The Minimum Size of 3-Graphs without a 4-Set Spanning No or Exactly Three Edges

Oleg Pikhurko*
Department of Mathematical Sciences
Carnegie Mellon University
Pittsburgh, PA 15213
Web: http://www.math.cmu.edu/~pikhurko

January 12, 2011

Abstract

Let G_i be the (unique) 3-graph with 4 vertices and i edges. Razborov [*On 3-Hypergraphs with Forbidden 4-Vertex Configurations*, SIAM J. Discr. Math. **24** (2010), 946–963] determined asymptotically the minimum size of a 3-graph on n vertices having neither G_0 nor G_3 as an induced subgraph. Here we obtain the corresponding stability result, determine the extremal function exactly, and describe all extremal hypergraphs for $n \geq n_0$. It follows that any sequence of almost extremal hypergraphs converges, which answers in the affirmative a question posed by Razborov.

1 Introduction

For a set X and an integer k, let $\binom{X}{k} = \{Y \subseteq X : |Y| = k\}$. A *$k$-graph G with vertex set V* is a subset of $\binom{V}{k}$, i.e., it is a collection of k-element subsets of V. Elements of V and G are called *vertices* and *edges* respectively. We will also call G a *hypergraph*.

Let $\mathcal{G}$ be a family of k-graphs. A k-graph F is *$\mathcal{G}$-free* if it contains no member of $\mathcal{G}$ as an induced subgraph. Let $t(n, \mathcal{G})$ be the minimum size of a $\mathcal{G}$-free k-graph on n vertices. This function is related to the Turán problem; we refer the reader to surveys by Füredi [Für91], Sidorenko [Sid95], and Keevash [Kee11].

*Partially supported by the National Science Foundation, Grant DMS-0758057.

If $\mathcal{G} = \{G\}$ consist of one k-graph G, we may abbreviate $t(n, \{G\})$ to $t(n, G)$, etc. For $0 \le i \le 4$, let G_i be the (unique) 3-graph with 4 vertices and i edges.

One of the most famous open questions in extremal combinatorics is to determine $t(n, G_0)$. It goes back to the fundamental paper by Turán [Tur41] who conjectured that

$$t(n, G_0) = t_n, \tag{1}$$

where t_n is defined as follows.

For pairwise disjoint sets V_0, V_1, and V_2, the *Turán pattern* T_{V_0,V_1,V_2} is the 3-graph on $V = V_0 \cup V_1 \cup V_2$ whose edges are triples $\{x, y, z\}$ with $x, y \in V_i$ and $z \in V_i \cup V_{i+1}$ for some $i \in \mathbb{Z}_3$. (Here $\mathbb{Z}_m$ denotes the additive group of residues modulo m.) Let t_{v_0,v_1,v_2} be the number of edges in T_{V_0,V_1,V_2} where $|V_i| = v_i$. The *Turán 3-graph* T_n is the (unique up to isomorphism) Turán pattern T_{V_0,V_1,V_2} with $v_0 + v_1 + v_2 = n$ and $|v_i - v_j| \le 1$ for all $i, j \in \mathbb{Z}_3$. It is not hard to show (see Lemma 4) that among all Turán patterns on n vertices, the Turán 3-graph T_n has the smallest size. Let

$$t_n = |T_n|.$$

We have $t_n = (\frac{4}{9} + o(1))\binom{n}{3}$ as $n \to \infty$. Also, any Turán pattern is G_0-free; thus $t(n, G_0) \le t_n$. The problem of obtaining a matching lower bound (even within a $(1+o(1))$-factor) seems to be extremely difficult. Successively better lower bounds on $t(n, G_0)$ were proved by de Caen [dC94], Giraud (unpublished, see [CL99]), and Chung and Lu [CL99]. Razborov [Raz07, Raz10a] presented a general framework for working with extremal problems of this kind. His solution of a certain semidefinite program with over 900 variables suggests that $t(n, G_0) \ge 0.43833 \binom{n}{3}$ for all sufficiently large n, see also Baber and Talbot [BT10]. One of many difficulties here is that, if Turán's conjecture (1) is correct, then there are many non-isomorphic extremal 3-graphs, see Brown [Bro83], Kostochka [Kos82], and Fon-Der-Flaass [FDF88]. Also, we refer the reader to Razborov [Raz10b] for some related results.

Note that T_n is also G_3-free; thus $t(n, \{G_0, G_3\}) \le t_n$. Applying his technique Razborov [Raz10a] proved the matching asymptotic lower bound. Thus

$$t(n, \{G_0, G_3\}) = \left(\frac{4}{9} + o(1)\right)\binom{n}{3}. \tag{2}$$

This result is interesting because there are very few non-trivial hypergraphs or hypergraph families for which the asymptotic of its Turán function is known. Also, it gives us a better understanding of the original conjecture of Turán. For example, if the conjecture is false, then any G_0-free 3-graph G on n vertices beating t_n has to contain an induced copy of G_3. (In fact, if $|G| \le (1 - \Omega(1))\, t_n$ as $n \to \infty$, then G contains $\Omega(n^4)$ G_3-subgraphs by the super-saturation technique of Erdős and Simonovits [ES83]).

Here, we prove for all $n \ge n_0$ that $t(n, \{G_0, G_3\}) = t_n$ and the Turán hypergraph T_n is the unique extremal 3-graph:

Theorem 1 (Exact Result) *There is* n_0 *such that every* $\{G_0, G_3\}$*-free* 3*-graph* F *on* $n \geq n_0$ *vertices has at least* t_n *edges with equality if and only if* $F \cong T_n$*.*

In particular, $t(n, \{G_0, G_3\}) = t_n$ *for* $n \geq n_0$*.*

Theorem 1 is also interesting in the context of the rapidly developing theory of graph and hypergraph limits, see e.g. [LS06, BCL$^+$08, ES08]. Although Razborov's proof of (2) is stated without any appeal to hypergraph limits, the flag algebras introduced by him provide a convenient and powerful language for manipulating limit objects. Also, any relations proved with the help of flag algebras or (hyper)graph limits hold only asymptotically as the order of the underlying (hyper)graph tends to infinity. So, at the first sight, this technique can give asymptotic results only. However, the proof of Theorem 1 gives an example of how a solution of the "limiting" case may lead to an exact result for all sufficiently large n. The key ingredient here is the stability property which states, roughly speaking, that all almost extremal hypergraphs have essentially the same unique structure. Here is the precise formulation for the $\{G_0, G_3\}$-problem:

Theorem 2 (Stability Property) *For every* $\varepsilon > 0$ *there is* $c > 0$ *such that the following holds. Let* G *be a* $\{G_0, G_3\}$*-free* 3*-graph on* $n > 1/c$ *vertices with at most* $t_n + cn^3$ *edges. Then we can make* G *isomorphic to* T_n *by changing at most* εn^3 *triples.*

Stability greatly helps in proving exact results (with one example being Theorem 1). This approach was pioneered by Simonovits [Sim68] in the late 1960s and has led to exact solutions of numerous extremal problems since then. In recent years it has been actively used to prove exact results for the hypergraph Turán problem, see e.g. [KM04, FS05, KS05a, KS05b, MP07, FMP08, Pik08].

As an extra bonus, Theorem 2 also implies the following result, which answers in the affirmative a question posed by Razborov [Raz10a, Section 5]. For $F \subseteq \binom{V}{k}$ and $H \subseteq \binom{U}{k}$ let $\mathrm{ind}(H, F)$ denote the *induced density* of H in F, that is, the probability that a random injection $U \to V$ preserves all edges and non-edges of H.

