{13} d_{23}(1-\delta)(1-\varepsilon)^3 \left(1 - \delta\frac{|V_1|}{|U_1|} - 8\varepsilon\frac{|V_1|}{|U_1|}\right)|U_1||U_2||U_3|.$$

Then $|U_1| > \omega_0|V_1|$ ensures

$$\begin{aligned} |\mathcal{H}[U_1, U_2, U_3]| &\geq \alpha d_{12} d_{13} d_{23}(1-\delta)(1-\varepsilon)^3 \left(1 - \frac{\delta}{\omega_0} - \frac{8\varepsilon}{\omega_0}\right)|U_1||U_2||U_3| \\ &\geq \alpha d_{12} d_{13} d_{23} \left(1 - \frac{2\delta}{\omega_0}\right)^5 |U_1||U_2||U_3|, \end{aligned}$$

where we used $8\varepsilon \leq \delta$ from (6.2). Since $\delta \leq \omega_0/144$ from (6.1), we have

$$\left(1 - \frac{2\delta}{\omega_0}\right)^5 \geq 1 - 10\frac{\delta}{\omega_0} - 80\frac{\delta^3}{\omega_0^3} - 32\frac{\delta^5}{\omega_0^5} \geq 1 - 122\frac{\delta}{\omega_0} \geq 1 - \omega_0,$$

which proves (6.3). $\square$

Chapter 7

Proof of Theorem 1.4.2

In this chapter, we prove Theorem 1.4.2, which will give the algorithm $\mathbb{A}_{\text{main}}$ for constructing, in time $O(n^3)$, an ω-weakly regular and t-equitable vertex partition $\Pi^{(1)} : V = V_0 \dot\cup \ldots \dot\cup V_t$ of an n-vertex 3-uniform hypergraph $\mathcal{G}$. The algorithm $\mathbb{A}_{\text{main}}$ is precisely given by $\mathbb{A}_{\text{linkreg}}$ of Theorem 5.0.5, which provides the stronger partition $\mathbf{\Pi} = \left(\Pi^{(1)}, \Pi^{(2)}\right)$ described in Chapter 5. It will be easy to show that the vertex partition $\Pi^{(1)}$ of $\mathbf{\Pi}$ is, in fact, ω-weakly regular. We proceed to the details, and begin by discussing the constants involved.

7.1 Constants

Let $\omega > 0$ and integer $t_0 \geq 1$ be given. We define a collection of auxiliary constants as follows:

$$\mu = \tfrac{\omega^8}{32}, \quad \alpha_0 = \omega_0 = \gamma = \delta_1 = \tfrac{\mu}{6}, \quad t_0' = t_0 \left\lceil \tfrac{12}{\mu} \right\rceil. \tag{7.1}$$

Now, with $\omega_0 > 0$ above, let

$$\delta_{\text{transf}} = \delta_{\text{transf}}(\omega_0) \tag{7.2}$$

be the constant guaranteed by Lemma 6.0.1. Set

$$\delta = \min\{\delta_1, \delta_{\text{transf}}\}. \tag{7.3}$$

For an integer variable $\ell \in \mathbb{N}$, set

$$d_0(\ell) = \frac{1}{2\ell} \text{ and define } \varepsilon(\ell) = \varepsilon_{\text{transf}}(d_0(\ell)) \tag{7.4}$$

to be the function (of $\ell \in \mathbb{N}$) guaranteed by Lemma 6.0.1. Now, with the integer $t'_0 \geq t_0$ from (7.1) and $\ell_0 = 1$, with constants $\alpha_0, \gamma > 0$ from (7.1) and $\delta > 0$ from (7.3) and with the function $\varepsilon : \mathbb{N} \to (0, 1)$ from (7.4), let

$$\begin{aligned} L_0 &= L_0(\alpha_0, \gamma, \delta, t'_0, \ell_0 = 1, \varepsilon), \\ T_0 &= T_0(\alpha_0, \gamma, \delta, t'_0, \ell_0 = 1, \varepsilon), \\ N_0 &= N_0(\alpha_0, \gamma, \delta, t'_0, \ell_0 = 1, \varepsilon) \end{aligned} \tag{7.5}$$

be the positive integer constants guaranteed by Theorem 5.0.5.

