Combinatorics, Probability and Computing (2005) 14, 211–224. © 2005 Cambridge University Press
DOI: 10.1017/S0963548304006613 Printed in the United Kingdom

On the Evolution of Triangle-Free Graphs

ANGELIKA STEGER

Institute of Theoretical Computer Science, ETH Zürich,
8092 Zürich, Switzerland
(e-mail: steger@inf.ethz.ch)

Received 21 August 2003; revised 29 October 2003

Dedicated to the memory of Walter Deuber

Let $\mathscr{T}(n,m)$ denote the set of all labelled triangle-free graphs with n vertices and exactly m edges. In this paper we give a short self-contained proof of the fact that there exists a constant $C > 0$ such that, for all $m \geqslant Cn^{3/2}\sqrt{\log n}$, a graph chosen uniformly at random from $\mathscr{T}(n,m)$ is with probability $1 - o(1)$ bipartite.

1. Introduction

Let $\mathscr{T}(n,m)$ denote the set of all labelled triangle-free graphs with vertex set $[n] = \{1,\dots,n\}$ and exactly m edges and let $T_{n,m}$ denote a graph chosen uniformly at random from $\mathscr{T}(n,m)$. In this paper we are concerned with the probability that the graph $T_{n,m}$ is bipartite. Note that this probability is equal to

$$\Pr[T_{n,m} \text{ is bipartite}] = \frac{|\mathscr{S}(n,m)|}{|\mathscr{T}(n,m)|},$$

where $\mathscr{S}(n,m)$ denote the set of all labelled bipartite graphs with vertex set $[n]$ and exactly m edges.

In 1906 Mantel proved that every triangle-free graph on n vertices has at most $\lfloor n^2/4 \rfloor$ edges and that *every* triangle-free graph with that many edges is bipartite. That is, for $m = \lfloor n^2/4 \rfloor$ we have $\Pr[T_{n,m} \text{ is bipartite}] = 1$. Moreover, Erdős, Kleitman and Rothschild [1] proved that a random triangle-free graph T_n (with no restriction on the number of edges) is almost surely bipartite and has about $m = \frac{1}{8}n^2$ edges. This implies that $\Pr[T_{n,m} \text{ is bipartite}] = 1 - o(1)$ whenever $m \geqslant \frac{1}{8}n^2$. In [8] Prömel and Steger observed that within the evolution from $m = 0$ to $m = \lfloor \frac{1}{4}n^2 \rfloor$ there might be two thresholds. More precisely, they showed the following.

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

Theorem 1.1. ([8]) *There exist constants c_1, c_2, c_3 such that*

$$\Pr[T_{n,m} \text{ is bipartite}] \to \begin{cases} 1 & \text{if } m = o(n), \\ 0 & \text{if } c_1 n \leqslant m \leqslant c_2 n^{3/2}, \\ 1 & \text{if } m \geqslant c_3 n^{7/4} \log n. \end{cases}$$

They also conjectured that the second threshold should in fact be located at $m = n^{3/2+o(1)}$. This conjecture was proved independently by Steger and Osthus, Prömel, Taraz. The latter were in fact able to show that this threshold is located precisely at $t(n) = \frac{\sqrt{3}}{4} n^{3/2} \sqrt{\log n}$.

Theorem 1.2. ([5]) *For any $\varepsilon > 0$,*

$$\Pr[T_{n,m} \text{ is bipartite}] \to \begin{cases} 1 & \text{if } m = o(n), \\ 0 & \text{if } \frac{1}{2} n \leqslant m \leqslant (1-\varepsilon)\frac{\sqrt{3}}{4} n^{3/2} \sqrt{\log n}, \\ 1 & \text{if } m \geqslant (1+\varepsilon)\frac{\sqrt{3}}{4} n^{3/2} \sqrt{\log n}. \end{cases}$$

The main difficulty of this theorem is the proof of the second 1-statement. Here the proof of Osthus, Prömel and Taraz is based on ingenious counting in combination with the correlation inequalities from [2], which bound the probabilities for slightly dependent events. They also rely on a result of Łuczak [4], who proved that for $m \geqslant Cn^{3/2}$ almost all triangle-free graphs are 'almost' bipartite. The proof of this latter result is based on the sparse version of Szemerédi's Regularity Lemma [3].

In this paper we give a short, self-contained proof of the following slightly weaker statement of the second 1-statement of Theorem 1.2.

Theorem 1.3. *There exists a constant $C > 0$ such that*

$$\Pr[T_{n,m} \text{ is bipartite}] \to 1 \qquad \text{if } m \geqslant Cn^{3/2} \sqrt{\log n}.$$

Notation and definitions. To simplify notation we will usually omit the explicit use of the Gauss brackets $\lfloor\ \rfloor$ in denoting the cardinality of vertex sets, but instead tacitly assume them. In this paper $\log n$ denotes the logarithm to base 2. The natural logarithm to base $e = 2.71..$ is denoted by $\ln n$.

The following inequalities are easily seen to be true and will be used hereafter without further reference.