Theorem 3 (Convergence) *Let* $n_i \to \infty$*. Let* $F_i \subseteq \binom{[n_i]}{3}$ *be a* $\{G_0, G_3\}$*-free 3-graph with* $|F_i| = (\frac{4}{9} + o(1)) \binom{n_i}{3}$ *as* $i \to \infty$*. Then, for every fixed* 3*-graph* H*, the limit* $\lim_{i\to\infty} \mathrm{ind}(H, F_i)$ *exists (and is equal to* $\lim_{m\to\infty} \mathrm{ind}(H, T_m)$*).*

Proof. By Theorem 2 we can change $o(n_i^3)$ edges in F_i and transform it into T_{n_i}. Relabel the vertices of F_i so that $V(F_i) = V(T_{n_i})$ and the symmetric difference $F_i \triangle T_{n_i}$ has $o(n^3)$ triples, where $V(F)$ denotes the vertex set of a hypergraph F.

For every fixed 3-graph H we have $|\mathrm{ind}(H, F_i) - \mathrm{ind}(H, T_{n_i})| = o(1)$ because the probability that a random injection $V(H) \to V(F_i)$ hits one of the triples where F_i

and T_{n_i} differ is $o(1)$. Also, $\mathrm{ind}(H, T_m)$ tends to an (explicitly computable) limit λ_H as $m \to \infty$. Thus $\mathrm{ind}(H, F_i) \to \lambda_H$, as required. ▮

Remark. A simple application of the Principle of Inclusion-Exclusion shows that the conclusion of Theorem 3 is equivalent to the statement that the sequence (F_i) of 3-graphs converges, as defined by Elek and Szegedy [ES08, Definition 2.5].

2 Some Notation

We denote $[n] = \{1, \ldots, n\}$. For brevity, we often omit punctuation signs when writing sets; for example, abc is a shorthand for $\{a, b, c\}$.

Let $G \subseteq \binom{V}{k}$ be a k-graph on V. For $A \subseteq V$, $G[A] = \{D \in G : D \subseteq A\}$ denotes the subgraph of G induced by A. For disjoint subsets $V_1, \ldots, V_k \subseteq V$, let

$$G[V_1, \ldots, V_k] = \{D \in G : \forall\, i \in [k]\ |D \cap V_k| = 1\}$$

denote the k-partite subgraph of G induced by the sets V_i. For $A \subseteq V$ with $a \leq k-1$ elements, the *link* $(k-a)$*-graph* of A is

$$G_A = \{D : D \subseteq V \setminus A,\ D \cup A \in G\}.$$

When $a = k-1$, we view G_A as a set of vertices rather than a set of 1-element sets. The *maximum degree* of G is $\Delta(G) = \max\{|G_x| : x \in V\}$.

Let G and H be two k-graphs with the same number of vertices. They are *isomorphic* (written as $G \cong H$) if there is a bijection $f : V(G) \to V(H)$ such $A \in G$ if and only if $f(A) \in H$ for every $A \in \binom{V(G)}{k}$. The *edit distance* $\delta_1(G, H)$ is the minimum of $|\sigma(G) \triangle H|$ over all bijections $\sigma : V(G) \to V(H)$. In other words, $\delta_1(G, H)$ is the smallest number of k-tuples whose inclusion into G one has to change in order to make G isomorphic to H.

3 Auxiliary Results

Here we list a few lemmas needed later. Their proofs are fairly straightforward and are included here for the sake of completeness.

Lemma 4 *Let $n \geq 3$. For every Turán pattern $T_{X,Y,Z}$ on $[n]$ we have $|T_{X,Y,Z}| \geq t_n$ and, if we have equality, then $T_{X,Y,Z} \cong T_n$.*

Proof. Let x, y, z be the cardinalities of X, Y, Z respectively. The claim is trivial for $n = 3$, so let us assume that $n \geq 4$.

It is enough to show that no two of x, y, z differ by more than by 1. Suppose on the contrary that this is false. We will give an example of a triple strictly better than (x, y, z), thus proving the lemma. Up to a symmetry, there are two cases.

Case 1 $x \geq y \geq z$ and $x \geq z + 2$.

Routine simplifications show that

$$\partial := t_{x,y,z} - t_{x-1,y,z+1} = \frac{x^2}{2} + xy - xz - \frac{y^2}{2} - \frac{3x}{2} - \frac{y}{2} + z + 1.$$

It is enough to show that this expression is strictly positive. This a linear function of z with the coefficient $1 - x < 0$, so it suffices to show that $\partial > 0$ under the additional assumption that $z = \min(x - 2, y)$.

If $z = x - 2$, then y can be one of x, $x - 1$, and $x - 2$ and ∂ is $x - 1$, $x - 1$, and $x - 2$ respectively. Since $n \geq 4$, we have $x \geq 2$. Also, if $x = 2$, then $z = 0$, $n = 4$, and $y = 2$. In all cases, ∂ is strictly positive, as desired.

If $z = y$, then $\partial = \frac{x^2}{2} - \frac{3x}{2} - \frac{y^2}{2} + \frac{y}{2} + 1$, which is an increasing function of $x \geq 2$. So it follows from the case $x = z + 2$ which we have just done.

Case 2 $x \geq z \geq y$ and $x \geq y + 2$.

Routine simplifications show that

$$\partial := t_{x,y,z} - t_{x-1,y+1,z} = -\frac{y^2}{2} + xy - yz + \frac{z^2}{2} - \frac{y}{2} - \frac{z}{2}.$$

This is a non-decreasing function of x, so it is enough to consider the case $x = \max(y + 2, z)$. If $y = x - 2$, then z is one of $x, x - 1, x - 2$ with ∂ being $x - 1$, $x - 2$ and $x - 2$ respectively. The assumption $n \geq 4$ implies that $\partial > 0$ in each case. If $z = x$, then $\partial = \frac{x^2}{2} - \frac{x}{2} - \frac{y^2}{2} - \frac{y}{2}$, which is increasing in $x \geq 2$, so it enough to assume that $x = z = y + 2$; we have $\partial = x - 1 > 0$ in this case. ∎

Lemma 5 *For every $\varepsilon > 0$ there is $c > 0$ such that for every $n > 1/c$ and for every non-negative integers v_0, v_1, v_2 with $v_0 + v_1 + v_2 = n$ and $t_{v_0,v_1,v_2} \leq (\frac{4}{9} + c)\binom{n}{3}$ we have $|v_i - n/3| \leq \varepsilon n$ for every $i \in \mathbb{Z}_3$.*

Proof. Since we are not interested in an explicit dependence of c on ε, we present a "non-constructive" but short proof. Suppose that the lemma is false, that is, there is $\varepsilon > 0$ such that for every integer m we have a counterexample (v_0, v_1, v_2) for $c = 1/m$. By choosing a subsequence of m, we can assume that v_i/n converges for each $i \in \mathbb{Z}_3$; let x_i be the limit of v_i/n. By Lemma 4, we have $t_{v_0,v_1,v_2} = (\frac{4}{9} + c)\binom{n}{3}$. Thus

$$P(x_0, x_1, x_2) = \frac{x_0^3 + x_1^3 + x_2^3}{6} + \frac{x_0^2 x_1 + x_1^2 x_2 + x_2^2 x_0}{2}$$

assumes value $\frac{4}{9} \times \frac{1}{3!} = \frac{2}{27}$.