7.2 The Algorithm $\mathbb{A}_{\textbf{main}} = \mathbb{A}_{\textbf{linkreg}}$

Let $\mathcal{G}$ be a given 3-uniform hypergraph on N vertices, where whenever needed, we assume N is sufficiently large. (In particular, we assume $N \geq N_0$ from (7.5).) The algorithm $\mathbb{A}_{\text{main}}$ is the algorithm $\mathbb{A}_{\text{linkreg}}$ of Theorem 5.0.5. In particular, and in time $O(N^3)$, an $(\ell, t, \gamma, \varepsilon(\ell))$-partition $\mathbf{\Pi} = \left(\Pi^{(1)}, \Pi^{(2)}\right)$ of the vertex set $V = V(\mathcal{G})$, for some integers $1 = \ell_0 \leq \ell = \ell_{\mathcal{G}} \leq L_0$ and $t'_0 \leq t = t_{\mathcal{G}} \leq T_0$, where $\mathcal{G}$ is (α_0, δ)-link regular with respect to $\mathbf{\Pi}$, we claim that the vertex partition $\Pi^{(1)} : V = V_1 \dot\cup \ldots \dot\cup V_t$ is ω-weakly regular, and the remainder of this chapter is devoted to these details. In particular, the remainder of this section is devoted to proving the following proposition.

Proposition 7.2.1. *With the constants $\mu, \omega > 0$ in* (7.1)*, there exists a subhypergraph $\mathcal{G}_0 \subseteq \mathcal{G}$ of size $|\mathcal{G}_0| \geq |\mathcal{G}| - \mu N^3$ with the property that, for every $1 \leq i < j < k \leq t$, $\mathcal{G}_0$ is $(2\omega_0)$-weakly regular with respect to (V_i, V_j, V_k).*

Observe that Proposition 7.2.1 implies Theorem 1.4.2. Indeed, we infer from Proposition 7.2.1 that

$$\sum_{1 \leq i < j < k \leq t} |(\mathcal{G} \setminus \mathcal{G}_0)[V_i, V_j, V_k]| \leq |\mathcal{G} \setminus \mathcal{G}_0| < \mu N^3. \tag{7.6}$$

Thus, all but $\sqrt{8\mu}t^3$ triples of indices $1 \le i < j < j \le t$ have

$$|(\mathcal{G} \setminus \mathcal{G}_0)\,[V_i, V_j, V_k]| \le \sqrt{8\mu}|V_i||V_j|V_k| = \sqrt{8\mu}\Big\lfloor \frac{N}{t} \Big\rfloor^3, \tag{7.7}$$

since otherwise we would have

$$\begin{aligned}|\mathcal{G} \setminus \mathcal{G}_0| &\ge \sum_{1 \le i < j < k \le t} |(\mathcal{G} \setminus \mathcal{G}_0)\,[V_i, V_j, V_k]| \ge \sqrt{8u}t^3 \cdot \sqrt{8u}\Big\lfloor \frac{N}{t} \Big\rfloor^3 \\ &\ge 8\mu t^3 (N/(2t))^3 = \mu N^3,\end{aligned}$$

contradicting (7.6). Now, we shall say that $\{i, j, k\} \in \binom{[t]}{3}$ is $\mathcal{G}$-*loyal* if (7.7) holds, and we denote by $\binom{[t]}{3}_{\text{loyal}}$ the set of all $\mathcal{G}$-loyal triples of indices $1 \le i < j < k \le t$. Theorem 1.4.2 is now concluded by the following fact.

Fact 7.2.2. *For each $\{i, j, k\} \in \binom{[t]}{3}_{loyal}$ we have that $\mathcal{G}$ is ω-weakly regular with respect to (V_i, V_j, V_k). Consequently, since all but $\sqrt{8\mu}t^3 < \omega t^3$ (cf. (7.1)) elements $i, j, k \in \binom{[t]}{3}$ are $\mathcal{G}$-loyal, we have that $V = V_0 \,\dot\cup\, V_1 \,\dot\cup \ldots \dot\cup\, V_t$ is an ω-weakly regular partition of $\mathcal{G}$.*

