



# THE TYPICAL STRUCTURE OF MAXIMAL TRIANGLE-FREE GRAPHS

JÓZSEF BALOGH, HONG LIU, ŠÁRKA PETŘÍČKOVÁ and  
MARYAM SHARIFZADEH

Department of Mathematical Sciences, University of Illinois at Urbana-Champaign,  
Urbana, IL 61801, USA;

email: jobal@math.uiuc.edu, hliu36@illinois.edu, petrckv2@illinois.edu, sharifz2@illinois.edu

Received 12 January 2015; accepted 21 August 2015

## Abstract

Recently, settling a question of Erdős, Balogh, and Petříčková showed that there are at most  $2^{n^2/8+o(n^2)}$   $n$ -vertex maximal triangle-free graphs, matching the previously known lower bound. Here, we characterize the typical structure of maximal triangle-free graphs. We show that almost every maximal triangle-free graph  $G$  admits a vertex partition  $X \cup Y$  such that  $G[X]$  is a perfect matching and  $Y$  is an independent set.

Our proof uses the Ruzsa–Szemerédi removal lemma, the Erdős–Simonovits stability theorem, and recent results of Balogh, Morris, and Samotij and Saxton and Thomason on characterization of the structure of independent sets in hypergraphs. The proof also relies on a new bound on the number of maximal independent sets in triangle-free graphs with many vertex-disjoint  $P_3$ s, which is of independent interest.

2010 Mathematics Subject Classification: 05C35, 05C30, 05C69

## 1. Introduction

Given a family of combinatorial objects with certain properties, a fundamental problem in extremal combinatorics is to describe the *typical* structure of these objects. This was initiated in a seminal work of Erdős, Kleitman and Rothschild [13] in 1976. They proved that almost all triangle-free graphs on  $n$  vertices are bipartite; that is, the proportion of  $n$ -vertex triangle-free graphs that are not bipartite goes to zero as  $n \rightarrow \infty$ . Since then, various extensions of this theorem

© The Author(s) 2015. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

have been established. The typical structure of  $H$ -free graphs has been studied when  $H$  is a large clique [3, 19],  $H$  is a fixed color-critical subgraph [23],  $H$  is a finite family of subgraphs [2], and  $H$  is an induced subgraph [4]. For sparse  $H$ -free graphs, analogous problems were examined in [9, 21]. In the context of other combinatorial objects, the typical structure of hypergraphs with a fixed forbidden subgraph is investigated for example in [10, 22], and the typical structure of intersecting families of discrete structures is studied in [5]; see also [1] for a description of the typical sum-free set in finite abelian groups.

In contrast to the family of all  $n$ -vertex triangle-free graphs, which has been well studied, very little was known about the subfamily consisting of all those that are maximal (under graph inclusion) triangle free. Note that the size of the family of triangle-free graphs on  $[n]$  is at least  $2^{n^2/4}$  (all subgraphs of a complete balanced bipartite graph), and at most  $2^{n^2/4+o(n^2)}$  by the result of Erdős, Kleitman, and Rothschild from 1976. Until recently, it was not even known if the subfamily of maximal triangle-free graphs is significantly smaller. As a first step, Erdős suggested the following problem (as stated in [26]): determine or estimate the number of maximal triangle-free graphs on  $n$  vertices. The following folklore construction shows that there are at least  $2^{n^2/8}$  maximal triangle-free graphs on the vertex set  $[n] := \{1, \dots, n\}$ .

*Lower bound construction.* Assume that  $n$  is a multiple of 4. Start with a graph on a vertex set  $X \cup Y$  with  $|X| = |Y| = n/2$  such that  $X$  induces a perfect matching and  $Y$  is an independent set (see Figure 1a). For each pair of a matching edge  $x_1x_2$  in  $X$  and a vertex  $y \in Y$ , add exactly one of the edges  $x_1y$  or  $x_2y$ . Since there are  $n/4$  matching edges in  $X$  and  $n/2$  vertices in  $Y$ , we obtain  $2^{n^2/8}$  triangle-free graphs. These graphs may not be maximal triangle free, but since no further edges can be added between  $X$  and  $Y$ , all of these  $2^{n^2/8}$  graphs extend to distinct maximal ones.

Balogh and Petříčková [11] recently proved a matching upper bound, that the number of maximal triangle-free graphs on vertex set  $[n]$  is at most  $2^{n^2/8+o(n^2)}$ . Now that the counting problem is resolved, one would naturally ask what most of the maximal triangle-free graphs look like; that is, what their typical structure is. Our main result provides an answer to this question.

**THEOREM 1.1.** *For almost every maximal triangle-free graph  $G$  on  $[n]$ , there is a vertex partition  $X \cup Y$  such that  $G[X]$  is a perfect matching and  $Y$  is an independent set.*

The proof of Theorem 1.1 consists of two parts. In the first part, we show an asymptotic version of Theorem 1.1, which implies that almost all maximal triangle-free graphs have a structure very close to the desired one (see the

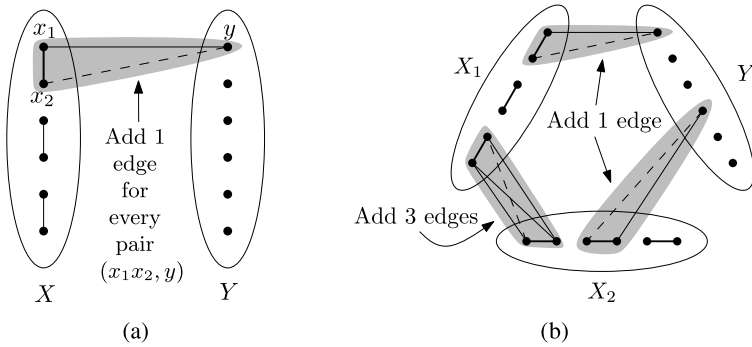


Figure 1. (a)  $r = 2$ , (b)  $r = 3$ . Lower bound construction for maximal  $K_{r+1}$ -free graphs.

beginning of Section 3 for an outline of the proof). In the second part, we compare directly the size of the family of ‘bad’ maximal triangle-free graphs, that is, those without the desired structure, with the size of the family of ‘good’ ones (see the beginning of Section 4 for the idea of the proof).

It is worth mentioning that, once a maximal triangle-free graph has the above partition  $X \cup Y$ , then there has to be exactly one edge between every matching edge of  $X$  and every vertex of  $Y$ . Thus Theorem 1.1 implies that almost all maximal triangle-free graphs have the same structure as the graphs in the lower bound construction above. Furthermore, our proof yields that the number of maximal triangle-free graphs without the desired structure is exponentially smaller than the number of maximal triangle-free graphs. Let  $\mathcal{M}_3(n)$  denote the set of all maximal triangle-free graphs on  $[n]$ , and let  $\mathcal{G}(n)$  denote the family of graphs from  $\mathcal{M}_3(n)$  that admit a vertex partition such that one part induces a perfect matching and the other is an independent set. Then there exists an absolute constant  $c > 0$  such that, for  $n$  sufficiently large,  $|\mathcal{M}_3(n) - \mathcal{G}(n)| \leq 2^{-cn} |\mathcal{M}_3(n)|$ .

It would be interesting to have similar results for  $\mathcal{M}_r(n)$ , the number of maximal  $K_r$ -free graphs on  $[n]$ . Alon pointed out that, if the number of maximal  $K_r$ -free graphs is  $2^{c_r n^2 + o(n^2)}$ , then  $c_r$  is monotone increasing in  $r$ , though not necessarily strictly monotone. For the lower bound, a discussion with Alon and Łuczak led to the following construction that gives  $2^{(1-1/r+o(1))n^2/4}$  maximal  $K_{r+1}$ -free graphs. Assume that  $n$  is a multiple of  $2r$ . Partition the vertex set  $[n]$  into  $r$  equal classes  $X_1, \dots, X_{r-1}, Y$ , and place a perfect matching into each of  $X_1, \dots, X_{r-1}$  (see Figure 1b). Between the classes we have the following connection rule: between the vertices of two matching edges from different classes  $X_i$  and  $X_j$  place exactly three edges, and between a vertex in  $Y$  and a matching edge in  $X_i$  put exactly one edge. For the upper bound, by Erdős, Frankl and

Rödl [12],  $\mathcal{M}_{r+1}(n) \leq 2^{(1-1/r+o(1))n^2/2}$ . A slightly improved bound is given in [11]: For every  $r$  there is  $\varepsilon(r) > 0$  such that  $|\mathcal{M}_{r+1}(n)| \leq 2^{(1-1/r-\varepsilon(r))n^2/2}$  for  $n$  sufficiently large. We suspect that the lower bound is the ‘correct value’; that is, that  $|\mathcal{M}_{r+1}(n)| = 2^{(1-1/r+o(1))n^2/4}$ .