Let us minimize $P(x, y, z)$ over non-negative reals x, y, z with $x + y + z = 1$. If, for example, $x \ge y \ge z$ with $x > z$, then the following difference of partial derivatives

$$\frac{\partial}{\partial z} P(x, y, z) - \frac{\partial}{\partial x} P(x, y, z) = (y - x)\,\frac{x + y}{2} + (z - y)x$$

is strictly negative (because at least one of $y - x \le 0$ and $z - y \le 0$ is strictly negative while $x \ge \frac{1}{3}$). Thus $P(x - \delta, y, z + \delta) < P(x, y, z)$ for all small $\delta > 0$. Likewise, if $x \ge z \ge y$ with $x > y$, then

$$\frac{\partial}{\partial y} P(x, y, z) - \frac{\partial}{\partial x} P(x, y, z) = (z - x)y + (y - z)\,\frac{y + z}{2} \le 0.$$

Moreover, if we have equality here, then $y = z = 0$, $x = 1$ and P assumes value $\frac{1}{6} > P(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) = \frac{2}{27}$. In any case, $P(x, y, z)$ is not minimum. This implies that the only extremal point is $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ and the minimum value of P is $\frac{2}{27}$.

It follows that $x_0 = x_1 = x_2 = \frac{1}{3}$, which contradicts the fact some two of the ratios v_0/n, v_1/n, and v_2/n differ by at least ε for every m. ∎

4 Stability for the $\{G_0, G_3\}$-Problem

In this section we will prove Theorem 2. Suppose on the contrary that it is false. Thus there is $\varepsilon > 0$ and a sequence (F_i) with $|F_i| \le (\frac{4}{9} + o(1))\binom{n_i}{3}$ as $i \to \infty$, where F_i is a $\{G_0, G_3\}$-free 3-graph on $n_i > i$ vertices that is εn_i^2-far in the edit distance from T_{n_i}.

Fix any such sequence (F_i). We will split the whole proof resulting in a final contradiction into a sequence of claims.

Let us call a 3-graph H *singular* if H is $\{G_0, G_3\}$-free but for every n the Turán graph T_n does not contain H as an induced subgraph. Clearly, it is enough to check this inclusion for $n = 3\,|V(H)|$ only. There are exactly 26 non-isomorphic singular 3-graphs on 6 vertices, denoted by $H_9, \ldots, H_{34}$ in [Raz10a].

Claim 1 *For every singular* 3*-graph* H *on* 6 *vertices we have* $\mathrm{ind}(H, F_i) \to 0$ *as* $i \to \infty$.

Proof of Claim. Although this claim is stated in [Raz10a, Section 5], let us sketch its proof very briefly. Let $n \to \infty$ and let F be an arbitrary $\{G_0, G_3\}$-free 3-graph on n vertices. Let $\rho = |F|/\binom{n}{3}$ be the edge density of F. Razborov [Raz10a, Section 3] derives the following identity:

$$\frac{5}{9}(\rho - 4/9) = [\![(e - 4/9)^2]\!]_1 + [\![Q_1(f_1, \ldots, f_4)]\!]_{\tau_1} + [\![Q_2(g_0, \ldots, g_5)]\!]_{\tau_2} + R + o(1). \quad (3)$$

Rather than formally defining all terms appearing here, we state only those properties that we need in order to prove Claim 1, referring the reader to [Raz10a] for all details.

- Each term involving the brackets $[\![\dots]\!]$ is non-negative by Inequality (6) in [Raz10a].
- The last term, which we denoted by R, is of the form $\sum_H \alpha_H \mathrm{ind}(H,F)$, where the following applies.
  - The sum runs over $\{G_0, G_3\}$-free 3-graphs H with 6 vertices.
  - α_H are explicit non-negative reals that are listed in Table 3 in [Raz10a].
  - $\alpha_H \geq 1/360$ for every singular H.

It follows that if $|F| \leq (\frac{4}{9}+\varepsilon)\binom{n}{3}$, then $\mathrm{ind}(H,F) \leq 360 \times \frac{5\varepsilon}{9} + o(1)$ for every singular 6-vertex hypergraph H. The claim follows. ▮

Now, we can apply the Strong Hypergraph Removal Lemma of Rödl and Schacht [RS09] to each 3-graph F_i with respect to induced singular subgraphs $H_9, \dots, H_{34}$. The lemma shows that we can change $o(n_i^3)$ edges in F_i as $i \to \infty$ and ensure that it contains no induced singular subgraph on 6 vertices. Hence, by making ε slightly smaller, it is enough to derive a contradiction under the additional assumption that F_i has no induced singular subgraph on 6 vertices.

Fix large i, and let $n = n_i$, $F = F_i$, and $V = V(F)$ for the remainder of this section. Let T be the logical predicate that takes three disjoint sets $U_0, U_1, U_2 \subseteq V$ as input and is true if and only if the induced subgraph $F[U_0 \cup U_1 \cup U_2]$ follows the Turán pattern, that is, its edges are precisely triples xyz with $xy \in U_j$ and $z \in U_j \cup U_{j+1}$ for some $j \in \mathbb{Z}_3$. Thus we have the following claim.

Claim 2 *For any set $U \subseteq V$ with $|U| \leq 6$, there is a partition $U = U_0 \cup U_1 \cup U_2$ such that $T(U_0, U_1, U_2)$ holds.* ▮

Let the logical predicate $S(ab, cd)$ state that the vertices $a, b, c, d \in V$ are pairwise distinct, $abc, abd \in F$, and $acd, bcd \notin F$. Also, for $a, b \in V$ let us write $a \sim b$ if $a = b$ or there are $c, d \in V$ satisfying $S(ab, cd)$. In the latter case, we call the pair cd a *witness* of $a \sim b$. Clearly, the binary relation $\sim$ is symmetric. The following claim can be checked by a trivial case analysis.

Claim 3 *If $T(U_0, U_1, U_2)$ holds and $a, b, c, d \in U_0 \cup U_1 \cup U_2$ satisfy $S(ab, cd)$, then for some $j \in \mathbb{Z}_3$ we have $a, b \in U_j$ and $c, d \in U_{j+1}$.* ▮

Claim 4 *The relation $\sim$ is transitive.*

Proof of Claim. Suppose that $a \sim b$ and $b \sim f$, which is witnessed by $S(ab, cd)$ and $S(bf, gh)$ respectively. Let $U = \{a, b, c, d, f, g, h\}$.

If $|U| \le 6$, then take a partition $U = U_0 \cup U_1 \cup U_2$ given by Claim 2. By Claim 3 and symmetry, we can assume that e.g. $ab \in U_0$ and $cd \in U_1$. Since $S(bf, gh)$ holds, the vertices b and f are in the same part U_j by Claim 3. Thus $f \in U_0$. Then $S(af, cd)$ holds, giving $f \sim a$ as required.

So suppose that $|U| = 7$ (i.e. all involved vertices are pairwise distinct). Consider $U' = U \setminus \{g\}$. It has 6 elements, so by Claim 2 there is a partition $U' = U_0' \cup U_1' \cup U_2'$ satisfying $T(U_0', U_1', U_2')$. Assume by Claim 3 that $a, b \in U_0'$ and $c, d \in U_1'$.

If $f \in U_0'$, then $S(af, cd)$ holds and thus $a \sim f$, as required. So suppose that $f \notin U_0'$.

If f or h is in U_1', say $h \in U_1'$, then we have $S(ab, dh)$. We can replace c by h in U, reducing its size to 6 and conclude as above that $a \sim f$. So suppose that $f, h \notin U_1'$.

Since $bfh \in F$, it must be the case that $f, h \in U_2'$. Thus $T(ab, cd, fh)$ holds, in particular, $afh \in F$. By the symmetry, we can swap f with a and cd with gh in the above analysis and conclude that either we are done or that $T(fb, gh, ac)$ holds. But the latter relation implies that $afh \notin F$. This contradiction proves Claim 4. ▮

Thus all vertices of F are partitioned into $\sim$-equivalence classes, say

$$V = V_1 \cup \cdots \cup V_k.$$

Let us call two vertices $a, b \in V$ *twins* if swapping a and b we get an automorphism of F.