(i) For all $a, b, x \in \mathbb{N}$: $\binom{a}{b} \leqslant \binom{a+x}{b}$, $\binom{a-x}{b-x} \leqslant \binom{a}{b}$, and $\binom{a}{b} \leqslant \left(\frac{ea}{b}\right)^b$.
(ii) For all $a, b, i, k \in \mathbb{N}$: $\binom{a}{i} \cdot \binom{b}{k} \leqslant \binom{a+b}{i+k}$.
(iii) For all $a, b, x \in \mathbb{N}$: $\binom{a-x}{b} \cdot e^{\frac{xb}{a}} \leqslant \binom{a}{b} \leqslant \binom{a-x}{b} \cdot e^{\frac{xb}{a-b-x}}$ and $\binom{a+x}{b} \leqslant \binom{a}{b} \cdot e^{x \ln a}$.
(iv) For all $a, b, x \in \mathbb{N}$: $\binom{a}{b-x} \leqslant \binom{a}{b} \cdot \left(\frac{b-x}{a-b}\right)^x \leqslant \binom{a}{b} \cdot 2^{x \log b}$.

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

2. Proof of Theorem 1.3

In [7] Prömel and Steger obtained precise bounds for the number of bipartite graphs on n vertices with m edges. For the purpose of this paper, however, the following weaker estimates suffice. See [8] for a short proof.

Lemma 2.1. ([8]) *For all m with $m \gg n \log n$ and n sufficiently large,*

$$2^{n-2\log n} \cdot \binom{\frac{1}{4}n^2}{m} \leqslant |\mathscr{S}(n,m)| \leqslant 2^n \cdot \binom{\frac{1}{4}n^2}{m}.$$

The following simple lemma was also proved in [8].

Lemma 2.2. ([8]) *For all $0 \leqslant m \leqslant \lfloor \frac{1}{8}n^2 \rfloor$,*

$$|\mathscr{T}(n,m)| \leqslant e^{2m} \cdot |\mathscr{S}(n,m)|.$$

Let $b \geqslant 1$ and $0 < \epsilon_{\mathscr{A}}, \epsilon_{\mathscr{B}}, \epsilon_{\mathscr{C}} < \frac{1}{100}$ be positive constants such that

$$\frac{40}{\log b} + \log(12e\epsilon_{\mathscr{A}}^{-1})\epsilon_{\mathscr{A}} < \frac{1}{2}, \qquad \frac{1}{2}\epsilon_{\mathscr{A}}\epsilon_{\mathscr{B}} > \frac{1700}{\log b} + \frac{10}{\sqrt{b}}, \tag{2.1}$$
$$\frac{1}{2}\epsilon_{\mathscr{A}}(\epsilon_{\mathscr{C}} - 3\epsilon_{\mathscr{B}}) \geqslant \frac{1700}{\log b} + \frac{4}{\sqrt{b}}, \qquad \frac{1}{1600}\sqrt{b} \geqslant 11.$$

Note that these inequalities are partly redundant. However, these are the requirements which we need for the rest of the proof and these relations should therefore ease reading further on. Note also that we may choose these constants such that $\varepsilon := \max\{\epsilon_{\mathscr{A}}, \epsilon_{\mathscr{B}}, \epsilon_{\mathscr{C}}\}$ is arbitrarily small. Furthermore we set

$$a := \tfrac{8\sqrt{b}}{\log b}, \qquad q = q(n,m) := \frac{1}{2\sqrt{b}}\frac{m}{n\log n}.$$

Before we continue with some definitions we would like to add some words on the general proof strategy. Assume the theorem is true. Then or $m \geqslant Cn^{3/2}\sqrt{\log n}$ most triangle-free graphs should be bipartite. In particular, they should have all the properties of a random bipartite graph. Now, what properties does a random bipartite graph have? For example, it should satisfy the following properties with high probability: both colour classes should have approximately size $\frac{1}{2}n$, every vertex should have degree about $2\frac{m}{n}$, and for every vertex v the neighbourhood $\Gamma(Q)$ of every suitably large subset $Q \subseteq \Gamma(v)$ in the neighbourhood of v should be almost equal to the colour class of v. Our strategy is now to define 'bad' sets, which capture properties that a random bipartite graph does *not* have, then to show (Lemma 2.3) that all graphs not contained in any of theses bad sets are indeed bipartite, and, finally, to show that for $m \geqslant Cn^{3/2}\sqrt{\log n}$ the sizes of these bad sets are negligible compared to the size of $|\mathscr{T}(n,m)|$.

Define bad sets $\mathscr{A}_i(n,m)$, $\mathscr{B}_i(n,m)$, $\mathscr{C}(n,m)$, and $\mathscr{D}(n,m)$ as follows.

$\mathscr{A}_1(n,m) :=$ set of all graphs G in $\mathscr{T}(n,m)$ which contain a vertex v of degree less than $\epsilon_{\mathscr{A}}\frac{m}{n}$.
$\mathscr{A}_2(n,m) :=$ set of all graphs G in $\mathscr{T}(n,m)$ which contain a vertex v of degree greater than $70\frac{m}{n}$.

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

$\mathscr{B}_1(n,m) :=$ set of all graphs G in $\mathscr{T}(n,m) \setminus \bigcup_{i=1,2} \mathscr{A}_i(n,m)$ which contain a vertex v and a set $Q \subseteq \Gamma(v)$ of size $|Q| = 2q$ such that $|\Gamma(Q)| \geqslant (\frac{1}{2} + \epsilon_{\mathscr{B}})n$.