Proof of Fact 7.2.2. Fix $\{i, j, k\} \in \binom{[t]}{3}_{\text{loyal}}$ and fix $V_i' \subseteq V_i$, $V_j' \subseteq V_j$, and $V_k' \subseteq V_k$ where $|V_i'| > \omega|V_i| \overset{(7.1)}{>} 2\omega_0|V_i|$, $|V_j'| > \omega|V_j| \overset{(7.1)}{>} 2\omega_0|V_j|$, and $|V_k'| > \omega|V_k| \overset{(7.1)}{>} 2\omega_0|V_k|$. Since

$$|\mathcal{G}[V_i, V_j, V_k]| = |\mathcal{G}_0[V_i, V_j, V_k]| + |(\mathcal{G} \setminus \mathcal{G}_0)\,[V_i, V_j, V_k]|$$

and similarly

$$\left|\mathcal{G}[V_i', V_j', V_k']\right| = \left|\mathcal{G}_0[V_i', V_j', V_k']\right| + \left|(\mathcal{G} \setminus \mathcal{G}_0)\,[V_i', V_j', V_k']\right|,$$

we infer

$$d_{\mathcal{G}}(V_i, V_j, V_k) = d_{\mathcal{G}_0}(V_i, V_j, V_k) + d_{\mathcal{G} \setminus \mathcal{G}_0}(V_i, V_j, V_k)$$

and

$$d_{\mathcal{G}}(V_i', V_j', V_k') = d_{\mathcal{G}_0}(V_i', V_j', V_k') + d_{\mathcal{G} \setminus \mathcal{G}_0}(V_i', V_j', V_k').$$

Subtracting, we obtain

$$\begin{aligned}
&d_{\mathcal{G}}(V_i, V_j, V_k) - d_{\mathcal{G}}(V_i', V_j', V_k') \\
&\qquad = \left[d_{\mathcal{G}_0}(V_i, V_j, V_k) - d_{\mathcal{G}_0}(V_i', V_j', V_k')\right] + \left[d_{\mathcal{G}\setminus\mathcal{G}_0}(V_i, V_j, V_k) - d_{\mathcal{G}\setminus\mathcal{G}_0}(V_i', V_j', V_k')\right].
\end{aligned}$$

Thus, the triangle inequality gives

$$\begin{aligned}
&|d_{\mathcal{G}}(V_i, V_j, V_k) - d_{\mathcal{G}}(V_i', V_j', V_k')| \\
&\qquad = \left|d_{\mathcal{G}_0}(V_i, V_j, V_k) - d_{\mathcal{G}_0}(V_i', V_j', V_k')\right| + \left|d_{\mathcal{G}\setminus\mathcal{G}_0}(V_i, V_j, V_k) - d_{\mathcal{G}\setminus\mathcal{G}_0}(V_i', V_j', V_k')\right| \\
&\qquad < 2\omega_0 + \max\left\{d_{\mathcal{G}\setminus\mathcal{G}_0}(V_i, V_j, V_k), d_{\mathcal{G}\setminus\mathcal{G}_0}(V_i', V_j', V_k')\right\} \\
&\qquad < 2\omega_0 + \max\left\{\mu, d_{\mathcal{G}\setminus\mathcal{G}_0}(V_i', V_j', V_k')\right\}, \qquad (7.8)
\end{aligned}$$

where we first use the $(2\omega_0)$-weak regularity of $\mathcal{G}_0$ with respect to (V_i, V_j, V_k), and second, used the $\mathcal{G}$-loyalty of $\{i, j, k\} \in \binom{[t]}{3}_{\text{loyal}}$. Moreover, again using the $\mathcal{G}$-loyalty of $\{i, j, k\} \in \binom{[t]}{3}_{\text{loyal}}$, we see

$$\left|(\mathcal{G} \setminus \mathcal{G}_0)\left[V_i', V_j', V_k'\right]\right| \leq |(\mathcal{G} \setminus \mathcal{G}_0)\left[V_i, V_j, V_k\right]| \leq \sqrt{8\mu}|V_i||V_j|V_k|,$$

and so

$$d_{\mathcal{G}\setminus\mathcal{G}_0}(V_i', V_j', V_k') < \sqrt{8\mu}\frac{|V_i||V_j|V_k|}{|V_i'||V_j'|V_k'|} < \frac{\sqrt{8\mu}}{\omega^3}.$$