Claim 5 *If $a \sim b$, then a and b are twins.*

Proof of Claim. Pick c and d such that $S(ab, cd)$ holds. Let $e, f \in V \setminus \{a, b\}$ be arbitrary. Let $U = \{a, b, c, d, e, f\}$. Apply Claim 3 to U to obtain a Turán partition $U = U_0 \cup U_1 \cup U_2$. Without loss of generality assume that $a, b \in U_0$. Wherever the vertices e, f fall, we have $aef \in F$ if and only if $bef \in F$ by the symmetry of U_0 with respect to $F[U]$, as required. ▮

Hence, we can define the *skeleton* 3-graph F' on $[k]$: namely, a triple $fgh \in \binom{[k]}{3}$ is an edge of F' if and only if the induced 3-partite 3-graph $F[V_f, V_g, V_h]$ is complete (or, equivalently by Claim 5 non-empty).

Claim 6 *F' does not contain an induced copy of G_0, G_2, nor G_3.*

Proof of Claim. If some 4-set $uwxy \in \binom{[k]}{4}$ spans exactly two triples in F', say $uwx, uwy \in F'$, then $S(ab, cd)$ holds for arbitrary representatives $a \in V_u$, $b \in V_w$, $c \in V_x$, and $d \in V_y$. Thus $a \sim b$, a contradiction to $u \neq w$.

The $\{G_0, G_3\}$-freeness of F' follows from the $\{G_0, G_3\}$-freeness of F and Claim 5. ▮

Claim 7 *Each V_j spans a complete 3-graph.*

Proof of Claim. Let $a, b, c \in V_j$ be distinct. Choose a witness fg to $a \sim b$. By Claim 5, the pair fg also witnesses $b \sim c$ and $a \sim c$. Thus $abf, acf, bcf \in F$. Since we do not have G_3, $abc \in F$. ∎

Claim 8 *If $fgh \in F'$, then $V_f \cup V_g \cup V_h$ spans the complete 3-graph in F.*

Proof of Claim. Take $c \in V_f$, $d \in V_g$ and distinct $a, b \in V_h$. We have $acd, bcd \in F$. Since $c \not\sim d$, at least one of abc or abd is an edge of F. But F does not contain G_3. Hence both abc and abd are edges of F. The result follows from Claim 7. ∎

Claim 9 *If two edges $D, E \in F'$ intersect in two vertices, then $D \cup E$ induces a complete subgraph in F'.*

Proof of Claim. This follows from Claim 6 (the $\{G_2, G_3\}$-freeness of F'). ∎

Claim 10 *If two edges $D, E \in F'$ intersect in one vertex, then $D \cup E$ induces a complete subgraph in F'.*

Proof of Claim. Let $D = abc$ and $E = cde$. The 4-set $abde$ spans at least one edge (since F' is G_0-free), say $abd \in F'$. By Claim 9 applied to $abc, abd \in F'$, the quadruple $abcd$ induces G_4. Since $cde \in F'$ intersects each of $acd, bcd \in F'$ in two vertices, we have $G[acde] \cong G[bcde] \cong G_4$ by Claim 9. This implies that every triple of $abcde$ is in F' except perhaps abe. But Claim 9 applied to $abc, bce \in F'$ shows that $abe \in F'$, as required. ∎

By the above claims, F' is a vertex-disjoint union of complete subgraphs on sets $W_1, \dots, W_l$ respectively. Let us agree that each isolated vertex of F', if there are any, forms a separate part W_j. The sets $W_1, \dots, W_l$ partition $[k] = V(F')$. Since F' is G_0-free, we have $l \le 3$ (otherwise pick one vertex from some four parts W_j to obtain a G_0-subgraph in F'). Moreover, every triple of F intersects at most two of the parts W_j. For $j \in [l]$, define $U_j = \cup_{h \in W_j} V_h$ and $u_j = |U_j|$. The sets $U_1, \dots, U_l$ partition $V = V(F)$.

Claim 11 *For each $j \in [l]$, the set $U_j \subseteq V$ spans a complete subgraph in F.*

Proof of Claim. If $W_j = \{h\}$ has only one element, then $U_j = V_h$ and the result follows from Claim 7. If $|W_j| \ge 3$, then W_j spans a (non-trivial) complete subgraph in F' and the result follows from Claim 8. Finally, it is impossible to have $|W_j| = 2$ for otherwise the two vertices of W_j would be isolated in F' and would form a separate part W_h each. ∎

Suppose first that $l = 3$. Then the following holds.

Claim 12 *For every $h \in [3]$ there is $j \neq h$ such that $abc \in F$ for every $ab \in \binom{U_h}{2}$ and $c \in U_j$.*

Proof of Claim. Without loss of generality assume $h = 3$. For $j = 1, 2$, let H_j be the 2-graph that consists of all pairs $ab \in \binom{U_3}{2}$ such that $abc \in F$ for every $c \in U_j$.

Let us show that the union of H_1 and H_2 is $\binom{U_3}{2}$. If, on the contrary, some pair $ab \in \binom{U_3}{2}$ is not in $H_1 \cup H_2$, then pick $c_j \in U_j$ with $abc_j \notin F$ for $j = 1, 2$ and observe that abc_1c_2 spans G_0 in F, a contradiction.

Also, for $j = 1, 2$, the 2-graph H_j contains no induced path of length 2. Indeed, if $ad, bd \in H_j$ but $ab \notin H_j$, then pick $c \in U_j$ with $abc \notin F$ and observe that $abcd$ spans G_3 (note that $abd \in F$ by Claim 11), a contradiction.

Hence, each of H_1 and H_2 is a union of vertex-disjoint cliques. Since $H_1 \cup H_2 = \binom{U_3}{2}$, it easily follows that H_1 or H_2 is equal to $\binom{U_3}{2}$, proving the claim. ▌

Claim 13 *For every distinct $j, h \in [3]$, F contains at least*

$$\frac{u_j^2 u_h + u_j u_h^2}{2} - \frac{u_j u_h \sqrt{u_j^2 + u_h^2}}{2} + O(n^2) \tag{4}$$

triples within $U_j \cup U_h$ that intersect both U_j and U_h.

Proof of Claim. Suppose without loss of generality that $j = 1$ and $h = 2$.

Let $i = 1$ or 2. Let m_i be the number of triples that do not belong to F and have exactly two vertices in U_i and one vertex in U_{3-i}. For $ab \in \binom{U_i}{2}$, let m_{ab} be the number of $c \in U_{3-i}$ such that $abc \notin F$. The sum

$$s_i = \sum_{ab \in \binom{U_i}{2}} \binom{m_{ab}}{2}$$

counts 4-tuples $abcd$ such that $ab \in \binom{U_i}{2}$, $cd \in \binom{U_{3-i}}{2}$, and $abc, abd \notin F$. Since F is G_0-free, no 4-tuple is counted twice (i.e. for both $i = 1$ and $i = 2$). Thus $s_1 + s_2 \leq \binom{u_1}{2}\binom{u_2}{2}$. On the other hand, the convexity of the function $\binom{x}{2}$ and the identity $m_i = \sum_{ab \in \binom{U_i}{2}} m_{ab}$ imply that

$$s_i \geq \binom{u_i}{2} \binom{m_i / \binom{u_i}{2}}{2} = \frac{m_i^2}{u_i^2} + O(n^3)$$

We conclude that

$$\frac{u_1^2 u_2^2}{4} \geq \frac{m_1^2}{u_1^2} + \frac{m_2^2}{u_2^2} + O(n^3). \tag{5}$$

If we ignore the error term and maximize $m = m_1 + m_2$ over non-negative reals m_1, m_2 satisfying (5), then any optimal assignment makes (5) equality. Using this to eliminate

m_2, we obtain that $m = m_1 + u_2(u_1^2u_2^2/4 - m_1^2/u_1^2)^{1/2}$. This function of m_1 has the unique maximum $\frac{1}{2}\, u_1u_2(u_1^2+u_2^2)^{1/2}$ attained at the unique positive root $m_1 = u_1^3u_2(u_1+u_2)^{1/2}/2$ of its derivative. This gives an upper bound on the number of triples between U_2 and U_3 that are missing from F, proving the claim. ∎

By Claim 12, fix $j(h)$ for each $h \in [3]$. Up to a symmetry we have two cases.