$\mathscr{B}_2(n,m) :=$ set of all graphs G in $\mathscr{T}(n,m) \setminus (\bigcup_{i=1,2} \mathscr{A}_i(n,m) \cup \mathscr{B}_1(n,m))$ which contain a set $Q \subseteq V$ of size $|Q| = q$ such that at least one of the following conditions is satisfied:

(i) $|\Gamma(Q)| \leqslant (\frac{1}{2} - \epsilon_{\mathscr{B}})n$,

(ii) there exists a set $X \subseteq \Gamma(Q)$ of size $|X| = 2\epsilon_{\mathscr{B}} n$ such that all vertices in Q have at most q neighbours in X.

$\mathscr{C}(n,m) :=$ set of all graphs G in $\mathscr{T}(n,m) \setminus (\bigcup_{i=1,2}(\mathscr{A}_i(n,m) \cup \mathscr{B}_i(n,m)))$ which contain a vertex v and sets $Q, Q' \subseteq \Gamma(v)$ of size $|Q| = |Q'| = q$ such that for every vertex $w \in Q'$ there exists a set $Q_w \subseteq \Gamma(w)$ of size $|Q_w| = q$ such that $|\Gamma(Q_w) \cap \Gamma(Q)| \geqslant \epsilon_{\mathscr{C}} n$.

$\mathscr{D}(n,m) :=$ set of all graphs G in $\mathscr{T}(n,m) \setminus (\bigcup_{i=1,2}(\mathscr{A}_i(n,m) \cup \mathscr{B}_i(n,m)) \cup \mathscr{C}(n,m))$ that are not bipartite but for which there exists a bipartition of the vertex set V into two sets V_1 and V_2 such that $|V_1|, |V_2| \geqslant (\frac{1}{2} - \frac{1}{100})n$ and such that the bipartite graph induced by this partition has the following two properties:

(i) it contains at least $m - nq$ edges, and

(ii) every vertex v has at most q neighbours in the same class as v.

Lemma 2.3. *If ε is sufficiently small, then for all sufficiently large n and all $m \geqslant \sqrt{b} n^{3/2} \sqrt{\log n}$,*

$$\mathscr{T}(n,m) \setminus \left(\bigcup_{i=1,2} (\mathscr{A}_i(n,m) \cup \mathscr{B}_i(n,m)) \cup \mathscr{C}(n,m) \cup \mathscr{D}(n,m) \right) \subseteq \mathscr{S}(n,m).$$

Proof. Let $G = (V,E)$ be an arbitrary but fixed graph in $\mathscr{T}(n,m) \setminus (\bigcup_{i=1,2}(\mathscr{A}_i(n,m) \cup \mathscr{B}_i(n,m)) \cup \mathscr{C}(n,m) \cup \mathscr{D}(n,m))$. In order to concentrate on the main ideas of the proof instead of technical details we will use the following two notations for subsets $X, Y \subseteq V$:

$$|X| \cong \frac{1}{2}n \iff \exists c_X \in \mathbb{N} \quad \text{s.t.} \quad \left(\frac{1}{2} - c_X\varepsilon\right)n \leqslant |X| \leqslant \left(\frac{1}{2} + c_X\varepsilon\right)n, \quad \text{and}$$

$$X \cong Y \iff |X| \cong \frac{1}{2}n \wedge |Y| \cong \frac{1}{2}n \wedge |X \cap Y| \cong \frac{1}{2}n.$$

In the following we assume that ε is sufficiently small so that $c_X\varepsilon \leqslant \frac{1}{100}$ for all constants c_X which occur in the proof. The reader is invited to observe that in the rest of the proof the constants c_X can always be chosen independently of $\varepsilon = \max\{\epsilon_{\mathscr{A}}, \epsilon_{\mathscr{B}}, \epsilon_{\mathscr{C}}\}$.

With this notation at hand we are ready to collect some properties of G which follow immediately from the definition of the sets $\mathscr{A}_1(n,m), \ldots, \mathscr{C}(n,m)$.

(1) For every vertex $v \in V$ all sets $Q_v \subseteq \Gamma(v)$ of size $|Q_v| = q$ are such that

- $|\Gamma(Q_v)| \cong \frac{1}{2}n$, and
- every set $X \subseteq \Gamma(v) \setminus Q_v$ of size $|X| = q$ satisfies $\Gamma(X) \cong \Gamma(Q_v)$.

(Proof: This follows immediately from the definition of the sets $\mathscr{B}_1(n,m)$ and $\mathscr{B}_2(n,m)$; for the second part observe that the set Q in the definition of $\mathscr{B}_1(n,m)$ has size $2q$.)

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

In the following we use Q_v to denote an arbitrary but fixed set $Q_v \subseteq \Gamma(v)$ of size $|Q_v| = q$. Note that the existence of these sets is guaranteed by the definition of the set $\mathscr{A}_1(n,m)$.

(2) For every vertex v there exist at most q many vertices $v_1, \dots, v_q \in \Gamma(v)$ that satisfy $|\Gamma(Q_{v_i}) \cap \Gamma(Q_v)| \geqslant \epsilon_{\mathscr{C}} n$.
(Proof: This follows immediately from the definition of the set $\mathscr{C}(n,m)$.)
In the following we use B_v to denote these at most q 'bad' vertices in $\Gamma(v)$.