*Related problem.* There is a surprising connection between the family of maximal triangle-free graphs and the family of maximal sum-free sets in  $[n]$ . Recently, Balogh, Liu, Sharifzadeh and Treglown [6] proved that the number of maximal sum-free sets in  $[n]$  is  $2^{(1+o(1))n/4}$ , settling a conjecture of Cameron and Erdős. Although neither of the results imply one another, the methods in both of the papers fall in the same general framework, in which a rough structure of the family is obtained first using the appropriate container lemma and removal lemma. These are Theorems 2.1 and 2.2 in this paper, and a group removal lemma of Green [16], and a granular theorem of Green and Ruzsa [17] in the sum-free case. Both problems can then be translated into bounding the number of maximal independent sets in some auxiliary link graphs. In particular, one of the tools here (Lemma 2.4) is also utilized in [7] to give an asymptotic of the number of maximal sum-free sets in  $[n]$ .

*Organization.* We first introduce all the tools in Section 2, then we prove Lemma 3.1, the asymptotic version of Theorem 1.1, in Section 3. Using this asymptotic result, we prove Theorem 1.1 in Section 4.

*Notation.* For a graph  $G$ , denote by  $|G|$  the number of vertices in  $G$ , and by  $e(G)$  the number of its edges. An  $n$ -vertex graph  $G$  is  $t$ -close to bipartite if  $G$  can be made bipartite by removing at most  $t$  edges. Denote by  $P_k$  the path on  $k$  vertices. Write  $\text{MIS}(G)$  for the number of maximal independent sets in  $G$ . The Cartesian product  $G \square H$  of graphs  $G$  and  $H$  is a graph with vertex set  $V(G) \times V(H)$  such that two vertices  $(u, u')$  and  $(v, v')$  are adjacent if and only if either  $u = v$  and  $u'v' \in E(H)$ , or  $u' = v'$  and  $uv \in E(G)$ . For a fixed graph  $G$ , let  $N(v)$  be the set of neighbors of a vertex  $v$  in  $G$ , and let  $d(v) := |N_G(v)|$  and  $\Gamma(v) := N(v) \cup \{v\}$ . For  $v \in V(G)$  and  $X \subseteq V(G)$ , denote by  $N_X(v)$  the set of all neighbors of  $v$  in  $X$  (that is,  $N_X(v) = N(v) \cap X$ ), and let  $d_X(v) := |N_X(v)|$ . Denote by  $\Delta(X)$  the maximum degree of the induced subgraph  $G[X]$ . For two disjoint vertex sets  $X, Y \subseteq V$ , the edges between  $X$  and  $Y$  are called  $[X, Y]$ -edges, and the number of  $[X, Y]$ -edges is denoted  $e(X, Y)$ . A (vertex) cut  $X \cup Y$  is a partition of the vertex set  $V$  into two disjoint subsets  $X$  and  $Y$ , and  $e(X, Y)$  is the size of the cut  $X \cup Y$ . A vertex cut  $X \cup Y$  is a max-cut if  $e(X, Y)$  is not smaller than the size of any other cut. Given a vertex cut  $X \cup Y$ , the inner edges (of  $X \cup Y$ ) are the edges in  $G[X]$  and  $G[Y]$ , the inner neighbors of a vertex  $v$  are its neighbors in the same partite set as  $v$  (that is,  $N_X(v)$  if  $v \in X$ ), and the inner degree of a vertex is the number of its inner neighbors. We say that a family  $\mathcal{F}$  of maximal

triangle-free graphs is *negligible* if there exists an absolute constant  $C > 0$  such that  $|\mathcal{F}| < 2^{-Cn} |\mathcal{M}_3(n)|$ .

## 2. Tools

Our first tool is a corollary of recent powerful counting theorems of Balogh, Morris and Samotij [8, Theorem 2.2.], and of Saxton and Thomason [25].

**THEOREM 2.1.** *For all  $\delta > 0$  there is  $c = c(\delta) > 0$  such that there is a family  $\mathcal{F}$  of at most  $2^{c \cdot \log n \cdot n^{3/2}}$  graphs on  $[n]$ , each containing at most  $\delta n^3$  triangles, such that for every triangle-free graph  $G$  on  $[n]$  there is an  $F \in \mathcal{F}$  such that  $G \subseteq F$ , where  $n$  is sufficiently large.*

The graphs in  $\mathcal{F}$  in the above theorem will be referred to as *containers*. A weaker version of Theorem 2.1, which can be concluded from the Szemerédi regularity lemma, could be used instead of Theorem 2.1 here. The only difference is that the upper bound on the size of  $\mathcal{F}$  is  $2^{o(n^2)}$ .

We need two well-known results. The first is the Ruzsa–Szemerédi triangle-removal lemma [24], and the second is the Erdős–Simonovits stability theorem [14].

**THEOREM 2.2.** *For every  $\varepsilon > 0$  there exist  $\delta = \delta(\varepsilon) > 0$  and  $n_0(\varepsilon) > 0$  such that any graph  $G$  on  $n > n_0(\varepsilon)$  vertices with at most  $\delta n^3$  triangles can be made triangle free by removing at most  $\varepsilon n^2$  edges.*

**THEOREM 2.3.** *For every  $\varepsilon > 0$  there exists  $\delta = \delta(\varepsilon) > 0$  and  $n_0(\varepsilon) > 0$  such that every triangle-free graph  $G$  on  $n > n_0(\varepsilon)$  vertices with at least  $n^2/4 - \delta n^2$  edges can be made bipartite by removing at most  $\varepsilon n^2$  edges.*

We also need the following lemma, which is an extension of results of Moon and Moser [20] and of Hujter and Tuza [18].

**LEMMA 2.4.** *Let  $G$  be an  $n$ -vertex triangle-free graph. If  $G$  contains at least  $k$  vertex-disjoint  $P_3$ s, then*

$$\text{MIS}(G) \leq 2^{n/2-k/25}. \quad (1)$$

*Proof.* The proof is by induction on  $n$ . The base case of the induction is  $n = 1$  with  $k = 0$ , for which  $\text{MIS}(G) = 1 \leq 2^{1/2-0/25}$ .

For the inductive step, let  $G$  be a triangle-free graph on  $n \geq 2$  vertices with  $k$  vertex-disjoint  $P_3$ s, and let  $v$  be any vertex in  $G$ . Observe that  $\text{MIS}(G - \Gamma(v))$

is the number of maximal independent sets containing  $v$ , and that  $\text{MIS}(G - \{v\})$  bounds from above the number of maximal independent sets not containing  $v$ . Therefore,

$$\text{MIS}(G) \leq \text{MIS}(G - \{v\}) + \text{MIS}(G - \Gamma(v)).$$

If  $G$  has  $k$  vertex-disjoint  $P_3$ s, then  $G - \Gamma(v)$  has at least  $k - d(v)$  vertex-disjoint  $P_3$ s, and so, by the induction hypothesis,

$$\begin{aligned} \text{MIS}(G) &\leq 2^{(n-1)/2-(k-1)/25} + 2^{(n-d(v)+1)/2-(k-d(v))/25} \\ &\leq 2^{n/2-k/25} (2^{-1/2+1/25} + 2^{-(d(v)+1)/2+d(v)/25}). \end{aligned}$$

The function  $f(x) = 2^{-1/2+1/25} + 2^{-(x+1)/2+x/25}$  is a decreasing function with  $f(3) \approx 0.9987 < 1$ . So, if there exists a vertex of degree at least 3 in  $G$ , then we have  $\text{MIS}(G) \leq 2^{n/2-k/25}$ , as desired.