Returning to (7.8) we conclude

$$|d_{\mathcal{G}}(V_i, V_j, V_k) - d_{\mathcal{G}}(V_i', V_j', V_k')| < 2\omega_0 + \frac{\sqrt{8\mu}}{\omega^3} < \omega,$$

as desired. □

It only remains to prove Proposition 7.2.1

Proof of Proposition 7.2.1. The subhypergraph $\mathcal{G}_0 \subseteq \mathcal{G}$ is obtained from $\mathcal{G}$ by deleting triples $\{x, y, z\} \in \mathcal{G}$ which meet any of the following conditions (which will be enumerated in a moment):

(i) $\{x, y, z\} \cap V_0 \neq \emptyset$ or $|\{x, y, z\} \cap V_i| \geq 2$ for some $1 \leq i \leq t$;

(ii) At least one of $\{x,y\}$, $\{y,z\}$, or $\{x,z\}$ belongs to a class $P_0^{ij} \in \Pi^{(2)}$;

(iii) At least one of $\{x,y\}$, $\{y,z\}$, or $\{x,z\}$, e.g. $\{x,y\}$ satisfies $x \in V_i$ and $y \in V_j$ where the indices $1 \leq i < j \leq t$ fail to satisfy $|P_0^{ij}| < \gamma |V_i||V_j|$ and $|P_a^{ij}| = \left(\frac{1}{\ell} \pm \varepsilon\right) |V_i||V_j|$ for all $1 \leq a \leq \ell_{ij}$;

(iv) At least one of $\{x,y\}$, $\{y,z\}$, or $\{x,z\}$ belongs to a class $P_a^{ij} \in \Pi^{(2)}$ which is not $\varepsilon(\ell)$-regular;

(v) $\{x,y,z\} \in \mathcal{K}_3(P_{xyz})$ for a triad $P_{xyz} = P_{abc}^{ijk} = P_a^{ij} \dot\cup P_b^{jk} \dot\cup P_c^{ik} \in$ Triad($\mathbf{\Pi}$) for which $\mathcal{G}_{xyz} = \mathcal{G} \cap \mathcal{K}_3(P_{xyz})$ satisfies $|\mathcal{G}_{xyz}| < \alpha_0 |\mathcal{K}_3(P_{xyz})|$;

(vi) $\{x,y,z\} \in \mathcal{K}_3(P_{xyz})$ for a $\left(\frac{1}{\ell}, \varepsilon\right)$-typical triad $P_{xyz} = P_{abc}^{ijk} \in$ Triad$_{\text{typ}}$($\mathbf{\Pi}$) (recall (5.2)) with density $d_{\mathcal{G}_{xyz}}(P_{xyz}) \geq \alpha_0$, but where $(\mathcal{G}_{xyz}, P_{xyz}, V_i \dot\cup V_j \dot\cup V_k)$ is not δ-link regular.

Conditions (i)-(vi) define the family $\mathcal{G} \setminus \mathcal{G}_0$ of triples that we delete to produce the desired subhypergraph $\mathcal{G}_0 \subseteq \mathcal{G}$. Now, it is easy to bound the number $|\mathcal{G} \setminus \mathcal{G}_0|$ of triples that were deleted. Indeed, by Definition 5.0.3, condition (i) deletes fewer than

$$|V_0|N^2 + t\left\lfloor \frac{N}{t} \right\rfloor^2 N \leq tN^2 + \frac{N^3}{t} \leq T_0 N^2 + \frac{1}{t_0'} N^3 \leq \frac{2}{t_0'} N^3 \overset{(7.1)}{\leq} \frac{\mu}{6} N^3$$

triples. By the same definition, condition (ii) deletes fewer than

$$\gamma \binom{t}{2} \left\lfloor \frac{N}{t} \right\rfloor^2 N + \binom{t}{2} \gamma \left\lfloor \frac{N}{t} \right\rfloor^2 \leq \gamma N^3 \overset{(7.1)}{\leq} \frac{\mu}{6} N^3$$

more triples. Again using Definition 5.0.3, condition (iii) similarly deletes fewer than

$$\gamma \binom{t}{2} \left\lfloor \frac{N}{t} \right\rfloor^2 N \leq \frac{\gamma}{2} N^3 \overset{(7.1)}{\leq} \frac{\mu}{6} N^3$$