(3) For every vertex v there exists a set $Y_v \subseteq \Gamma(Q_v)$ such that $|Y_v| \cong |\Gamma(Q_v)| \cong \frac{1}{2}n$ and such that

$$\Gamma(Q_y) \cong \Gamma(Q_v) \qquad \text{for all} \quad y \in Y_v.$$

(Proof: Let $\widetilde{Q}_v$ denote a set of size q within $\Gamma(v) \setminus (B_v \cup Q_v)$. Recall that we know from the first property that $\Gamma(\widetilde{Q}_v) \cong \Gamma(Q_v)$. Consider a vertex w in $\widetilde{Q}_v$. Then, according to the second property, $\Gamma(Q_w) \cong V \setminus \Gamma(Q_v)$. Repeating this argument we deduce that all vertices y in $\Gamma(w) \setminus B_w$ satisfy $\Gamma(Q_y) \cong V \setminus \Gamma(Q_w) \cong \Gamma(Q_v)$. Now consider the set $X := \{x \in \Gamma(\widetilde{Q}_v) \mid \forall w \in \Gamma(x) \cap \widetilde{Q}_v : x \in B_w\}$. Then the definition of X implies that every vertex $w \in \widetilde{Q}_v$ has at most $|B_w| \leqslant q$ neighbours in X. From the definition of the set $\mathscr{B}_2(n,m)$ it therefore follows that $|X| < 2\epsilon_{\mathscr{B}} n$. The claim follows as the definition of X also implies that for every vertex $y \in Y := \Gamma(\widetilde{Q}_v) \setminus X$ there exists a $w \in \widetilde{Q}_v$ such that $y \in \Gamma(w) \setminus B_w$ and hence $\Gamma(Q_y) \cong \Gamma(Q_v)$.)

(4) For every pair of vertices x, y we have

$$\Gamma(Q_x) \cong \Gamma(Q_y) \qquad \text{or} \qquad \Gamma(Q_x) \cong V \setminus \Gamma(Q_y).$$

(Proof: This follows easily from the previous property, as the assumption that $|\Gamma(Q_x) \cap \Gamma(Q_v)| > \mu n$ for an appropriately defined constant $\mu > 0$ implies that there exists a vertex $z \in \Gamma(Q_x) \cap \Gamma(Q_v)$ such that $\Gamma(Q_z) \cong \Gamma(Q_x)$ and $\Gamma(Q_z) \cong \Gamma(Q_y)$.)

Choose an arbitrary vertex v_0. Let $A := \Gamma(Q_{v_0})$ and $B := V \setminus A$. Recall that $|A|, |B| \cong \frac{1}{2}n$. We use these two sets to define a bipartition of the vertex set:

$$V_1 := \{v \in V \mid \Gamma(Q_v) \cong A\},$$
$$V_2 := \{v \in V \mid \Gamma(Q_v) \cong B\}.$$

Observe that according to property (4) this defines indeed a *partition* of the vertex set V. As a last step observe that property (2) implies that, for $i = 1, 2$, every vertex $v \in V_i$ has at most q neighbours in V_i. That is, (V_1, V_2) induces a bipartite graph that contains at least $m - qn$ edges and in which every vertex has at most q neighbours inside its class. As by assumption the graph G is not contained in $\mathscr{D}(n,m)$, this implies that G is indeed bipartite. □

Theorem 2.4. *There exists a constant ξ_0 such that, for all $n \in \mathbb{N}$,*

$$|\mathscr{T}(n,m)| \leqslant (1 + 2^{\xi(n,m)} + 2^{-\frac{1}{800}m^2/n^3}) \cdot |\mathscr{S}(n,m)| \qquad \text{for all} \quad 0 \leqslant m \leqslant \binom{n}{2},$$

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

where

$$\xi(n,m) = an^{3/2}\sqrt{\log n}\cdot\log\frac{bn^{3/2}\sqrt{\log n}}{m} + \xi_0.$$

Observe that Theorem 2.4 implies Theorem 1.3 for, *e.g.*, $C = 2b$. Note also that for $m \ll n^{3/2}\sqrt{\log n}$ the bound on $|\mathscr{T}(n,m)|$ is rather weak. Better bounds can be found in Wormald [9]. Bounds on the chromatic number of almost all graphs in the non-bipartite interval can be found in Osthus, Prömel and Taraz [6].

Proof. Choose n_0 in such a way that all inequalities below hold for all $n \geqslant n_0$. Subsequently choose ξ_0 in such a way that the claim of the theorem holds for all $n \leqslant n_0$. (The reader is invited to observe that the inequalities later used to define n_0 will not depend on ξ_0.)

We proceed by induction on n. So assume the theorem holds for all $n' < n$ for some $n \in \mathbb{N}$. We will show that it then also holds for n. Let m be an arbitrary integer between 1 and $\binom{n}{2}$. Observe that if

$$an^{3/2}\sqrt{\log n}\cdot\log\frac{bn^{3/2}\sqrt{\log n}}{m} > 2\log e\cdot m,$$

then by Lemma 2.2 there is nothing to show. So we may assume without loss of generality that

$$\sqrt{b}n^{3/2}\sqrt{\log n} \leqslant m < \frac{1}{4}n^2 \tag{2.2}$$

(*cf.* the definition of a and the fact at every triangle-free graph with $\frac{1}{4}n^2$ edges is bipartite).