It remains to verify (1) for graphs with  $\Delta(G) \leq 2$ . Observe that we can assume that  $G$  is connected. Indeed, if  $G_1, \dots, G_l$  are maximal components of  $G$ , and each of  $G_i$  has  $n_i$  vertices and  $k_i$  vertex-disjoint  $P_3$ s, then

$$\text{MIS}(G) = \prod_i \text{MIS}(G_i) \leq \prod_i 2^{n_i/2-k_i/25} = 2^{\sum_i n_i/2 - \sum_i k_i/25} = 2^{n/2-k/25}.$$

Every connected graph with  $\Delta(G) \leq 2$  and  $n \geq 2$  vertices is either a path or a cycle. Suppose first that  $G$  is a path  $P_n$ . We have  $\text{MIS}(P_2) = 2 \leq 2^{2/2-0/25}$ ,  $\text{MIS}(P_3) = 2 \leq 2^{3/2-1/25}$ . By Füredi [15, Example 1.1],  $\text{MIS}(P_n) = \text{MIS}(P_{n-2}) + \text{MIS}(P_{n-3})$  for all  $n \geq 4$ . By the induction hypothesis, thus

$$\begin{aligned} \text{MIS}(P_n) &\leq 2^{(n-2)/2-(k-1)/25} + 2^{(n-3)/2-(k-1)/25} \\ &\leq 2^{n/2-k/25} (2^{-1+1/25} + 2^{-3/2+1/25}) \leq 2^{n/2-k/25}. \end{aligned}$$

Let now  $G$  be a cycle  $C_n$ . We have  $\text{MIS}(C_4) = 2 \leq 2^{4/2-1/25}$  and  $\text{MIS}(C_5) = 5 \leq 2^{5/2-1/25}$ . By Füredi [15, Example 1.2],  $\text{MIS}(C_n) = \text{MIS}(C_{n-2}) + \text{MIS}(C_{n-3})$  for all  $n \geq 6$ . Therefore, by the induction hypothesis,

$$\text{MIS}(C_n) \leq 2^{(n-2)/2-(k-1)/25} + 2^{(n-3)/2-(k-1)/25} \leq 2^{n/2-k/25}. \quad \square$$

REMARK 2.5. A disjoint union of  $C_5$ s together with a matching shows that the constant  $c$  for which  $\text{MIS}(G) \leq 2^{n/2-k/c}$  in Lemma 2.4 cannot be smaller than 5.6.

### 3. Asymptotic result

In this section, we prove an asymptotic version of Theorem 1.1:

LEMMA 3.1. Fix any  $\gamma > 0$ . Almost every maximal triangle-free graph  $G$  on the vertex set  $[n]$  satisfies the following: for any max-cut  $V(G) = X \cup Y$ , there exist  $X' \subseteq X$  and  $Y' \subseteq Y$  such that the following hold.

- (i)  $|X'| \leq \gamma n$  and  $G[X - X']$  is an induced perfect matching.
- (ii)  $|Y'| \leq \gamma n$  and  $Y - Y'$  is an independent set.

The outline of the proof is as follows. We observe that every maximal triangle-free graph  $G$  on  $[n]$  can be built in the following three steps.

- (S1) Choose a max-cut  $X \cup Y$  for  $G$ .
- (S2) Choose triangle-free graphs  $S$  and  $T$  on the vertex sets  $X$  and  $Y$ , respectively.
- (S3) Extend  $S \cup T$  to a maximal triangle-free graph by adding edges between  $X$  and  $Y$ .

We give an upper bound on the number of choices for each step. First, there are at most  $2^n$  ways to fix a max-cut  $X \cup Y$  in (S1). For (S2), we show (Lemma 3.5) that almost all maximal triangle-free graphs on  $[n]$  are  $o(n^2)$ -close to bipartite, which implies that the number of choices for most of these graphs in (S2) is at most  $2^{o(n^2)}$ . For fixed  $X, Y, S, T$ , we bound, using Claim 3.4, the number of choices in (S3) by the number of maximal independent sets in some auxiliary link graph  $L$ . This enables us to use Lemma 2.4 to force the desired structure on  $S$  and  $T$ .

DEFINITION 3.2 (Link graph). Given edge-disjoint graphs  $S$  and  $A$  on  $[n]$ , define the link graph  $L := L_S[A]$  of  $S$  on  $A$  as follows:

$$\begin{aligned} V(L) &:= E(A), \\ E(L) &:= \{a_1 a_2 : \exists s \in E(S) \text{ such that } \{a_1, a_2, s\} \text{ forms a triangle}\}. \end{aligned}$$

CLAIM 3.3. Let  $S$  and  $A$  be two edge-disjoint graphs on  $[n]$ . If  $S$  is triangle free, then  $L_S[A]$  is triangle free.

*Proof.* Indeed, otherwise there exist  $a_1, a_2, a_3 \in E(A)$  and  $s_1, s_2, s_3 \in E(S)$  such that the 3-sets  $\{a_1, a_2, s_1\}$ ,  $\{a_2, a_3, s_2\}$ , and  $\{a_1, a_3, s_3\}$  span triangles. Since  $S$  and  $A$  are edge disjoint, the edges  $a_1, a_2, a_3$  share a common endpoint, and  $\{s_1, s_2, s_3\}$  spans a triangle. This is a contradiction, since  $S$  is triangle free.  $\square$

CLAIM 3.4. Let  $S$  and  $A$  be two edge-disjoint triangle-free graphs on  $[n]$  such that there is no triangle  $\{a, s_1, s_2\}$  in  $S \cup A$  with  $a \in E(A)$  and  $s_1, s_2 \in E(S)$ . Then

the number of maximal triangle-free subgraphs of  $S \cup A$  containing  $S$  is exactly  $\text{MIS}(L_S[A])$ .

*Proof.* First observe that, by our assumption, every triangle in  $S \cup A$  consists of two edges from  $E(A)$  and one edge from  $E(S)$ . It follows that, for a subgraph  $A' \subseteq A$ , the graph  $G = S \cup A'$  is triangle free if and only if  $E(A')$  spans an independent set in  $L := L_S[A]$ .

A triangle-free graph  $G = S \cup A'$  is not a maximal triangle-free subgraph of  $S \cup A$  if and only if there exists an edge  $a \in E(A) - (A')$  such that, for any two edges  $a' \in E(A')$  and  $s \in E(S)$ ,  $\{a, a', s\}$  does not form a triangle. By the definition of a link graph  $L_S[A]$ , this is exactly when there exists a vertex  $a \in E(A) - (A')$  such that the set  $E(A') \cup \{a\}$  is an independent set, that is, when  $E(A')$  is not maximal independent set in  $L_S[A]$ . □

We fix the following parameters that will be used throughout the rest of the paper. Let  $\gamma, \beta, \varepsilon, \varepsilon' > 0$  be sufficiently small constants satisfying the following hierarchy:

$$\varepsilon' \ll \delta_{2,3}(\varepsilon) \ll \varepsilon \ll \beta \ll \delta_{2,3}(\gamma^3) \ll \gamma \ll 1, \tag{2}$$

where  $\delta_{2,3}(x) > 0$  is the constant returned from Theorem 2.3 with input  $x$ . The notation  $x \ll y$  above means that  $x$  is a sufficiently small function of  $y$  to satisfy some inequalities in the proof. In the following proof,  $\delta_{2,2}(x)$  is the constant returned from Theorem 2.2 with input  $x$ , and, in the rest of the paper, we shall always assume that  $n$  is sufficiently large, even when this is not explicitly stated.

**LEMMA 3.5.** *Almost all maximal triangle-free graphs on  $[n]$  are  $2\varepsilon n^2$ -close to bipartite.*

*Proof.* Let  $\mathcal{F}$  be the family of graphs obtained from Theorem 2.1 using  $\delta_{2,2}(\varepsilon')$ . Then every triangle-free graph on  $[n]$  is a subgraph of some container  $F \in \mathcal{F}$ .

We first show that the family of maximal triangle-free graphs in small containers is negligible. Consider a container  $F \in \mathcal{F}$  with  $e(F) \leq n^2/4 - 6\varepsilon n^2$ . Since  $F$  contains at most  $\delta_{2,2}(\varepsilon')n^3$  triangles, by Theorem 2.2, we can find  $A$  and  $B$ , subgraphs of  $F$ , such that  $F = A \cup B$ , where  $A$  is triangle free, and  $e(B) \leq \varepsilon n^2$ . For each  $F \in \mathcal{F}$ , fix such a pair  $(A, B)$ . Then every maximal triangle-free graph in  $F$  can be built in two steps.