additional triples. Using the same definition a final time, condition (iv) deletes fewer than

$$\gamma N^2 N \overset{(7.1)}{\leq} \frac{\mu}{6} N^3$$

more triples. Now, trivially, condition (v) deletes

$$\sum_{P \in \text{Triad}(\mathbf{\Pi})} |\{\{x,y,z\} \in \mathcal{G}_P : \mathcal{G}_P = \mathcal{G} \cap \mathcal{K}_3(P) \text{ has } |\mathcal{G}_P| < \alpha_0 |\mathcal{K}_3(P)|\}|$$
$$< \alpha_0 \sum_{P \in \text{Triad}(\mathbf{\Pi})} |\mathcal{K}_3(P)| \leq \alpha_0 N^3 \stackrel{(7.1)}{\leq} \frac{\mu}{6} N^3$$

more triples. Finally, by Definition 5.0.4, condition (vi) deletes fewer than

$$\delta N^3 \stackrel{(7.1)}{\leq} \frac{\mu}{6} N^3$$

additional triples. Thus, conditions i)-vi) deleted a total of $|\mathcal{G} \setminus \mathcal{G}_0| \leq 6 \cdot (\mu/6) N^3 = \mu N^3$ triples from $\mathcal{G}$ so that the remaining subhypergraph $\mathcal{G} \setminus \mathcal{G}_0$ has the desired size $|\mathcal{G}_0| \geq |\mathcal{G}| - \mu N^3$.

It remains to show that $\mathcal{G}_0 \subset \mathcal{G}$, defined above, satisfies that, for every fixed $1 \leq i < j < k \leq t$, the triple (V_i, V_j, V_k) from $\Pi^{(1)}$ is $(2\omega_0)$-weakly regular with respect to $\mathcal{G}_0$. For that, we note the basic identity

$$|\mathcal{G}_0[V_i, V_j, V_k]| = \sum_{a=0}^{\ell_{ij}} \sum_{b=0}^{\ell_{jk}} \sum_{c=0}^{\ell_{ik}} \left| \mathcal{G}_0 \cap \mathcal{K}_3 \left(P_a^{ij} \dot\cup P_b^{jk} \dot\cup P_c^{ik} \right) \right|. \tag{7.9}$$

However, by conditions i)-vi) above, non-zero terms in the sum occur only when their indices (a,b,c) satisfy $0 \notin \{a,b,c\}$, and when P_{abc}^{ijk}, and when P_{abc}^{ijk} is $\left(\frac{1}{\ell}, \varepsilon\right)$-typical (because each of P_a^{ij}, P_b^{jk}, P_c^{ik} is $\left(\frac{1}{\ell}, \varepsilon\right)$-regular), and when

$$\left(\mathcal{G}_0 \left(P_{abc}^{ijk} \right) = \mathcal{G}_0 \cap \mathcal{K}_3 \left(P_{abc}^{ijk} \right), P_{abc}^{ijk}, V_i \dot\cup V_j \dot\cup V_k \right)$$

is an $\left(\alpha_{abc}^{ijk}, \vec{d}, \varepsilon \right)$-triad which is δ-link regular, where $\alpha_{abc}^{ijk} = d_{\mathcal{G}_0(P_{abc}^{ijk})} \left(P_{abc}^{ijk} \right) \geq \alpha_0$ and $\vec{d} = \left(\frac{1}{\ell}, \frac{1}{\ell}, \frac{1}{\ell} \right)$. For simplicity of terminology and notation, we refer to such indices $(a,b,c) \in [\ell_{ij}] \times [\ell_{jk}] \times [\ell_{ik}]$ as the $\mathcal{G}_0 [V_i, V_j, V_k]$-support and denote the set of them by

$$\text{supp}_{\mathcal{G}_0} (\{i,j,k\}) = \left\{ (a,b,c) \in [\ell_{ij}] \times [\ell_{jk}] \times [\ell_{ik}] : \mathcal{G}_0 \left(P_{abc}^{ijk} \right) \neq \emptyset \right\}.$$

Thus, (7.9) reduces to

$$|\mathcal{G}_0[V_i, V_j, V_k]| = \sum_{(a,b,c)\in \mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})} \left|\mathcal{G}_0\left(P^{ijk}_{abc}\right)\right| \tag{7.10}$$