As a first step we convert the bounds on $|\mathscr{T}(n-x,m-y)|$ provided by the induction hypothesis into a somewhat nicer form. First observe that for all $n, x \in \mathbb{N}$ with $x < n$ we have

$$(n-x)^{3/2}\cdot\sqrt{\log(n-x)} \leqslant (n-x)\sqrt{n}\sqrt{\log n} = n^{3/2}\sqrt{\log n} - x\sqrt{n\log n}, \tag{2.3}$$

and that for all $1 \leqslant x \leqslant q$ and $0 \leqslant y \leqslant \frac{1}{2}m$ and n sufficiently large,

$$\begin{aligned}|\mathscr{S}(n-x,m-y)| &\leqslant 2^n\cdot\binom{\frac{1}{4}(n-x)^2}{m-y} \leqslant 2^n\cdot\binom{\frac{1}{4}n^2-\frac{1}{2}nx}{m-y}\cdot e^{\frac{1}{4}x^2\ln(\frac{1}{4}n^2)}\\ &\leqslant 2^{n+x^2\log n}\cdot\binom{\frac{1}{4}n^2-\frac{1}{2}nx}{m-y} \leqslant 2^{n+\frac{x}{\log b}\frac{m}{n}}\cdot\binom{\frac{1}{4}n^2-\frac{1}{2}nx}{m-y}\end{aligned} \tag{2.4}$$

(*cf.* Lemma 2.1 and the definition of q) and that

$$\begin{aligned}\xi(n-x,m-y) &\leqslant \xi(n,m) + ax\sqrt{n\log n}\cdot\log\frac{m}{bn^{3/2}\sqrt{\log n}} + \frac{3ayn^{3/2}\sqrt{\log n}}{m}\\ &\leqslant \xi(n,m) + \frac{8x}{\log b}\frac{m}{n} + \frac{24y}{\log b}.\end{aligned}$$

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

This follows from equations (2.2) and (2.3), the definition of a, and the observation that, for all $y \leqslant \frac{1}{2}m$,

$$\begin{aligned}\log \frac{b(n-x)^{3/2}\sqrt{\log(n-x)}}{m-y} &\leqslant \log\left(\frac{bn^{3/2}\sqrt{\log n}}{m}\cdot\left(1+\frac{y}{m-y}\right)\right)\\ &\leqslant \log\left(\frac{bn^{3/2}\sqrt{\log n}}{m}\cdot\left(1+\frac{2y}{m}\right)\right)\\ &= \log\frac{bn^{3/2}\sqrt{\log n}}{m}+\log\left(1+\frac{2y}{m}\right)\\ &\leqslant \log\frac{bn^{3/2}\sqrt{\log n}}{m}+\frac{3y}{m}\end{aligned}$$

(to see the last inequality use $\log(1+x) \leqslant \frac{x}{\ln 2} \leqslant \frac{3}{2}x$). We therefore conclude that the induction hypothesis and (2.4) imply that, for all $1 \leqslant x \leqslant q$ and $0 \leqslant y \leqslant \frac{1}{2}m$,

$$\begin{aligned}|\mathscr{T}(n-x,m-y)| &\leqslant (1+2^{\xi(n,m)+\frac{8x}{\log b}\frac{m}{n}+\frac{24y}{\log b}}+2^{-\frac{1}{800}\frac{(m-y)^2}{(n-x)^3}})\cdot 2^{n+\frac{x}{\log b}\frac{m}{n}}\binom{\frac{1}{4}n^2-\frac{1}{2}xn}{m-y}\\ &\leqslant 2^{\xi(n,m)+\frac{10x}{\log b}\frac{m}{n}+\frac{24y}{\log b}}\cdot 2^n\cdot\binom{\frac{1}{4}n^2-\frac{1}{2}nx}{m-y}. \qquad (2.5)\end{aligned}$$

We also note that whenever $y \leqslant 70x\frac{m}{n}$, the following slightly weaker but shorter bound also holds:

$$|\mathscr{T}(n-x,m-y)| \leqslant 2^{\xi(n,m)+\frac{1700}{\log b}x\frac{m}{n}}\cdot 2^n\cdot\binom{\frac{1}{4}n^2-\frac{1}{2}nx}{m-y}. \qquad (2.6)$$

These last two inequalities will be heavily used throughout the upcoming parts of the proof. Before we do so we first recall that, by Lemma 2.3,

$$|\mathscr{T}(n,m)| \leqslant \sum_{i=1,2}(|\mathscr{A}_i(n,m)|+|\mathscr{B}_i(n,m)|)+|\mathscr{C}(n,m)|+|\mathscr{D}(n,m)|+|\mathscr{S}(n,m)|.$$

It hence suffices to show that all summands on the right-hand side except the last one are smaller than or equal to

$$\frac{1}{6}\cdot\left(2^{\xi(n,m)}+2^{-\frac{1}{800}\frac{m^2}{n^3}}\right)\cdot|\mathscr{S}(n,m)|. \qquad (2.7)$$

Moreover, by Lemma 2.1 it also suffices to show that each of these summands is smaller than or equal to

$$\frac{1}{6}\cdot 2^{\xi(n,m)}\cdot 2^{n-2\log n}\cdot\binom{\frac{1}{4}n^2}{m}. \qquad (2.8)$$

This is what we will do now. (Recall that we assumed that n is sufficiently large, so that all inequalities below do hold.)