Case 1 $(j(1), j(2), j(3)) = (2, 1, 1)$.

Here $U_1 \cup U_2$ spans a complete subgraph in F. By Claim 13, the number of edges in F is at least $P(u_1, u_2, u_3) + O(n^2)$, where

$$P(u_1, u_2, u_3) = \frac{(u_1+u_2)^3 + u_3^3}{6} + \frac{u_1u_3^2}{2} + \frac{u_2u_3^2 + u_2^2u_3}{2} - \frac{u_2u_3\sqrt{u_2^2+u_3^2}}{2}.$$

Claim 14 *The minimum value of $P(x, y, z)$ over non-negative x, y, z with $x + y + z = 1$ is strictly larger than $\frac{4}{9} \times \frac{1}{6} = \frac{2}{27}$.*

Proof of Claim. Let

$$Q(y, z) = P(1-y-z, y, z) - \frac{2}{27} = \frac{y^2z}{2} - \frac{z^3}{2} - \frac{yz\sqrt{y^2+z^2}}{2} + z^2 - \frac{z}{2} + \frac{5}{54}.$$

Let us minimize Q over

$$S = \{(y, z) \in \mathbb{R}^2 : y \geq 0,\ z \geq 0,\ y + z \leq 1\}.$$

The derivative

$$\frac{\partial Q(y,z)}{\partial y} = -\frac{z\left(y - \sqrt{y^2+z^2}\right)^2}{2\sqrt{y^2+z^2}},$$

is non-positive. Hence, there is an optimal assignment with $y = 1 - z$. Note that

$$Q(1-z, z) = \frac{5}{54} + \frac{(z^2-z)\sqrt{1-2z+2z^2}}{2}.$$

Let $I = [0, 1] \subseteq \mathbb{R}$ denote the closed unit interval. It is enough to show that $(5/54)^2 - (Q(1-z,z) - 5/54)^2$ is positive on I. The last expression is a polynomial and factorizes as $(18z^2 - 18z + 5)R(z)/2916$, where $R(z) = -81z^4 + 162z^3 - 99z^2 + 18z + 5$. Clearly, it remains to show that $R(z)$ is positive on I. The derivative $R'(z)$ has three simple roots $1/2$ and $(3 \pm \sqrt{5})/6$, all of which are in I. So the potential minima of f on I are restricted to values $f(0)$, $f(1/2)$, or $f(1)$. But each of these is positive. This proves Claim 14. ∎

By Claim 14, we have that $|F|$ is strictly larger than $\left(\frac{2}{27} + o(1)\right) n^3 = \left(\frac{4}{9} + o(1)\right)\binom{n}{3}$, a contradiction.

Case 2 $(j(1), j(2), j(3)) = (2, 3, 1)$.

Thus F contains the Turán pattern $T = T_{U_1,U_2,U_3}$ plus perhaps some extra edges. By Lemma 5, the 3-graph T alone has at least $(\frac{4}{9} + o(1))\binom{n}{3}$ edges. Since we have assumed that $|F| \le (\frac{4}{9}+o(1))\binom{n}{3}$, each of F and T has size $(\frac{4}{9}+o(1))\binom{n}{3}$ and $|F \setminus T| = o(n^3)$. Thus F can be converted into T by changing $o(n^3)$ edges. By the second part of Lemma 5, $u_j = (\frac{1}{3} + o(1))\, n$ for each $j \in [3]$. Therefore, $\delta_1(T, T_n) = o(n^3)$. Thus F and T_n are $o(n^3)$-close in the edit distance, a contradiction to our assumption.

Since $l = 1$ is impossible (otherwise $F = \binom{[n]}{3}$), it remains to consider the case $l = 2$. By Claim 13 and the routine fact that the maximum of $x(1-x)(x^2+(1-x)^2)^{1/2}$ for $x \in I$ is attained for $x = 1/2$, we have

$$|F| \ge \binom{n}{3} - \frac{u_1 u_2 \sqrt{u_1^2 + u_2^2}}{2} + O(n^2) \ge \frac{n^3}{6} - \frac{n^3}{8\sqrt{2}} + O(n^2),$$

which is strictly larger than $(\frac{4}{9} + o(1))\binom{n}{3}$. This final contradiction proves Theorem 2.

5 Exact Result for the $\{G_0, G_3\}$-Problem

First, we will obtain the conclusion of Theorem 1 under the additional assumptions that G is close to a Turán pattern and its maximum degree is at most that of T_n:

Theorem 6 *There is $\varepsilon > 0$ such that the following holds. Let G be a $\{G_0, G_3\}$-free 3-graph on $n \ge 1/\varepsilon$ vertices such that $|G| \le t_n$, $\Delta(G) \le \Delta(T_n)$, and G is εn^3-close in the edit distance to some T_{V_0,V_1,V_2}. Then G is isomorphic to T_n.*

Then, in Section 5.2, we will show that Theorems 2 and 6 imply Theorem 1.

5.1 Proof of Theorem 6

Suppose on the contrary that Theorem 6 is false. Then for every $\varepsilon > 0$ there is a counterexample. In fact, there are infinitely many counterexamples (otherwise by Lemma 4, we would have eliminated all of them by making ε sufficiently small). Thus we may assume that $\varepsilon \to 0$ and that n, the number of vertices, is arbitrarily large with respect to $1/\varepsilon$.

In order to make the proof more readable, we use the asymptotic notation where all terms depending on ε are hidden. For example, $a = o(n)$ means that $|a| \le f(\varepsilon)n$ for some function $f(\varepsilon)$ that depends on ε only and tend to 0 as $\varepsilon \to 0$.

We will use the following constants that are chosen in this order, each being sufficiently small positive number depending on the previous ones:

$$c_1 \gg c_2 \gg c_3 \gg c_4 \gg c_5 \gg c_6 \gg c_7.$$

We do not try to optimize the inequalities that we derive in the course of the proof.

Let $G \subseteq \binom{[n]}{3}$ satisfy all assumptions of Theorem 6. Choose a *best-fit* partition, that is, a partition $[n] = V_0 \cup V_1 \cup V_2$ such that $|T \setminus G|$ is smallest possible, where $T = T_{V_0,V_1,V_2}$.

By our assumptions, there is a Turán pattern T' with $|T' \setminus G| \le |T' \triangle G| \le \varepsilon n^3$. By the extremality of T, we have

$$|T \setminus G| \le |T' \setminus G| \le \varepsilon n^3 = o(n^3).$$

We conclude that

$$|T| \le |T \cap G| + |T \setminus G| \le |G| + \varepsilon n^3 \le t_n + o(n^3).$$

So, by Lemma 5, we have

$$v_i = (1/3 + o(1))\, n. \tag{6}$$

Let $B = T \setminus G$ and $S = G \setminus T$. We call triples in B *bad* and triples in S *superfluous*. Let $b = |B| \le \varepsilon n^3$ and $s = |S|$.

Since each v_i is at least 4, we cannot remove any triple from T without creating G_0. If $s = 0$, then $G \subseteq T$ and, in fact, $G = T$ by the G_0-freeness of G; thus $|T| = t_n$ and $G \cong T \cong T_n$ by Lemma 4, satisfying Theorem 6. So assume that $s > 0$. Also, $t_n \ge |G| = |T| + s - b \ge t_n + s - b$, so $s \le b$. Summarizing:

$$0 < s \le b = o(n^3). \tag{7}$$

Let $P = \{xy \in \binom{[n]}{2} : |B_{xy}| \ge n/20\}$ be the set of pairs of vertices that belong to at least $n/20$ bad triples. Let $p = |P|$.