Moreover, since $\mathcal{G}_0\left(P^{ijk}_{abc}\right) = \emptyset$ on every $(a,b,c) \notin \mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})$, the analogous identity

$$\left|\mathcal{G}_0\left[V_i', V_j', V_k'\right]\right| = \sum_{(a,b,c)\in \mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})} \left|\mathcal{G}_0\left(P^{ijk}_{abc}\right)[V_i, V_j, V_k]\right| \tag{7.11}$$

holds for all subsets $V_i' \subseteq V_i$, $V_j' \subseteq V_j$, $V_k' \subseteq V_k$. By the definition of $\mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})$, and by our choice of constants (see (7.1)-(7.4)) we may apply Lemma 6.0.1 to each term $(a,b,c) \in \mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})$ whenever $|V_i'| > \omega_0|V_i|$, $|V_j'| > \omega_0|V_j|$, and $|V_k'| > \omega_0|V_k|$. Doing precisely this to (7.10) (with $|V_i'| = |V_i|$, $|V_j'| = |V_j|$, and $|V_k'| = |V_k|$) yields

$$|\mathcal{G}_0[V_i, V_j, V_k]| = (1 \pm \omega_0)\frac{|V_i||V_j||V_k|}{\ell^3} \sum_{(a,b,c)\in \mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})} \alpha^{ijk}_{abc}, \tag{7.12}$$

while doing this to 7.11 with fixed subsets $V_i' \subseteq V_i$, $V_j' \subseteq V_j$, $V_k' \subseteq V_k$ with $|V_i'| > \omega_0|V_i|$, $|V_j'| > \omega_0|V_j|$, and $|V_k'| > \omega_0|V_k|$ yields

$$\left|\mathcal{G}_0\left[V_i', V_j', V_k'\right]\right| = (1 \pm \omega_0)\frac{|V_i'||V_j'||V_k'|}{\ell^3} \sum_{(a,b,c)\in \mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})} \alpha^{ijk}_{abc}. \tag{7.13}$$

We infer from (7.12) and (7.13) that

$$\ell^3 d_{\mathcal{G}_0}(V_i, V_j, V_k) = (1 \pm \omega_0) \sum_{(a,b,c)\in \mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})} \alpha^{ijk}_{abc},$$

and

$$\ell^3 d_{\mathcal{G}_0}(V_i', V_j', V_k') = (1 \pm \omega_0) \sum_{(a,b,c)\in \mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})} \alpha^{ijk}_{abc}.$$

Subtracting and using the triangle inequality gives

$$\ell^3 \left| d_{\mathcal{G}_0}(V_i', V_j', V_k') - d_{\mathcal{G}_0}(V_i, V_j, V_k) \right| \leq 2\omega_0 \sum_{(a,b,c) \in \mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})} \alpha_{abc}^{ijk}. \tag{7.14}$$

However, every $\alpha_{abc}^{ijk} \leq 1$ over all $\left|\mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})\right| \leq \ell_{ij}\ell_{jk}\ell_{ik} \leq \ell^3$ terms $(a,b,c) \in \mathrm{supp}_{\mathcal{G}_0}(\{i,j,k\})$, and so (7.14) implies

$$\left| d_{\mathcal{G}_0}(V_i', V_j', V_k') - d_{\mathcal{G}_0}(V_i, V_j, V_k) \right| \leq 2\omega_0$$

as desired. $\square$

Chapter 8

Conclusion and Future Work

The techniques in this dissertation prompt several interesting questions which we shall consider in the near future. Since it is not clear whether or not the following questions admit positive answers, we shall refrain from discussing potential applications of these problems. (Applications would ensue any confirmations below.)

The most natural question concerns whether or not Theorem 1.4.2 can be extended to k-uniform hypergraphs, for an arbitrary uniformity $k \geq 2$.