The sets $\mathscr{A}(n,m)$. All graphs in $\mathscr{A}_1(n,m)\cup\mathscr{A}_2(n,m)$ can be constructed as follows. First choose the vertex v (at most n choices). For all integers y such that $0 \leqslant y < \epsilon_{\mathscr{A}}\frac{m}{n}$ or $70\frac{m}{n} < y < n$, then choose a triangle-free graph with $m-y$ edges on the remaining vertices ($|\mathscr{T}(n-1,m-y)|$ ways). Finally, choose the y neighbours of v (at most $\binom{n-1}{y}$

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

ways). Together, this gives

$$|\mathscr{A}_1(n,m)| \leqslant 2^{\log n} \cdot \sum_{y=0}^{\epsilon_{\mathscr{A}}\frac{m}{n}} \binom{n-1}{y} |\mathscr{T}(n-1,m-y)|$$
$$\leqslant 2^{2\log n} \cdot \max_{y \leqslant \epsilon_{\mathscr{A}}\frac{m}{n}} \binom{n}{y} |\mathscr{T}(n-1,m-y)|,$$

respectively

$$|\mathscr{A}_2(n,m)| \leqslant 2^{\log n} \cdot \sum_{y=70\frac{m}{n}}^{n-1} \binom{n-1}{y} |\mathscr{T}(n-1,m-y)|$$
$$\leqslant 2^{2\log n} \cdot \max_{y \geqslant 70\frac{m}{n}} \binom{n}{y} |\mathscr{T}(n-1,m-y)|.$$

Using the bound on $|\mathscr{T}(n-1,m-y)|$ from (2.5), we deduce

$$|\mathscr{A}_i(n,m)| \leqslant 2^{2\log n} \cdot \max_{y \leqslant \epsilon_{\mathscr{A}}\frac{m}{n} | y \geqslant 70\frac{m}{n}} \binom{n}{y} \binom{\frac{1}{4}n^2 - \frac{1}{2}n}{m-y} \cdot 2^n \cdot 2^{\xi(n,m) + \frac{10}{\log b}\frac{m}{n} + \frac{24}{\log b}y}.$$

First we consider the case $m \leqslant \frac{1}{6}n^2$. As we assumed that $m \geqslant \sqrt{b}n^{3/2}\sqrt{\log n}$ we know that, with room to spare, $y < n \leqslant \frac{1}{3}m$ and the binomial coefficients can therefore be bounded as follows:

$$\binom{n}{y}\binom{\frac{1}{4}n^2 - \frac{1}{2}n}{m-y} \leqslant \left(\frac{en}{y}\right)^y \cdot \binom{\frac{1}{4}n^2}{m-y} \cdot e^{-\frac{1}{2}n \cdot \frac{m-y}{\frac{1}{4}n^2}}$$
$$\leqslant \left(\frac{en}{y}\right)^y \cdot \binom{\frac{1}{4}n^2}{m} \left(\frac{m-y}{\frac{1}{4}n^2 - m}\right)^y \cdot e^{-\frac{m}{n}}$$
$$\leqslant \binom{\frac{1}{4}n^2}{m} \cdot \left(\frac{12em}{yn}\right)^y \cdot e^{-\frac{m}{n}}.$$

For $y \geqslant 70\frac{m}{n}$ we have

$$\left(\frac{12em}{yn}\right)^y \leqslant 2^{-y}.$$

For $y \leqslant \epsilon_{\mathscr{A}}\frac{m}{n}$ we observe that the maximum is attained at $y = \epsilon_{\mathscr{A}}\frac{m}{n}$, and that therefore

$$\left(\frac{12em}{yn}\right)^y \leqslant 2^{\log\frac{12e}{\epsilon_{\mathscr{A}}} \cdot \epsilon_{\mathscr{A}}\frac{m}{n}}.$$

In both cases we easily check that the definitions in (2.1) imply that we indeed have that the cardinalities of the sets $\mathscr{A}_i(n,m)$ are smaller than the term in (2.8).

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

Now consider the case $m > \frac{1}{6}n^2$. For $y < \epsilon_{\mathscr{A}} \frac{m}{n} \leqslant \frac{1}{4}\epsilon_{\mathscr{A}} n$ (as $m \leqslant \frac{1}{4}n^2$ for triangle-free graphs) we bound the binomial coefficients as follows:

$$\begin{aligned}\binom{n}{y}\binom{\frac{1}{4}n^2 - \frac{1}{2}n}{m-y} &\leqslant \left(\frac{en}{y}\right)^y \cdot \binom{\frac{1}{4}n^2 - (\frac{1}{2}n - y)}{m} \\ &\leqslant \left(\frac{en}{y}\right)^y \cdot \binom{\frac{1}{4}n^2}{m} e^{-(\frac{1}{2}n-y)\frac{m}{\frac{1}{4}n^2}} \\ &\leqslant \binom{\frac{1}{4}n^2}{m} \cdot \left(\frac{en}{y}\right)^y \cdot e^{-\frac{m}{n}}.\end{aligned}$$

Again we easily check that the maximum is attained at $y = \epsilon_{\mathscr{A}} \frac{m}{n} > \frac{1}{6}\epsilon_{\mathscr{A}} n$ (recall that we assumed $m > \frac{1}{6}n^2$), and that therefore

$$\left(\frac{en}{y}\right)^y \leqslant (6e/\epsilon_{\mathscr{A}})^{\epsilon_{\mathscr{A}} \frac{m}{n}} \leqslant 2^{\log \frac{6e}{\epsilon_{\mathscr{A}}} \cdot \epsilon_{\mathscr{A}} \frac{m}{n}}.$$

As above, the definitions in (2.1) therefore imply that the cardinality of the set $\mathscr{A}_1(n,m)$ is smaller than the term in (2.8). Finally, observe that for $m > \frac{1}{6}n^2$ we have that $70\frac{m}{n} > n$ and hence $|\mathscr{A}_2(n,m)| = 0$.