- (i) Choose a triangle-free  $S \subseteq B$ .
- (ii) Extend  $S$  in  $A$  to a maximal triangle-free graph.



The number of choices in (i) is at most  $2^{e(B)} \leq 2^{\varepsilon' n^2}$ . Observe that any edge  $A \in E(A)$  that is in a triangle containing two edges from  $S$  cannot be added in step (ii). Therefore we remove all such edges from  $A$  and call the resulting graph  $A'$ . Let  $L := L_S[A']$  be the link graph of  $S$  on  $A'$ . By Claim 3.3,  $L$  is triangle free. Claim 3.4 implies that the number of maximal triangle-free graphs in  $S \cup A$  containing  $S$  (that is, the number of extensions in (ii)) is at most  $\text{MIS}(L)$ . Thus, by Lemma 2.4,

$$\text{MIS}(L) \leq 2^{|A'|/2} \leq 2^{n^2/8-3\varepsilon' n^2}.$$

Therefore, the number of maximal triangle-free graphs in small containers is at most

$$|\mathcal{F}| \cdot 2^{\varepsilon' n^2} \cdot 2^{n^2/8-3\varepsilon' n^2} \leq 2^{n^2/8-\varepsilon' n^2}.$$

From now on, we may consider only maximal triangle-free graphs contained in containers of size at least  $n^2/4 - 6\varepsilon' n^2$ . Let  $F$  be any large container. Recall that, by Theorem 2.2,  $F = A \cup B$ , where  $A$  is triangle free with  $e(A) \geq n^2/4 - 7\varepsilon' n^2$  and  $e(B) \leq \varepsilon' n^2$ . Since  $\varepsilon' \ll \delta_{2,3}(\varepsilon)$ , by Theorem 2.3,  $A$  can be made bipartite by removing at most  $\varepsilon n^2$  edges. Since  $\varepsilon' \ll \varepsilon$ ,  $F$  can be made bipartite by removing at most  $(\varepsilon' + \varepsilon)n^2 \leq 2\varepsilon n^2$  edges. Therefore, every maximal triangle-free graph contained in  $F$  is  $2\varepsilon n^2$ -close to bipartite.  $\square$

Fix  $X, Y, S, T$  as in steps (S1) and (S2). Let  $A$  be the complete bipartite graph with parts  $X$  and  $Y$ . By Claim 3.4, the number of ways to extend  $S \cup T$  in (S3) is at most  $\text{MIS}(L_{S \cup T}[A])$ . The number of ways to fix  $X$  and  $Y$  is at most  $2^n$ , and, by Lemma 3.5, the number of ways to fix  $S$  and  $T$  is at most  $\binom{n^2}{2\varepsilon n^2}$ . It follows that, if  $\text{MIS}(L_{S \cup T}[A])$  is smaller than  $2^{n^2/8-cn^2}$  for some  $c \gg \varepsilon$ , then the family of maximal triangle-free graphs with such  $(X, Y, S, T)$  is negligible.

CLAIM 3.6.  $L_{S \cup T}[A] = S \square T$ .

*Proof.* Note that  $V(L_{S \cup T}[A]) = E(A) = \{(x, y) : x \in X, y \in Y\} = V(S \square T)$ . Using the definition of the Cartesian product,  $(x, y)$  and  $(x', y')$  are adjacent in  $S \square T$  if and only if  $x = x'$  and  $\{y, y'\} \in E(T)$ , or  $y = y'$  and  $\{x, x'\} \in E(S)$ , that is, if and only if  $\{x = x', y, y'\}$  or  $\{x, x', y = y'\}$  form a triangle in  $S \cup A$ . But, by the definition of  $L_{S \cup T}[A]$ , this is exactly when  $(x, y)$  and  $(x', y')$  are adjacent in  $L_{S \cup T}[A]$ .  $\square$

Claim 3.6 allows us to rule out certain structures of  $S$  and  $T$  since, by Lemma 2.4, if  $S \square T$  has many vertex-disjoint  $P_3$ s, then the number of maximal triangle-free graphs with  $S = G[X]$  and  $T = G[Y]$  is much smaller than  $2^{n^2/8}$ .

CLAIM 3.7. For almost all maximal triangle-free  $n$ -vertex graphs  $G$  with a max-cut  $X \cup Y$  the following hold.

- (i)  $|X|, |Y| \geq n/2 - \beta n$ .
- (ii)  $\Delta(X), \Delta(Y) \leq \beta n$ .

*Proof.* Let  $G$  be a maximal triangle-free graph with a max-cut  $X \cup Y$ . By Lemma 3.5, almost all maximal triangle-free graphs are  $2\epsilon n^2$ -close to bipartite, which implies that the number of choices for  $G[X]$  and  $G[Y]$  is at most  $\binom{n^2}{2\epsilon n^2}$ . Denote by  $A$  the complete bipartite graph with partite sets  $X$  and  $Y$ .

For (i), suppose that  $|X| \leq n/2 - \beta n$ . Then  $|X||Y| \leq n^2/4 - \beta^2 n^2$ , and, for any fixed  $S$  on  $X$  and  $T$  on  $Y$ , Lemma 2.4 implies that  $\text{MIS}(L_{S \cup T}[A]) \leq 2^{n^2/8 - \beta^2 n^2/2}$ . Since  $\beta \gg \epsilon$ , it follows from the discussion before Claim 3.6 that the family of maximal triangle-free graphs with such max-cut  $X \cup Y$  is negligible.

For (ii), suppose that  $G$  has a vertex  $x \in X$  of inner degree at least  $\beta n$ . Since  $X \cup Y$  is a max-cut,  $|N_Y(x)| \geq |N_X(x)| \geq \beta n$ . Since  $G$  is triangle free, there is no edge in between  $N_X(x)$  and  $N_Y(x)$ . Let  $A' \subseteq A$  be a graph formed by deleting all edges between  $N_X(x)$  and  $N_Y(x)$  from  $A$ . Define a link graph  $L' := L_{S \cup T}[A']$  of  $S \cup T$  on  $A'$ . In this case, the number of choices for (S3) is at most  $\text{MIS}(L')$ . Since  $L'$  is triangle free (Claim 3.3) and  $|L'| = e(A') \leq |X||Y| - |N_X(x)||N_Y(x)| \leq n^2/4 - \beta^2 n^2$ , it follows from Lemma 2.4 that

$$\text{MIS}(L') \leq 2^{|L'|/2} \leq 2^{n^2/8 - \beta^2 n^2/2}. \quad \square$$

*Proof of Lemma 3.1.* First, we show that, for almost every maximal triangle-free graph  $G$  on  $[n]$  with max-cut  $X \cup Y$  and with  $G[X] = S$  and  $G[Y] = T$ , there are very few vertex-disjoint  $P_3$ s in  $S \cup T$ . Suppose that there exist  $\beta n$  vertex-disjoint  $P_3$ s in  $S$  or in  $T$ , say in  $S$ . Since  $L_{S \cup T}[A] = S \square T$  by Claim 3.6, and, for each of the  $\beta n$  vertex-disjoint  $P_3$ s in  $S$  we obtain  $|T|$  vertex-disjoint  $P_3$ s in  $S \square T$ , the number of vertex-disjoint  $P_3$ s in  $L_{S \cup T}[A]$  is at least  $\beta n|T| = \beta n|Y|$ . By Claim 3.7(i),  $\beta n|Y| \geq \beta n(n/2 - \beta n) \geq \beta n^2/3$ . Then, by Lemma 2.4,

$$\text{MIS}(L_{S \cup T}[A]) \leq 2^{|S \square T|/2 - \beta n^2/75} \leq 2^{n^2/8 - \beta n^2/75}.$$

Since  $\beta \gg \epsilon$ , the family of maximal triangle-free graphs with such  $(X, Y, S, T)$  is negligible. Hence, for almost every maximal triangle-free graph  $G$  with some  $(X, Y, S, T)$ , we can find some induced subgraphs  $S' \subseteq S$  and  $T' \subseteq T$  with  $|S'| \leq 3\beta n$  and  $|T'| \leq 3\beta n$  such that both  $S - S'$  and  $T - T'$  are  $P_3$  free. This implies that each of  $S - S'$  and  $T - T'$  is a union of a matching and an independent set.