Question 8.0.1. *Does there exist an algorithm which, for an arbitrary integer $k \geq 2$, a real $\varepsilon > 0$, and a given n-vertex k-uniform hypergraph $\mathcal{H}^{(k)} = (V, E)$, constructs in time $O(n^k)$ an ε-regular partition $V = V_1 \cup \cdots \cup V_t$ for some integer $t = t(k, \varepsilon)$ depending on k and ε but not on n?*

Similarly to our approach with Theorem 1.4.3, we believe confirming Question 8.0.1 may involve proving an algorithm for a stronger notion of hypergraph regularity. This theme prompts our remaining questions.

Recall from the Introduction that we discussed the existence of *strong hypergraph regularity and counting lemmas*, which were established among the works [9, 11, 12, 23, 24, 25, 26]. In particular, Nagle, Rödl, and Schacht [25, 26] established an algorithmic strong hypergraph regularity lemma which, for a given n-vertex k-uniform hypergraph $\mathcal{H}^{(k)} = (V, E)$, constructs in time $O(n^{3k})$ a 'strongly regular' partition of $\mathcal{H}^{(k)}$. It is likely that their running time $O(n^{3k})$ can be improved, at least some. Perhaps it can even be reduced to an optimal order of magnitude $O(n^k)$.

Question 8.0.2. *Does there exist an algorithm which, for an arbitrary integer $k \geq 2$ and a given n-vertex k-uniform hypergraph $\mathcal{H}^{(k)} = (V, E)$, constructs in time $O(n^k)$ a 'strongly regular' partition of $\mathcal{H}^{(k)}$?*

The proofs in [25, 26] are lengthy and very technical, so a positive answer to Question 9.2 is far from clear. It would be of significant interest to determine the outcome of Question 8.0.2 even when $k = 3$.

Question 8.0.3. *Does there exist an algorithm which, for a given n-vertex 3-uniform hypergraph $\mathcal{H}^{(3)} = (V, E)$, constructs in time $O(n^3)$ a 'strongly regular' partition of $\mathcal{H}^{(3)}$?*

We mention that our Theorem 1.4.3 is a step in the direction of Question 8.0.3.

References

[1] Alon, N., Duke, R., Lefmann, H., Rödl, V., Yuster, R., "The algorithmic aspects of the Regularity Lemma", *Journal of Algorithms* Vol. 16, pp. 80–109, 1994

[2] Alon, N., Duke, R., Lefmann, H., Rödl, V., Yuster, R., "The algorithmic aspects of the Regularity Lemma (extended abstract)", *Proceedings of the IEEE Symposium on Foundations of Computer Science*, pp. 473–481, 1992

[3] Bollobás, B. (1998), *Modern Graph Theory*, Springer Science & Business Media, Graduate Texts in Mathematics, Volume 184, 394pp.

[4] Czygrinow, A., Rödl, V. (2000), *An algorithmic regularity lemma for hypergraphs*, SIAM J. Comput. **30**, 1041–1066.

[5] Diestel, R. (2010), *Graph Theory* (fourth edition), Springer-Verlag, Heidelberg, Graduate Texts in Mathematics, Volume 173, 451pp.

[6] Duke, R., Lefmann, H., Rödl, V. (1995), *A fast algoithm for computing the frequencies of subgraphs in a given graph*, SIAM J. Comput. **24**, 598–620.

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

[8] Fox, J. (2011), *A new proof of the graph removal lemma*, Annals of Mathematics **174**, 561–579.

[9] Frankl, P., Rödl, V. (2002), *Extremal problems on set systems*, Random Structures Algorithms **20**, no. 2, 131-164.

[10] Frankl, P., Rödl, V. (1992), *The uniformity lemma for hypergraphs*, Graphs and Combinatorics **8**, no. 4, 309–312.

[11] Gowers, W.T. (2006), *Quasirandomness, counting and regularity for 3-uniform hypergraphs*, Combinatorics, Probability and Computing **15**, 143–184.

[12] Gowers, W.T. (2007), *Hypergraph regularity and the multidimensional Szemerédi Theorem*, Annals of Mathematics **166**, 897–946.

[13] Guruswami, V., Håstad, J., Sudan, M. (2002), *Hardness of approximate hypergraph coloring*, SIAM J. Comput. **31**, no. 6, 1663–1686.

[14] Haxell, P., Nagle, B., Rödl, V. (2005), *An algorithmic version of the hypergraph regularity method (extended abstract)*, Proceedings of the IEEE Symposium on Foundations of Computer Science, 439–446.