The set $\mathscr{B}_1(n,m)$. All graphs in $\mathscr{B}_1(n,m)$ can be constructed as follows. First choose the vertex v (at most n ways). Let y denote the degree of v. Note that $\epsilon_{\mathscr{A}} \frac{m}{n} \leqslant y \leqslant 70\frac{m}{n}$. For all possible values choose a triangle-free graph with $m - y$ edges on the remaining vertices ($|\mathscr{T}(n-1, m-y)|$ ways). Next choose an appropriate set Q of size $2q$ (at most $\binom{n-1}{2q} \leqslant 2^{2q \log n}$ ways). As v cannot be connected to any vertex in $\Gamma(Q)$ (this would close a cycle of length 3) there are at most $\binom{n-|\Gamma(Q)|}{y}$ ways to choose the remaining neighbours of v. Together we obtain

$$\begin{aligned}|B_1(n,m)| &\leqslant 2^{\log n + 2q \log n} \cdot \sum_{\epsilon_{\mathscr{A}} \frac{m}{n} \leqslant y \leqslant 70\frac{m}{n}} \binom{n - |\Gamma(Q)|}{y} |\mathscr{T}(n-1, m-y)| \\ &\leqslant 2^{3q \log n} \cdot \max_{\epsilon_{\mathscr{A}} \frac{m}{n} \leqslant y \leqslant 70\frac{m}{n}} \binom{(\frac{1}{2} - \epsilon_{\mathscr{B}})n}{y} |\mathscr{T}(n-1, m-y)|.\end{aligned}$$

We now use the bound on $|\mathscr{T}(n-x, m-y)|$ from (2.6) for $x = 1$ and y to obtain

$$|\mathscr{B}_1(n,m)| \leqslant 2^{3q \log n} \cdot \max_{\epsilon_{\mathscr{A}} \frac{m}{n} \leqslant y \leqslant 70\frac{m}{n}} \binom{(\frac{1}{2} - \epsilon_{\mathscr{B}})n}{y} \cdot \binom{\frac{1}{4}n^2 - \frac{1}{2}n}{m-y} \cdot 2^n \cdot 2^{\xi(n,m) + \frac{1700}{\log b}\frac{m}{n}}.$$

We bound the binomial coefficients:

$$\begin{aligned}\binom{(\frac{1}{2} - \epsilon_{\mathscr{B}})n}{y}\binom{\frac{1}{4}n^2 - \frac{1}{2}n}{m-y} &\leqslant \binom{\frac{1}{4}n^2 - \epsilon_{\mathscr{B}} n}{m} \\ &\leqslant \binom{\frac{1}{4}n^2}{m} \cdot e^{-\epsilon_{\mathscr{B}} n \cdot \frac{4m}{n^2}} = \binom{\frac{1}{4}n^2}{m} \cdot e^{-4\epsilon_{\mathscr{B}} \frac{m}{n}}.\end{aligned}$$

Using the definitions (2.1) we easily deduce that (2.8) is an upper bound for $|\mathscr{B}_1(n,m)|$.

The set $\mathscr{B}_2(n,m)$. All graphs in $\mathscr{B}_2(n,m)$ can be constructed as follows. First choose the set Q (less than 2^n ways). Let e_q denote the number of edges inside Q and let e_r denote the

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

number of edges connecting Q with $\Gamma(Q)$. As the degree of every vertex in Q is between $\epsilon_{\mathscr{A}}\frac{m}{n}$ and $70\frac{m}{n}$, we know that $e_r \geqslant q\cdot(\epsilon_{\mathscr{A}}\frac{m}{n}-q) \geqslant \frac{1}{2}\epsilon_{\mathscr{A}} q\frac{m}{n}$ and $e_q + e_r \leqslant 70q\frac{m}{n}$. For all possible values choose a triangle-free graph with $m - e_q - e_r$ edges on the remaining vertices ($|\mathscr{T}(n-q, m-e_q-e_r)|$ ways). Next choose the set $\Gamma(Q)$ (less than 2^n ways) and the edges inside the set Q (less than $\binom{\frac{1}{2}q^2}{e_q} \leqslant 2^{q^2}$ ways).

In order to choose the edges between Q and $\Gamma(Q)$ we distinguish between the two properties in the definition of the set $\mathscr{B}_2(n,m)$.

If property (i) is satisfied then there are clearly only

$$\begin{aligned}\binom{q\cdot(\frac{1}{2}-\epsilon_{\mathscr{B}})n}{e_r} &\leqslant e^{-\epsilon_{\mathscr{B}} qn\frac{e_r}{\frac{1}{2}qn}}\cdot\binom{\frac{1}{2}qn}{e_r} = e^{-2\epsilon_{\mathscr{B}} e_r}\cdot\binom{\frac{1}{2}qn}{e_r} \leqslant e^{-\epsilon_{\mathscr{A}}\epsilon_{\mathscr{B}} q\frac{m}{n}}\cdot\binom{\frac{1}{2}qn}{e_r}\\ &\leqslant e^{-\epsilon_{\mathscr{A}}\epsilon_{\mathscr{B}} q\frac{m}{n}}\cdot\binom{\frac{1}{2}qn}{e_q+e_r}\cdot 2^{2e_q\log n} \leqslant e^{-\epsilon_{\mathscr{A}}\epsilon_{\mathscr{B}} q\frac{m}{n}}\cdot 2^{q^2\log n}\cdot\binom{\frac{1}{2}qn}{e_q+e_r}\end{aligned}$$

ways to choose the edges between Q and $\Gamma(Q)$.