Claim 1 *$p \ge b/2n$.*

Proof of Claim. Suppose on the contrary that $p < b/2n$. Let $L \subseteq B$ consist of bad triples that do not contain pairs in P. We have $|L| \ge b - p(n-2) \ge b/2$.

Let l be the number of pairs $(D, E) \in L \times S$ such that $|D \cap E| = 2$. Every bad triple $xyz \in L$ contributes at least $(\frac{1}{3} + o(1))n$ to l. Indeed, if $x, y \in V_i$ and $z \in V_i \cup V_{i+1}$ for some $i \in \mathbb{Z}_3$, then for every $w \in V_{i-1}$ at least one of wxy, wxz, wyz belongs to S (otherwise the quadruple $wxyz$ spans a copy of G_0 in G). Thus $l \ge (b/2)\,(\frac{1}{3} + o(1))n$. On the other hand, the 2-shadow of S has at most $3s$ pairs, so some pair xy is covered by at least $l/(3s)$ triples of L. Thus, by (7),

$$|B_{xy}| \geq |L_{xy}| \geq \frac{(b/2)\,(\frac{1}{3}+o(1))n}{3b} > \frac{n}{20}.$$

Thus $xy \in P$, which contradicts the definition of L. ∎

Claim 2 *There is a vertex x with $|B_x| \geq c_2 n^2$.*

Proof of Claim. Each pair in P can either lie inside some part V_i or connect two parts. We distinguish two cases depending on where the majority of pairs in P go.

Case 1 $|P \cap (\cup_{i=0}^{2} \binom{V_i}{2})| \geq p/2$.

Without loss of generality, suppose that $|P^0| \geq p/6$, where $P^0 = P \cap \binom{V_0}{2}$. Define

$$P' = \{xy \in P^0 : |B_{xy} \cap V_1| \geq n/40\}.$$

Case 1.1 $|P'| \geq p/12$.

For each quadruple $uwxy$ with $xy \in P'$ and $u, w \in B_{xy} \cap V_1$ (at least $(p/12) \times \binom{n/40}{2}$ choices), uwx or uwy is superfluous (otherwise $G[uwxy] \cong G_0$). Therefore, some triple, say $uwx \in S$ with $x \in V_0$, appears for at least

$$\frac{(p/12) \times \binom{n/40}{2}}{s} \geq \frac{(b/(24n)) \times \binom{n/40}{2}}{b} \geq c_1 n$$

choices of y, where we used (7) and Claim 1. This vertex x is in at least $\frac{1}{2} \times c_1 n \times (n/20) \geq c_2 n^2$ bad triples, as required.

Case 1.2 $|P'| < p/12$.

For each pair $xy \in P^0 \setminus P'$ (at least $p/12$ choices), $u \in V_0 \cap B_{xy}$ (at least $|B_{xy}| - n/40 \geq n/40$ choices), and $w \in V_1 \setminus B_{xy}$ (at least $(1/3 - 1/40 + o(1))n$ choices), we have $wxy \in G$ and $uxy \notin G$. Thus, in order to avoid G_3, we have that uwx or uwy is in B. By averaging, some triple, say, $uwx \in B$ with $w \in V_1$ appears for at least $c_1 n$ choices of y in this way. Out of these $c_1 n$ P-pairs connecting such vertices y to $u, x \in V_0$, at least half go to the same vertex, which necessarily has B-degree at least $c_2 n^2$.

Case 2 More than half of edges of P connect two different parts V_i.

Without loss of generality, suppose that at least $p/6$ pairs of P connect V_0 to V_1. Let the 2-graph P^{01} consist of these pairs. Note that any bad triple xyz with $xy \in P^{01}$ satisfies $z \in V_0$. Define

$$P'' = \{xy \in P^{01} : |V_2 \setminus S_{xy}| \geq n/6\}.$$

Case 2.1 $|P''| \geq p/12$.

For every choice of $xy \in P''$, $z \in B_{xy} \subseteq V_0$, and $w \in V_2 \setminus S_{xy}$ (at least $(p/12) \times (n/20) \times (n/6)$ choices), at least one of wxz or wyz is superfluous (to prevent $G[wxyz] \cong G_0$). By averaging, some triple, say, $wxz \in S$ with $w \in V_2$ appears for at least $c_1 n$ choices of y, implying that x has the required B-degree.

Case 2.2 $|P''| < p/12$.

For every choice of $xy \in P^{01} \setminus P''$, say $x \in V_0$ and $y \in V_1$, and distinct $u, w \in S_{xy} \cap V_2$, we have $uwx \in B$ or $uwy \in S$ (to avoid G_3). One of these alternatives occurs at least half of the time. Averaging gives a triple (in B or in S) that appears at least $c_1 n$ times this way. As above, this gives a vertex incident to at least $c_2 n^2$ edges of B. The claim is proved. ▮

Fix some vertex x with $|B_x| \geq c_2 n^2$. The following definitions and assumptions will apply to the rest of the section. Assume without loss of generality that $x \in V_0$. Partition the link 2-graphs B_x and S_x as $B^0 \cup B^1 \cup B^2$ and $S^0 \cup S^1 \cup S^2$, where

$$\begin{aligned} B^0 &= B_x \cap \binom{V_0}{2}, \\ B^1 &= \{yz \in B_x : y \in V_0,\ z \in V_1\}, \\ B^2 &= B_x \cap \binom{V_2}{2}, \\ S^0 &= \{yz \in S_x : y \in V_1,\ z \in V_2\}, \\ S^1 &= S_x \cap \binom{V_1}{2}, \\ S^2 &= \{yz \in S_x : y \in V_0,\ z \in V_2\}. \end{aligned}$$

For $i \in \mathbb{Z}_3$, let $b_i = |B^i|$ and $s_i = |S^i|$.

Let A be a largest subset of V_1 with the property that

$$|S^1[A]| \geq \binom{|A|}{2} - c_3 n^2. \tag{8}$$

Since $A = \emptyset$ satisfies (8), A is well-defined. Let $\alpha = |A|/n$. Also, let

$$C = \{y \in V_0 : |B^1_y| \geq c_4 n\},$$

where $B^1_y = (B^1)_y$ is the set of neighbors of y in the 2-graph B^1. Let $\gamma = |C|/n$.

Let us state a few easy inequalities relating some of the parameters that have just been defined.

By the definition of A, we have

$$s_1 \geq \binom{\alpha n}{2} - c_3 n^2. \tag{9}$$

Also, let us show that

$$b_1 \le \alpha\gamma n^2 + c_4 n^2 + o(n^2). \tag{10}$$

The vertices in $V_0 \setminus C$ are incident to at most $|V_0| \times c_4 n < c_4 n^2$ edges of B^1. Let $C' = \{y \in C : |B^1_y| > \alpha n\}$. Clearly, $C \setminus C'$ is incident to at most $\alpha\gamma n^2$ edges of B^1. For every $y \in C'$, we have $|\overline{S^1}[B^1_y]| > c_3 n^2$ by the definition of α; moreover, for every distinct $u, w \in B^1_y$ with $uw \notin S^1$, we have $uwy \in S$ (to avoid $G[uwxy] \cong G_0$). Thus $|C'| \times c_3 n^2 \le |S|$. By (7), $|C'| = o(n)$, and (10) follows.