[15] Haxell, P., Nagle, B., Rödl, V. (2008), *An algorithmic version of the hypergraph regularity method*, SIAM J. Comput. **37**, no. 6, 1728–1776.

[16] Janson, S., Luczak, T., Rucinski A. (2001), *Random Graphs*, John Wiley & Sons, 2001

[17] Kohayakawa, Y., Rödl, V., Thoma, L. (2002), *An optimal algorithm for checking regularity (extended abstract)*, The 13^{th} Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2002), San Francisco, CA, 6 to 8 January 2002, 277–286.

[18] Kohayakawa, Y., Rödl, V., Thoma, L. (2003), *An optimal algorithm for checking regularity*, SIAM J. Comput. **32**, 1210–1235.

[19] Komlós, J., Shokoufandeh, A., Simonovits, M., Szemerédi, E. (2002), *The regularity lemma and its applications in graph theory*, Theoretical aspects of computer science (Tehran, 2000), Lecture Notes in Comput. Sci. **2292**, Springer–Berlin, 84–112.

[20] Komlós, J., Simonovits, M. (1996), *Szemerédi's Regularity Lemma and its applications in graph theory*, Combinatorics, Paul Erdös is eighty **2**, 295-352.

[21] Lovász, L. (1973), *Coverings and colorings of hypergraphs*, Proceedings of the Fourth Southeastern Conference on Combinatorics, Graph Theory, and Computing (Florida Atlantic Univ., Boca Raton, Fla., 1973) (Winnipeg, Man.), Utilitas Math., 3–12.

[22] Molla, T., Nagle, B., Theado, J. (2019), *On some algorithmic hypergraph problems*, manuscript.

[23] Nagle, B., Rödl, V. (2003), *Regularity properties for triple systems*, Random Structures Algorithms **23**, no. 3, 264–332.

[24] Nagle, B., Rödl, V., Schacht, M. (2006), *The counting lemma for regular k-uniform hypergraphs*, Random Structures Algorithms **26**, no. 2, 1–67.

[25] Nagle, B., Rödl, V., Schacht, M. (2016), *An algorithmic hypergraph regularity lemma [Extended Abstract]*, In: Robert Krauthgamer (editor): Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2016), 1765–1773. ACM Press.

[26] Nagle, B., Rödl, V., Schacht, M. (2018), *An algorithmic hypergraph regularity lemma*, Random Structures Algorithms **52**, no. 2, 301–353.

[27] Nešetřil, J., Poljak, S. (1985), *On the complexity of the subgraph problem*, Comment. Math. Univ. Carolin. **26**, no. 2, 415–419.

[28] Person, Y., Schacht, M. (2011), *An expected polynomial time algorithm for coloring 2-colorable 3-graphs*, Discrete Mathematics and Theoretical Computer Science **13**, no. 2, 1–18.

[29] Rödl, V., Skokan, J. (2004), *Regularity lemma for k-uniform hypergraphs*, Random Structures Algorithms **25**, 1–42.

[30] Roth, K.F. (1953), *On certain sets of integers*, J. London Math. Soci. **28**, 104–109.

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

[32] Szemerédi, E. (1975), *On sets of integers containing no k elements in arithmetic progression*, Acta Arithmetica **27**, 199–245.

[33] Szemerédi, E. (1978), *Regular partitions of graphs*, in: Problèmes en Combinatoire et Théorie des Graphes, Proc. Colloque Inter. CNRS, (Bermond, J.-C., Fournier, J.-C., Las Vergnas, M., Sotteau, D., eds), 399–401.

[34] Tao, T., Vu, V. (2006), *Additive Combinatorics* (Vol. 105), Cambridge University Press.

[35] Williams, V.V. (2012), *Multiplying matrices faster than Coppersmith-Winograd*, Proceedings of the 44th Symposium on Theory of Computing, 887–898.

About the Author

John Theado was born in Sarasota, FL and grew up in Western New York. A self-taught computer programmer, he went back to school as an older student to begin a degree in computer science at South Florida Community College. After enjoying the required math courses, he decided to declare as a mathematics major after transfering to the University of South Florida in 2010. He earned his bachelor's degree in 2012 and then entered the mathematics graduate program at USF in 2013. He earned his Ph.D. in Mathematics in 2019.