Next, we show that at most one of the graphs  $S$  and  $T$  can have a large matching. Suppose that both  $S$  and  $T$  have a matching of size at least  $\beta n$ ; then there are

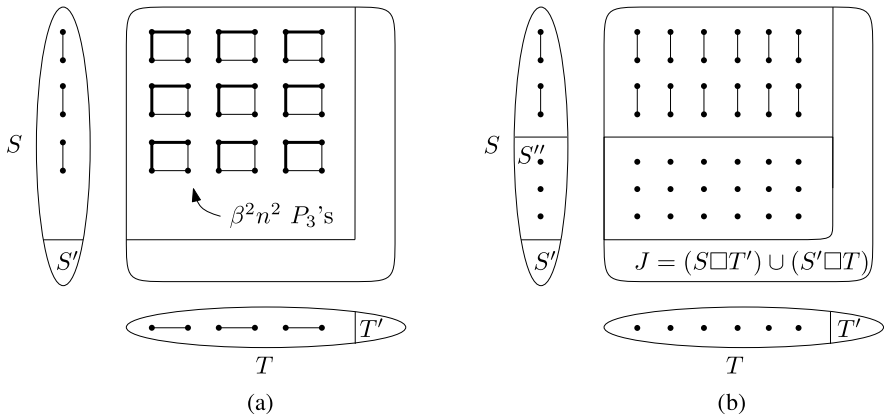


Figure 2. Forbidden structures in  $S$  and  $T$ .

at least  $\beta^2 n^2$  vertex-disjoint  $C_4$ s in  $S \square T$ , each of which contains a copy of  $P_3$  (see Figure 2a). It follows that the family of such graphs is negligible, since  $\text{MIS}(L_{S \cup T}[A]) \leq 2^{n^2/8 - \beta^2 n^2/25}$  and  $\beta \gg \varepsilon$ . Hence, we can assume that all but  $2\beta n$  vertices in  $T$  form an independent set. Redefine  $T'$  so that  $|T'| \leq 2\beta n$  and  $V(T - T')$  is an independent set.

Lastly, we show that there are very few isolated vertices in the graph  $S - S'$ . Suppose that there are  $\gamma n/2$  isolated vertices in  $S - S'$ , spanning a subgraph  $S''$  of  $S$ . We count  $\text{MIS}(S \square T)$  as follows. Let  $J := (S \square T') \cup (S' \square T)$  and  $L' := S \square T - J$ . Every maximal independent set in  $S \square T$  can be built as follows.

- (i) Choose an independent set in  $J$ .
- (ii) Extend it to a maximal independent set in  $L'$ .

Since  $|J| \leq |S'| |T| + |T'| |S| \leq 3\beta n \cdot n + 2\beta n \cdot n = 5\beta n^2$ , there are at most  $2^{|J|} = 2^{5\beta n^2}$  choices for (i). Note that  $L'$  consists of isolated vertices from  $S'' \square (T - T')$  and an induced matching from  $(S - S' - S'') \square (T - T')$  (see Figure 2b). Thus the number of extensions in (ii) is at most  $\text{MIS}((S - S' - S'') \square (T - T'))$ . The graph  $(S - S' - S'') \square (T - T')$  is a perfect matching with at most

$$\frac{1}{2} |S - S''| |T| \leq \frac{1}{2} \left( |S| - \frac{\gamma n}{2} \right) (n - |S|) \leq \frac{1}{2} \left( \frac{n}{2} - \frac{\gamma n}{4} \right)^2 \leq \frac{n^2}{8} - \frac{\gamma n^2}{16}$$

edges, and so choosing one vertex for each matching edge gives at most  $2^{n^2/8 - \gamma n^2/16}$  maximal independent sets. Since  $\beta \ll \gamma$ , it follows that

$\text{MIS}(S \square T) \leq 2^{5\beta n^2} \cdot 2^{n^2/8-\gamma n^2/16} \leq 2^{n^2/8-\gamma n^2/17}$ . Thus, such a family of maximal triangle-free graphs is negligible, and we may assume that  $|S''| \leq \gamma n/2$ .

The statement of Lemma 3.1 follows by setting  $X' := V(S' \cup S'')$  and  $Y' := V(T')$ . Indeed,  $|X'| \leq 3\beta n + \gamma n/2 \leq \gamma n$ ,  $|Y'| \leq 2\beta n \leq \gamma n$ ,  $G[X - X'] = S - S' - S''$  is a perfect matching, and  $Y - Y' = V(T) - V(T')$  is an independent set.  $\square$

#### 4. Proof of Theorem 1.1

In this section, we will prove Theorem 1.1. Recall the hierarchy of parameters fixed in Section 3:

$$\varepsilon' \ll \delta_{2,3}(\varepsilon) \ll \varepsilon \ll \beta \ll \delta_{2,3}(\gamma^3) \ll \gamma \ll 1, \quad (3)$$

We will in fact show that there are exponentially fewer ‘bad’ graphs, that is, maximal triangle-free graphs without the desired structure. We do so by first grouping graphs by some triple  $(X, Y, M)$  (see the definitions below). Then we compare the number of ‘bad’ graphs to the number of ‘good’ graphs within each group by showing that there are not many ‘bad’ ones (Lemmas 4.2 and 4.3), while there are many ‘good’ ones (Lemma 4.5). There might be an overcounting issue due to overlaps among groups. This is taken care of by Lemma 4.4.

**DEFINITION 4.1.** Fix a vertex partition  $V = X \cup Y$ , a perfect matching  $M$  on the vertex set  $X$  (when  $|X|$  is odd,  $M$  is an almost perfect matching covering all but one vertex of  $X$ ), and nonnegative integers  $r, s$  and  $t$ .

- (1) Denote by  $\mathcal{B}(X, Y, M, s, t)$  the class of maximal triangle-free graphs  $G$  with max-cut  $X \cup Y$  satisfying the following three conditions.
  - (i) The subgraph  $G[X]$  has a maximum matching  $M' \subseteq M$  covering all but at most  $\gamma n$  vertices in  $X$ .
  - (ii) The size of a largest family of vertex-disjoint  $P_3$ s in  $S := G[X]$  is  $s$ .
  - (iii) The size of a maximum matching in  $T := G[Y]$  is  $t$ .
- (2) Denote by  $\mathcal{B}(X, Y, M, r) \subseteq \mathcal{B}(X, Y, M, 0, 0)$  the subclass consisting of all graphs in  $\mathcal{B}(X, Y, M, 0, 0)$  with exactly  $r$  isolated vertices in  $G[X]$ .
- (3) When  $|X|$  is even, denote by  $\mathcal{G}(X, Y, M)$  the class of all maximal triangle-free graphs  $G$  with max-cut  $X \cup Y$ ,  $G[X] = M$ , and  $Y$  an independent set.
- (4) When  $|X|$  is even, denote by  $\mathcal{H}(X, Y, M)$  the class of maximal triangle-free graphs  $G$  that are constructed as follows.

- (P1) Add  $M$  to  $X$ .
- (P2) For every edge  $x_1x_2 \in M$  and every vertex  $y \in Y$ , add either the edge  $x_1y$  or  $x_2y$ .
- (P3) Extend each of the  $2^{|X||Y|/2}$  resulting graphs to a maximal triangle-free graph by adding edges in  $X$  and/or  $Y$ .

By Lemmas 3.1, 3.5 and Claim 3.7, throughout the rest of the proof, we may only consider maximal triangle-free graphs in  $\bigcup_{X,Y,M,s,t} \mathcal{B}(X, Y, M, s, t)$  that are  $\beta n^2$ -close to bipartite,  $|X|, |Y| \geq n/2 - \beta n$ , and  $\Delta(X), \Delta(Y) \leq \beta n$ . We may further assume from the proof of Lemma 3.1 that  $s, t \leq \beta n$ .

Notice that graphs from  $\mathcal{G}(X, Y, M) = \mathcal{B}(X, Y, M, 0)$  are precisely those with the desired structure. We will show that the number of graphs without the desired structure is exponentially smaller. The set of ‘bad’ graphs consists of the following two types.