Let us estimate l, the number of pairs (E, D) with $E \in B^2$, $D \in S^0$, and $|E \cap D| = 1$ plus the number of pairs (E, w) with $E \in B^2$, $w \in V_1$, and $E \cup \{w\} \in S$. On one hand, every $yz \in B^2$ contributes at least v_1 to l: for every $w \in V_1$ at least one of wxy, wxz, wyz is in S (to prevent $G[wxyz] \cong G_0$). On the other hand, each $D \in S^0$ contributes at most $v_2 - 1$ to l while the number of pairs (E, w) is at most $|S| = o(n^3)$. By (6), we have $b_2 n/3 \le s_0 n/3 + o(n^3)$, i.e.

$$s_0 \ge b_2 + o(n^2). \tag{11}$$

Similarly to above, let us estimate l, the number of pairs (E, D) with $E \in B^0$, $D \in B^1$, and $|E \cap D| = 1$ plus the number of pairs (E, w) where $E \in B^0$, $w \in V_1$, and $E \cup \{w\} \in B$. Each $yz \in B^0$ contributes at least v_1 to l: for every $w \in V_1$, at least one of wxy, wxz, wzy is in B (to avoid G_3). On the other hand, each $D \in B^1$ contributes at most $v_0 - 1$ to l while there are at most $|B| = o(n^3)$ required pairs (E, w). This implies that

$$b_1 \ge b_0 + o(n^2). \tag{12}$$

By (6) every vertex of T (as well as of T_n) has degree $(4/9 + o(1)) \binom{n}{2}$. Since

$$\Delta(T_n) \ge \Delta(G) \ge |G_x| = |T_x| + |S_x| - |B_x|,$$

by the maximum degree assumption of Theorem 6, we conclude that

$$b_0 + b_1 + b_2 \ge s_0 + s_1 + s_2 + o(n^2). \tag{13}$$

The number of triples in $T \setminus G$ that contain x is

$$|(T \setminus G)_x| = |B_x| = b_0 + b_1 + b_2. \tag{14}$$

If we change T by moving x to V_1, then $|(T \setminus G)_x|$ becomes $b_0 + (\binom{n/3}{2} - s_1) + ((n/3)^2 - s_0) + o(n^2)$. By the best-fit property of T, this is at least (14), which implies that

$$b_1 + b_2 + s_0 + s_1 \le \frac{n^2}{6} + o(n^2). \tag{15}$$

If we move x to V_2, then $|(T \setminus G)_x|$ becomes $(\binom{n/3}{2} - s_1) + b_2 + ((n/3)^2 - s_2) + o(n^2)$. Again by the best-fit property of T, we have

$$b_0 + b_1 + s_1 + s_2 \le \frac{n^2}{6} + o(n^2). \tag{16}$$

Claim 3 $s_0 \geq c_5 n^2$.

Proof of Claim. Let us suppose on the contrary that $s_0 < c_5 n^2$. By (11), we have $b_2 \leq c_5 n^2 + o(n^2)$. Thus

$$b_0 + b_1 = |B_x| - b_2 \geq c_2 n^2 - c_5 n^2 + o(n^2) \geq 3c_2 n^2/4 + o(n^2)$$

and by (12), we have $b_1 \geq 3c_2 n^2/8 + o(n^2)$. This, (10), and $\max(\alpha, \gamma) \leq 1/3 + o(1)$ imply that both α and γ are at least $(3c_2/8 - c_4)/(1/3) + o(1) > c_2$.

Let A be as in (8). Take $y \in V_0 \setminus C$. Let $A' = A \setminus B_y^1$. We have $|A'| \geq (\alpha - c_4)n$ and

$$|S^1[A']| \geq \binom{|A'|}{2} - c_3 n^2 \geq \binom{(c_2 - c_4)n}{2} - c_3 n^2 > c_2^2 n^2/3.$$

For every $wz \in S^1[A']$, we have $wyz \in S$ (to avoid G_3 on $wxyz$). Thus $|V_0 \setminus C| \times c_2^2 n^2/3 \leq |S| = o(n^3)$ and, by (6), $\gamma = 1/3 + o(1)$.

Pick $y \in C$, $z \in B_y^1$, and $w \in V_2$. There are at least $\gamma n \times c_4 n \times v_2$ such triples. By (7) and the assumption on s_0, $o(n^3)$ choices satisfy $wyz \in S$ and at most $c_5 n^3$ choices satisfy $wz \in S^0$. For all remaining triples wyz, we have $wy \in S^2$ (to avoid G_0 on $wxyz$). Let $\overline{S} = \{wy : y \in C,\ w \in V_2,\ wy \notin S^2\}$ be the bipartite complement of $S^2[C, V_2]$. Since for each $wy \in \overline{S}$ there are at least $c_4 n$ choices of z, we have $|\overline{S}| c_4 n \leq c_5 n^3 + o(n^3)$. Thus e.g. $|\overline{S}| \leq (c_4 + o(1))\, n^2$. We conclude that

$$s_2 \geq |C| \times |V_2| - |\overline{S}| \geq (1/9 - c_4 + o(1))\, n^2.$$

By (11) and (13), we have

$$(1/9 - c_4)n^2 + s_1 \leq b_0 + b_1 + o(n^2). \tag{17}$$

Inequalities (9), (10), (17), and $b_0 \leq \binom{v_0}{2}$ imply that

$$\frac{1}{9} - c_4 + \frac{\alpha^2}{2} - c_3 \leq \frac{1}{18} + \frac{\alpha}{3} + c_4 + o(1). \tag{18}$$

Inequalities $\frac{1}{9} + \frac{\alpha^2}{2} \leq \frac{1}{18} + \frac{\alpha}{3}$ and $0 \leq \alpha \leq \frac{1}{3}$ imply that $\alpha = 1/3$. It follows from (18) that, for example, $\alpha \geq 1/3 - c_2$, $b_0 \geq (1/18 - c_2)n^2$, $s_1 \geq (\alpha^2/2 - c_2)n^2$, and $b_1 \geq (\alpha/3 - c_2)n^2$. But this contradicts (16). The claim is proved. ∎

Claim 4 $s_1 \geq (1/18 - c_6^2)n^2$.

Proof of Claim. Let

$$V_1' = \left\{ y \in V_1 : \left| S_y \cap \binom{V_2}{2} \right| \leq c_7 n^2 \right\}. \tag{19}$$

By (7), $|V_1 \setminus V_1'| = o(n)$. In particular, the number of S^0-edges intersecting $V_1 \setminus V_1'$ is $o(n^2)$. By Claim 3 and (6), the average S^0-degree of a vertex in V_1' is at least $(3c_5 + o(1))\, n$. Take a vertex $y \in V_1'$ whose S^0-degree is at least this average. Let $D = S^0_y$. For every distinct $u, w \in D$ with $uwy \notin G$, we have $uw \in B^2$ (to avoid G_3). Thus

$$|B^2[D]| \ge \binom{|D|}{2} - c_7 n^2. \tag{20}$$

Fix this D. Let $z \in V_1'$ be arbitrary. Let $D' = D \setminus S^0_z$. For every pair $uw \in \binom{D'}{2}$, we have $uw \notin B^2$ or $uwz \in S$ (to avoid G_0 on $uwxz$). Thus $\binom{|D'|}{2} \le (c_7 + c_7 + o(1))\, n^2$ by (19) and (20). Hence,

$$|S^0_z \cap D| \ge |D| - (2\sqrt{c_7} + o(1))\, n$$

for every $z \in V_1'$. We conclude that

$$|S^0[D, V_1']| \ge (|D| - 2\sqrt{c_7}\, n) \times |V_1'| + o(n^2) \ge |D|\,|V_1'| - \sqrt{c_7}\, n^2. \tag{21}$$

Define $D'' = \{z \in D : |B_z \cap \binom{V_1}{2}| \le c_7 n^2\}$. Then D'' contains all but $o(n)$ vertices of D. Pick $z \in D''$ whose S^0-degree is at least the average, which is at least $(1/3 - \sqrt{c_7}/(3c_5) + o(1))n$ by (21). For every distinct $u, w \in S^0_z$ we have $uwz \in B$ or $uw \in S^1$ (to avoid G_3 on $uwxz$). Thus the edges in the complement $\overline{S^1}$ of S^1 are restricted to pairs that intersect $V_1 \setminus S^0_z$ and to $B_z \cap \binom{V_1}{2}$. Thus

$$|\overline{S^1}| \le (\sqrt{c_7}/(3c_2) + c_7 + o(1))\, n^2 \le c_6^2 n^2/2,$$

proving the claim. ∎

Claim 5 *$b_1 \ge (1/9 - 2c_6)n^2$.*

Proof of Claim. Let $V_0' = \{z \in V_0 : |S_z \cap \binom{V_1}{2}| \le c_7 n^2\}$. By (6) and (7), $|V_0'| = (1/3 + o(1))\, n$. Take $z \in V_0'$. Let $D = V_1' \setminus B^1_z$, where V_1' is defined by (19). For every distinct $u, w \in D$, we have $uwz \in S$ or $uw \notin S^1$ (to avoid G_3 on $uwxz$). Thus

$$\binom{|D|}{2} \le |\overline{S^1}| + |S_z| \le (c_6^2 + c_7 + o(1))\, n^2,$$

where we used Claim 4. Thus $|D| \le ((2c_6^2 + 2c_7)^{1/2} + o(1))\, n$. Since $z \in V_0'$ was arbitrary, it follows that $|\overline{B_1}[V_0, V_1]|$, the number of pairs connecting V_0 to V_1 that are not in B^1, is at most

$$|V_0 \setminus V_0'| \times v_1 + (2c_6^2 + 2c_7)^{1/2}\, n \times (n/3) + o(n^2) \le 2c_6 n^2,$$

giving the claim. ∎

Claims 3, 4, and 5 imply that $b_1 + s_0 + s_1 \ge (1/9 - 2c_6 + c_5 + 1/18 - c_6^2 + o(1))\, n^2$, contradicting (15). This final contradiction proves Theorem 6.

5.2 Proof of Theorem 1

Let $\varepsilon > 0$ be the constant returned by Theorem 6. Let $c = c(\varepsilon) > 0$ be the constant returned by Theorem 2 on input ε. Assume that $c \le \varepsilon$. Let us show that $n_0 = (1/c)^3$ suffices. Let G be an arbitrary $\{G_0, G_3\}$-free 3-graph on $n \ge n_0$ vertices with at most t_n edges.

Initially, define $G_n = G$ and $m = n$. If $\Delta(G_m) \le \Delta(T_m)$, then we stop. Otherwise, pick a vertex x of G_m of degree at least $\Delta(T_m) + 1$, let $G_{m-1} = G_m - x$ be obtained from G_m by removing this vertex x (and all edges that contain it), decrease m by 1, and repeat.

When we stop, then $m \ge 2$ and we have

$$0 \le |G_m| \le t_m - (n-m) \le \binom{m}{3} + m - n < m^3 - n. \tag{22}$$

Thus $m > n^{1/3} \ge 1/c$. Theorem 2 implies that G_m is εm^3-close to T_m in the edit distance. (Note that $|G_m| \le t_m$ by (22).) Since $m \ge 1/c \ge 1/\varepsilon$, Theorem 6 implies that $G_m \cong T_m$. By (22), we have $m = n$. Thus $G \cong T_n$, proving Theorem 1.

Acknowledgement

The author thanks Peter Keevash and Alexander Razborov for helpful comments.

References

[BCL+08] C. Borgs, J. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, *Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing*, Adv. Math. **219** (2008), 1801–1851.

[Bro83] W. G. Brown, *On an open problem of Paul Turán concerning* 3*-graphs*, Studies in pure mathematics, Birkhäuser, Basel, 1983, pp. 91–93.

[BT10] R. Baber and J. Talbot, *Hypergraphs do jump*, To appear in Combin. Probab. Computing, 2010.

[CL99] F. Chung and L. Lu, *An upper bound for the Turán number* $t_3(n, 4)$, J. Combin. Theory (A) **87** (1999), 381–389.

[dC94] D. de Caen, *The current status of Turán's problem on hypergraphs*, Extremal problems for finite sets (Visegrád, 1991), Bolyai Soc. Math. Stud., vol. 3, János Bolyai Math. Soc., Budapest, 1994, pp. 187–197.

[ES83] P. Erdős and M. Simonovits, *Supersaturated graphs and hypergraphs*, Combinatorica **3** (1983), 181–192.

[ES08] G. Elek and B. Szegedy, *A measure-theoretic approach to the theory of dense hypergraphs*, E-print `arxiv.org:0810.4062`, 2008.

[FDF88] D. G. Fon-Der-Flaass, *A method for constructing* $(3,4)$*-graphs*, Math. Zeitschrift **44** (1988), 546–550.

[FMP08] Z. Füredi, D. Mubayi, and O. Pikhurko, *Quadruple systems with independent neighborhoods*, J. Combin. Theory (A) **115** (2008), 1552–1560.

[FS05] Z. Füredi and M. Simonovits, *Triple systems not containing a Fano configuration*, Combin. Probab. Computing **14** (2005), 467–488.

[Für91] Z. Füredi, *Turán type problems*, Surveys in Combinatorics, London Math. Soc. Lecture Notes Ser., vol. 166, Cambridge Univ. Press, 1991, pp. 253–300.

[Kee11] P. Keevash, *Hypergraph Turán problem*, Preprint, 2011.

[KM04] P. Keevash and D. Mubayi, *Stability results for cancellative hypergraphs*, J. Combin. Theory (B) **92** (2004), 163–175.

[Kos82] A. V. Kostochka, *A class of constructions for Turán* $(3,4)$*-problem*, Combinatorica **2** (1982), 187–192.

[KS05a] P. Keevash and B. Sudakov, *On a hypergraph Turán problem of Frankl*, Combinatorica **25** (2005), 673–706.

[KS05b] ______, *The Turán number of the Fano plane*, Combinatorica **25** (2005), 561–574.

[LS06] L. Lovász and B. Szegedy, *Limits of dense graph sequences*, J. Combin. Theory (B) **96** (2006), 933–957.

[MP07] D. Mubayi and O. Pikhurko, *A new generalization of Mantel's theorem to* k*-graphs*, J. Combin. Theory (B) **97** (2007), 669–678.

[Pik08] O. Pikhurko, *An exact Turán result for the generalized triangle*, Combinatorica **28** (2008), 187–208.

[Raz07] A. Razborov, *Flag algebras*, J. Symb. Logic **72** (2007), 1239–1282.

[Raz10a] ______, *On 3-hypergraphs with forbidden 4-vertex configurations*, SIAM J. Discr. Math. **24** (2010), 946–963.

[Raz10b] ______, *On the Fon-der-Flaass interpretation of extremal examples for Turán's* $(3,4)$*-problem*, Manuscript, 2010.

[RS09] V. Rödl and M. Schacht, *Generalizations of the Removal Lemma*, Combinatorica **29** (2009), 467–501.

[Sid95] A. Sidorenko, *What we know and what we do not know about Turán numbers*, Graphs Combin. **11** (1995), 179–199.

[Sim68] M. Simonovits, *A method for solving extremal problems in graph theory, stability problems*, Theory of Graphs (Proc. Colloq., Tihany, 1966), Academic Press, 1968, pp. 279–319.

[Tur41] P. Turán, *On an extremal problem in graph theory (in Hungarian)*, Mat. Fiz. Lapok **48** (1941), 436–452.