- (i) When  $|X|$  is even,  $\bigcup_{s,t} \mathcal{B}(X, Y, M, s, t) - \mathcal{B}(X, Y, M, 0)$ .
- (ii) When  $|X|$  is odd,  $\bigcup_{s,t} \mathcal{B}(X, Y, M, s, t)$ .

Fix an arbitrary choice of  $(X, Y, M)$ . For simplicity, let  $\mathcal{B}(s, t) := \mathcal{B}(X, Y, M, s, t)$  and  $\mathcal{B}(r) := \mathcal{B}(X, Y, M, r)$ . Let  $A$  be the complete bipartite graph with parts  $X$  and  $Y$ .

LEMMA 4.2. *If  $s + t \geq 1$ , then  $|\mathcal{B}(s, t)| \leq 2^{|X||Y|/2 - n/200}$ .*

*Proof.* Let  $s$  and  $t$  be two nonnegative integers, at least one of which is nonzero. We first bound the number of ways to choose  $S$  and  $T$ , that is, the number of ways to add inner edges. The number of ways to choose the vertex set of the  $s$  vertex-disjoint  $P_3$ s in  $S$  and the  $t$  matching edges in  $T$  is at most  $\binom{n}{3s} \binom{n}{2t}$ . Recall that, by the definition of  $\mathcal{B}(s, t)$ , the maximum matching  $M' \subseteq M$  covers all but at most  $\gamma n$  vertices of  $X$ . So the number of ways to choose the independent vertices in  $X$  is at most  $\binom{n}{\gamma n}$ . Since  $\Delta(X), \Delta(Y) \leq \beta n$ , each of the  $3s + 2t$  chosen vertices has inner degree at most  $\beta n$ . Therefore, the number of ways to choose their inner neighbors is at most

$$\binom{n}{\beta n}^{3s+2t} \leq \left( \left( \frac{en}{\beta n} \right)^{\beta n} \right)^{3s+2t} \leq 2^{\beta \log(e/\beta) \cdot (3s+2t)n}.$$

The number of ways to add the  $[X, Y]$ -edges is  $\text{MIS}(L_{S \cup T}(A))$ . We claim that the link graph  $L := L_{S \cup T}(A) = S \square T$  has at least  $(s + t)n/5$  vertex-disjoint  $P_3$ s. Indeed, recall that  $|S| = |T| \geq n/2 - \beta n$  and  $s, t \leq \beta n$ ; thus, in  $S \square T$  (see

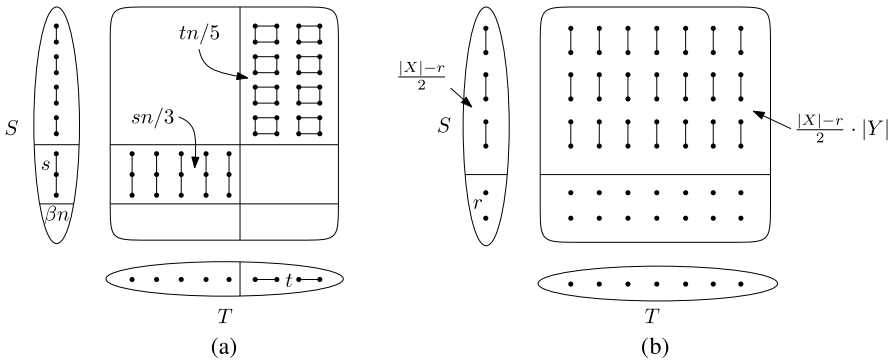


Figure 3. (a) The number of vertex-disjoint  $P_3$ s in  $S \square T$  is at least  $sn/3 + tn/5$  (Lemma 4.2). (b)  $\text{MIS}(S \square T) \leq 2^{(|X|-r)|Y|/2}$  if  $s = t = 0$  and  $X$  has  $r$  isolated vertices (Lemma 4.3).

Figure 3a), we have at least  $s(|T| - 2t) \geq sn/3$  vertex-disjoint  $P_3$ s coming from  $s$  vertex-disjoint  $P_3$ s in  $S$ , and at least  $\frac{1}{2}(|S| - \beta n - 3s) \cdot t \geq tn/5$  vertex-disjoint  $P_3$ s coming from the Cartesian product of a matching in  $S$  and a matching in  $T$ . So, by Lemma 2.4,

$$\text{MIS}(L) \leq 2^{|X||Y|/2 - (s+t)n/125}.$$

Since  $s + t \geq 1$  and  $\gamma$  and  $\beta$  are sufficiently small,

$$\begin{aligned} |\mathcal{B}(s, t)| &\leq \binom{n}{3s} \binom{n}{2t} \binom{n}{\gamma n} \cdot 2^{\beta \log(e/\beta) \cdot (3s+2t)n} \cdot 2^{|X||Y|/2 - (s+t)n/125} \\ &\leq 2^{|X||Y|/2 - n/200}. \end{aligned} \quad \square$$

LEMMA 4.3. *If  $s = t = 0$  and  $r \in \mathbb{Z}^+$ , then  $|\mathcal{B}(r)| \leq 2^{|X||Y|/2 - n/6}$ .*

*Proof.* By the definition of  $\mathcal{B}(r)$ ,  $X$  consists of  $r$  isolated vertices and a matching of size  $(|X| - r)/2$ , and  $Y$  is an independent set. Hence the graph  $L_{S \cup T}(A) = S \square T$  consists of a matching of size  $(|X| - r)|Y|/2$  and isolated vertices (see Figure 3b). There are at most  $\binom{n}{r}$  ways to pick the isolated vertices in  $X$ , and at most  $\text{MIS}(L_{S \cup T}(A))$  ways to choose the  $[X, Y]$ -edges. Recall that  $|Y| \geq n/2 - \beta n$ . Thus we have

$$\begin{aligned} |\mathcal{B}(r)| &\leq \binom{n}{r} \cdot 2^{(|X|-r)|Y|/2} \leq 2^{|X||Y|/2 + r \log n - rn/5} \\ &\leq 2^{|X||Y|/2 - rn/6} \leq 2^{|X||Y|/2 - n/6}. \end{aligned} \quad \square$$

Case 1:  $|X|$  is even.

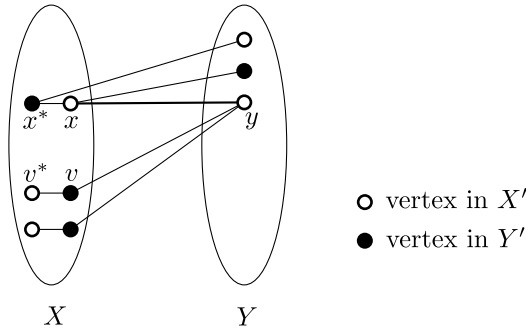


Figure 4.  $(X', Y', M')$  is uniquely determined after choosing  $xy \in M'$  (Lemma 4.4).

LEMMA 4.4. A maximal triangle-free graph  $G$  on  $[n]$  is in at most  $n^2$  different classes  $\mathcal{G}(X, Y, M)$ .

*Proof.* Let  $G \in \mathcal{G}(X, Y, M)$ . Recall that  $G[X] = M$  and  $Y$  is an independent set. Thus  $G$  can be in a different class  $\mathcal{G}(X', Y', M')$  if and only if  $X' \neq X$ ,  $Y' \neq Y$  and  $M' \neq M$ . Since  $M' \neq M$  and  $Y$  is an independent set, there exists an edge  $xy \in M'$  with  $x \in X$  and  $y \in Y$ . There are at most  $n^2$  ways to choose such an edge. We claim that, once we pick the edge  $xy \in M'$ , the sets  $X'$  and  $Y'$  (and thus also  $M' = G[X']$ ) are already decided. Recall that, since  $G$  is a maximal triangle-free graph, every vertex in  $Y$  is adjacent to exactly one vertex from each edge in  $M$ .

Observe that the neighbor  $x^*$  of  $x$  in  $X$  has to be in  $Y'$ , since otherwise there would be a path  $x^*xy$  in  $G[X']$  (see Figure 4). Let  $vv^* \in M - \{xx^*\}$  such that  $vy \in E(G)$ . Then  $v \in Y'$ , since otherwise there would be a path  $vyx$  in  $X'$ . The set  $Y'$  is independent, and so  $v^* \in X'$ . It remains to decide whether  $w \in X'$  or  $w \in Y'$  for every vertex  $w \in Y - \{y\}$ . If  $xw \in E(G)$ , then  $w \in Y'$ , since otherwise we would have a path  $wxy$  in  $G[X']$ . Otherwise  $x^*w \in E(G)$ , and so  $w \in X'$ , since otherwise there would be an edge  $wx^*$  in  $G[Y']$ .  $\square$

By Lemma 4.4, it is sufficient to show that, for any choice of  $(X, Y, M)$  with  $|X|$  even,

$$\frac{|\bigcup_{s,t} \mathcal{B}(X, Y, M, s, t) - \mathcal{B}(X, Y, M, 0)|}{|\mathcal{G}(X, Y, M)|} \leq 2^{-n/300}. \tag{4}$$

For simplicity, let  $\mathcal{G} := \mathcal{G}(X, Y, M)$  and  $\mathcal{H} := \mathcal{H}(X, Y, M)$ .

LEMMA 4.5. We have  $|\mathcal{G}| \geq (1 + o(1))2^{|X||Y|/2}$ .

*Proof.* Recall that  $|X|, |Y| \geq n/2 - \beta n$ , and therefore  $|\mathcal{H}| = 2^{|X||Y|/2} \gg 2^{n^2/8 - \beta n^2}$ . Running the same proof as that of Lemma 3.5 (start the proof by invoking Theorem 2.1 with  $\delta_{2,2}(\beta)$ , and replace  $\varepsilon'$  by  $\beta$  and  $\varepsilon$  by  $\gamma^3$ ) implies that almost all graphs in  $\mathcal{H}$  are  $2\gamma^3 n^2$ -close to bipartite. Let  $\mathcal{H}' \subseteq \mathcal{H}$  be the subfamily consisting of all those that are  $2\gamma^3 n^2$ -close to bipartite. Then it is sufficient to show that  $|\mathcal{H}' - \mathcal{G}| = o(2^{|X||Y|/2})$ . There are two types of graph in  $\mathcal{H}' - \mathcal{G}$ .

- (i)  $\mathcal{H}_1$ : those that are not maximal after (P2).
- (ii)  $\mathcal{H}_2$ : those that are maximal after (P2), but  $X \cup Y$  is not one of its max-cuts.

We first bound the number of graphs in  $\mathcal{H}_1$ . For any graph  $G \in \mathcal{H}_1$ , some inner edges were added in (P3). Suppose that  $[X, Y]$ -edges added in (P2) were chosen randomly (each of  $x_1 y$  and  $x_2 y$  with probability  $1/2$ ). Clearly,  $uv$  can be added in (P3) if and only if  $u$  and  $v$  have no common neighbor. Consider the case when  $u, v \in X$ , and let  $uu', vv'$  be the corresponding edges in  $M$  (see Figure 5a). Every  $y \in Y$  is adjacent to exactly one of  $u, u'$  and exactly one of  $v, v'$ . Thus the probability that  $y$  is a common neighbor of  $u$  and  $v$  is  $1/4$ , which implies that  $uv$  can be added with probability  $(3/4)^{|Y|}$ . Let now  $u, v \in Y$ . Then  $u$  and  $v$  have no common neighbor if and only if, for every  $x_1 x_2 \in M$ ,  $u$  and  $v$  chose different neighbors among  $x_1$  and  $x_2$ . So in this case we can add  $u, v$  with probability  $(1/2)^{|X|/2}$ . Summing over all possible outcomes of (P2) and all possible choices for  $uv$  implies that

$$|\mathcal{H}_1| \leq 2^{|X||Y|/2} \cdot \binom{n}{2} \cdot \left( \left( \frac{1}{2} \right)^{|X|/2} + \left( \frac{3}{4} \right)^{|Y|} \right) \ll 2^{|X||Y|/2 - n/5}.$$

We can bound  $|\mathcal{H}_2|$  in a similar way. It suffices to show that, if the  $[X, Y]$ -edges added in (P2) were chosen uniformly at random, then the probability that  $X \cup Y$  is not a max-cut is  $o(1)$ . Let  $X' \cup Y'$  be a different vertex cut, where we may assume that  $|X' \cap X| \geq |Y' \cap X|$  (see Figure 5b). Then  $|X' \cap X| \geq |X|/2 > n/5$ . Let  $M_{X',Y'}$  be the event that  $X' \cup Y'$  is a cut greater than  $X \cup Y$ , and let  $a := |X' \cap Y|$  and  $b := |Y' \cap X|$ . Recall that the number of inner edges of  $X \cup Y$  is  $e(G[X]) = |X|/2 < n/3$ . If  $a \geq 200$ , then the expected number of edges in  $G[X']$  is at least  $\mathbb{E}[e([X' \cap X, X' \cap Y])] \geq \frac{1}{2} \cdot |X' \cap X|/2 \cdot a \geq 10n$ . Therefore, by the Chernoff bound,  $\mathbb{P}[M_{X',Y'}] \leq \mathbb{P}[e(G[X']) < n/3] = o(2^{-n})$ . We may thus assume that  $a \leq 200$ , which implies that  $|Y' \cap Y| = |Y| - a > n/5$ . If  $b \geq 200$ , then  $\mathbb{E}[e(G[Y'])] \geq \frac{1}{2} \cdot b/2 \cdot |Y' \cap Y| \geq 10n$ , and so  $\mathbb{P}[M_{X',Y'}] \leq \mathbb{P}[e(G[Y']) < n/3] = o(2^{-n})$ . Hence, we may further assume that  $b \leq 200$ . Note now that both  $X' \cap X$  and  $Y' \cap Y$  have size at least  $n/2 - \beta n - 200 \geq n/2 - 2\beta n$ . Since  $X' \cup Y' \neq X \cup Y$ , we have  $a + b \geq 1$ . Hence, the expected number of inner edges of  $X' \cup Y'$  is at least



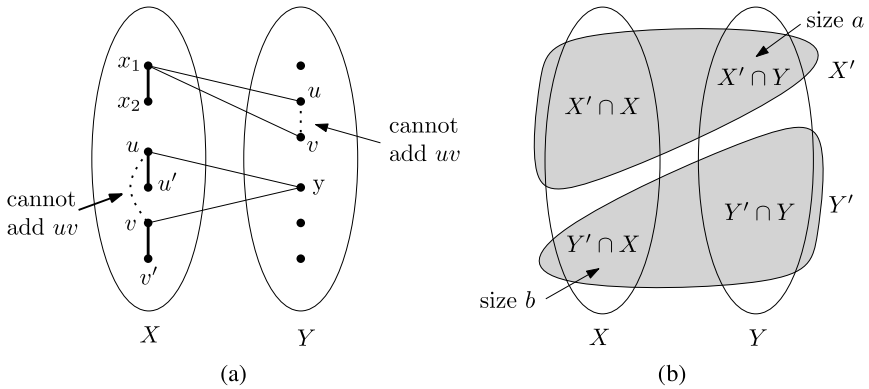


Figure 5. (a) Two examples of  $uv$  that cannot be added when forming a graph from  $\mathcal{H}_1$ . (b) Bounding  $|\mathcal{H}_2|$ : the probability that  $e(X', Y') > e(X, Y)$  is  $o(1)$ . Bounding the number of graphs in  $\mathcal{H}_1$  and  $\mathcal{H}_2$  in Lemma 4.5.

$$\begin{aligned} & \mathbb{E}[e([X' \cap X, X' \cap Y])] + \mathbb{E}[e(G[Y'])] + e(G[X' \cap X]) \\ & \geq \frac{1}{2} \cdot \frac{|X' \cap X|}{2} \cdot a + \frac{1}{2} \cdot \frac{b}{2} \cdot |Y' \cap Y| + (e(G[X]) - b) \\ & \geq \frac{1}{4} \cdot (a + b) \cdot \left(\frac{n}{2} - 2\beta n\right) + \frac{|X|}{2} - b \geq \frac{3n}{8} - 300\beta n. \end{aligned}$$

Thus, by the Chernoff bound,  $\mathbb{P}[M_{X',Y'}] \leq \mathbb{P}[e(G[X']) + e(G[Y']) < n/3] \leq 2^{-n/1000}$ . The number of  $X' \cup Y'$  with  $a, b \leq 200$  is only at most  $\binom{n}{200}^2 \leq n^{400}$ .  $\square$

Since  $s, t, r \leq n$ , Lemmas 4.2, 4.3 and 4.5 imply (4):

$$\begin{aligned} \frac{|\bigcup_{s,t} \mathcal{B}(s, t) - \mathcal{B}(0)|}{|\mathcal{G}|} &= \frac{\sum_{s+t \geq 1} |\mathcal{B}(s, t)| + \sum_{r \geq 1} |\mathcal{B}(r)|}{|\mathcal{G}|} \\ &\leq \frac{(n^2 + n) \cdot 2^{|X||Y|/2 - n/200}}{(1 + o(1))2^{|X||Y|/2}} \leq 2^{-n/300}. \end{aligned}$$

Case 2:  $|X|$  is odd.

Fix an arbitrary choice of  $X, Y, M$  with  $|X|$  odd, and let  $x \in X$  be the vertex not covered by  $M$ . By Lemmas 4.2 and 4.3,

$$\begin{aligned} \left| \bigcup_{s,t} \mathcal{B}(X, Y, M, s, t) \right| &\leq \sum_{s,t:s+t \geq 1} |\mathcal{B}(X, Y, M, s, t)| + \sum_{r \geq 1} |\mathcal{B}(X, Y, M, r)| \\ &\leq 2^{|X||Y|/2 - n/300}. \end{aligned}$$

Pick an arbitrary vertex  $y \in Y$ , and define  $X_0 = X \cup \{y\}$ ,  $Y_0 = Y - \{y\}$ , and  $M_0 = M \cup \{xy\}$ . Then, by Lemma 4.5, we have

$$|\mathcal{G}(X_0, Y_0, M_0)| \geq (1 + o(1))2^{|X_0||Y_0|/2} \geq 2^{|X||Y|/2 - (|X| - |Y|)/2 - 1} \geq 2^{|X||Y|/2 - 2\beta n},$$

since  $|X| - |Y| \leq 2\beta n$ . Notice that any  $(X_0, Y_0, M_0)$  with  $|X_0|$  even can be obtained from at most  $n$  different triples  $(X, Y, M)$  with  $|X|$  odd in this way. Together with Lemma 4.4, it is sufficient to show that  $\bigcup_{s,t} \mathcal{B}(X, Y, M, s, t)$  is negligible compared to  $\mathcal{G}(X_0, Y_0, M_0)$ :

$$\frac{|\bigcup_{s,t} \mathcal{B}(X, Y, M, s, t)|}{|\mathcal{G}(X_0, Y_0, M_0)|} \leq \frac{2^{|X||Y|/2 - n/300}}{2^{|X||Y|/2 - 2\beta n}} \leq 2^{-n/400}.$$

This completes the proof of Theorem 1.1.

### Acknowledgements

The authors would like to thank the anonymous referee for useful comments and suggestions, particularly on the simplification of the proof of Lemma 4.5. The first author's research is partially supported by a Simons Fellowship, NSF CAREER Grant DMS-0745185, Marie Curie FP7-PEOPLE-2012-IIF 327763, and Arnold O. Beckmann Research Award RB15006.

### References

- [1] N. Alon, J. Balogh, R. Morris and W. Samotij, 'Counting sum-free sets in Abelian groups', *Israel J. Math.* **199**(1) (2014), 309–344.
- [2] J. Balogh, B. Bollobás and M. Simonovits, 'The typical structure of graphs without given excluded subgraphs', *Random Structures Algorithms* **34** (2009), 305–318.
- [3] J. Balogh, N. Bushaw, M. Collares Neto, H. Liu, R. Morris and M. Sharifzadeh, 'The typical structure of graphs with no large cliques', *Combinatorica*, to appear.
- [4] J. Balogh and J. Butterfield, 'Excluding induced subgraphs: critical graphs', *Random Structures Algorithms* **38** (2011), 100–120.
- [5] J. Balogh, S. Das, M. Delcourt, H. Liu and M. Sharifzadeh, 'Intersecting families of discrete structures are typically trivial', *J. Combin. Theory Ser. A* **132** (2015), 224–245.
- [6] J. Balogh, H. Liu, M. Sharifzadeh and A. Treglown, 'The number of maximal sum-free subsets of integers', *Proc. Amer. Math. Soc.* **143**(11) (2015), 4713–4721.
- [7] J. Balogh, H. Liu, M. Sharifzadeh and A. Treglown, 'Sharp bound on the number of maximal sum-free subsets of integers', submitted.
- [8] J. Balogh, R. Morris and W. Samotij, 'Independent sets in hypergraphs', *J. Amer. Math. Soc.* **28**(3) (2015), 669–709.
- [9] J. Balogh, R. Morris, W. Samotij and L. Warnke, 'The typical structure of sparse  $K_{r+1}$ -free graphs', *Trans. Amer. Math. Soc.*, to appear.
- [10] J. Balogh and D. Mubayi, 'Almost all triple systems with independent neighborhoods are semi-bipartite', *J. Combin. Theory Ser. A* **118** (2011), 1494–1518.

- [11] J. Balogh and Š. Petříčková, ‘Number of maximal triangle-free graphs’, *Bull. Lond. Math. Soc.* **46**(5) (2014), 1003–1006.
- [12] P. Erdős, P. Frankl and V. Rödl, ‘The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent’, *Graphs Combin.* **2** (1986), 113–121.
- [13] P. Erdős, D. J. Kleitman and B. L. Rothschild, ‘Asymptotic enumeration of  $K_n$ -free graphs’, in *Colloquio Internazionale sulle Teorie Combinatorie, Tomo II*, Atti dei Convegni Lincei, no. 17 (Accad. Naz. Lincei, Rome, 1976), 19–27.
- [14] P. Erdős and M. Simonovits, ‘A limit theorem in graph theory’, *Studia Sci. Math. Hungar.* **1** (1966), 51–57.
- [15] Z. Füredi, ‘The number of maximal independent sets in connected graphs’, *J. Graph Theory* **11**(4) (1987), 463–470.
- [16] B. Green, ‘A Szemerédi-type regularity lemma in abelian groups, with applications’, *Geom. Funct. Anal.* **15** (2005), 340–376.
- [17] B. Green and I. Ruzsa, ‘Counting sumsets and sum-free sets modulo a prime’, *Studia Sci. Math. Hungar.* **41** (2004), 285–293.
- [18] M. Hujter and Z. Tuza, ‘The number of maximal independent sets in triangle-free graphs’, *SIAM J. Discrete Math.* **6**(2) (1993), 284–288.
- [19] P. G. Kolaitis, H. J. Prömel and B. L. Rothschild, ‘ $K_{\ell+1}$ -free graphs: asymptotic structure and a 0-1 law’, *Trans. Amer. Math. Soc.* **303** (1987), 637–671.
- [20] J. W. Moon and L. Moser, ‘On cliques in graphs’, *Israel J. Math.* **3** (1965), 23–28.
- [21] D. Osthus, H. J. Prömel and A. Taraz, ‘For which densities are random triangle-free graphs almost surely bipartite?’, *Combinatorica* **23** (2003), 105–150.
- [22] Y. Person and M. Schacht, ‘Almost all hypergraphs without Fano planes are bipartite’, in *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms* (2009), 217–226.
- [23] H. J. Prömel and A. Steger, ‘The asymptotic number of graphs not containing a fixed color-critical subgraph’, *Combinatorica* **12** (1992), 463–473.
- [24] I. Z. Ruzsa and E. Szemerédi, ‘Triple systems with no six points carrying three triangles’, in *Combinatorics (Keszthely, 1976)*, Coll. Math. Soc. J. Bolyai 18, Volume II 939–945.
- [25] D. Saxton and A. Thomason, ‘Hypergraph containers’, *Invent. Math.* **201**(3) (2015), 925–992.
- [26] M. Simonovits, *Paul Erdős’ Influence on Extremal Graph Theory*, The Mathematics of Paul Erdős, II (Springer, Berlin, 1996), 148–192.