If property (ii) is satisfied we argue as follows. First we choose the set X (2^n ways). Observe that there are at most

$$\left(\sum_{i=0}^{q}\binom{|X|}{i}\right)^{|Q|} \leqslant 2^{q|Q|\log n} = 2^{q^2\log n}$$

ways to connect Q to X. By the definition of $\mathscr{B}_1(n,m)$ we know that $|\Gamma(Q)| \leqslant (\frac{1}{2}+\epsilon_{\mathscr{B}})n$. Hence, the number of ways to connect Q to $\Gamma(Q)\setminus X$ can be bounded by

$$\begin{aligned}\sum_{i=e_r-q^2}^{e_r}\binom{|Q|\cdot|\Gamma(Q)\setminus X|}{i} &\leqslant \binom{|Q|\cdot|\Gamma(Q)\setminus X|}{e_r}\cdot 2^{q^2\log(qn)} \leqslant \binom{\frac{1}{2}qn-\epsilon_{\mathscr{B}} qn}{e_r}\cdot 2^{2q^2\log n}\\ &\leqslant e^{-2\epsilon_{\mathscr{B}} e_r}\cdot 2^{2q^2\log n}\cdot\binom{\frac{1}{2}qn}{e_r} \leqslant e^{-2\epsilon_{\mathscr{B}} e_r}\cdot 2^{3q^2\log n}\cdot\binom{\frac{1}{2}qn}{e_q+e_r}.\end{aligned}$$

Summarizing, we conclude that if property (ii) is satisfied the number of ways to choose the edges between Q and $\Gamma(Q)$ is bounded by

$$\begin{aligned}2^{n+4q^2\log n}\cdot e^{-2\epsilon_{\mathscr{B}} e_r}\cdot\binom{\frac{1}{2}qn}{e_q+e_r} &\leqslant 2^{n+4q^2\log n}\cdot e^{-2\epsilon_{\mathscr{B}} e_r}\cdot\binom{\frac{1}{2}qn}{e_q+e_r}\\ &\leqslant 2^{n+4q^2\log n}\cdot e^{-\epsilon_{\mathscr{A}}\epsilon_{\mathscr{B}} q\frac{m}{n}}\cdot\binom{\frac{1}{2}qn}{e_q+e_r}.\end{aligned}$$

(Here we used (2.1) and the fact that $e_r \geqslant \frac{1}{2}\epsilon_{\mathscr{A}} q\frac{m}{n}$.)

Together we obtain

$$\begin{aligned}|B_2(n,m)| &\leqslant 2^{2n+q^2}\cdot\sum_{\frac{1}{2}\epsilon_{\mathscr{A}} q\frac{m}{n}\leqslant e_q+e_r\leqslant 70q\frac{m}{n}} 2\cdot 2^{n+4q^2\log n}\cdot e^{-\epsilon_{\mathscr{A}}\epsilon_{\mathscr{B}} q\frac{m}{n}}\cdot\\ &\qquad\qquad\binom{\frac{1}{2}qn}{e_q+e_r}|\mathscr{T}(n-q,m-e_q-e_r)|\\ &\leqslant 2^{20q^2\log n}\cdot\max_{\frac{1}{2}\epsilon_{\mathscr{A}} q\frac{m}{n}\leqslant e_q+e_r\leqslant 70q\frac{m}{n}} e^{-\epsilon_{\mathscr{A}}\epsilon_{\mathscr{B}} q\frac{m}{n}}\cdot\binom{\frac{1}{2}qn}{e_q+e_r}|\mathscr{T}(n-q,m-e_q-e_r)|.\end{aligned}$$

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

(Note that $m \geqslant \sqrt{b}n^{3/2}\sqrt{\log n}$ and the definition of q imply that $q^2 \log n \geqslant \frac{1}{4}n$.) We again use the bound on $|\mathscr{T}(n-x, m-y)|$ from (2.6), this time for $x = q$ and $y = e_q + e_r$ to obtain (note that $y \leqslant \frac{1}{2}m$ follows immediately from the definition of b and q the fact that, trivially, $m \leqslant n^2$)

$$|\mathscr{B}_2(n,m)| \leqslant 2^{20x^2 \log n} \cdot \max_{\frac{1}{2}\epsilon_{\mathscr{A}} x\frac{m}{n} \leqslant y \leqslant 70x\frac{m}{n}} e^{-\epsilon_{\mathscr{A}}\epsilon_{\mathscr{B}} x\frac{m}{n}} \cdot \binom{\frac{1}{2}xn}{y}\binom{\frac{1}{4}n^2 - \frac{1}{2}nx}{m-y} \cdot 2^n \cdot 2^{\xi(n,m)+\frac{1700}{\log b}x\frac{m}{n}}$$
$$\leqslant 2^{\frac{10}{\sqrt{b}}x\frac{m}{n}} \cdot e^{-\epsilon_{\mathscr{A}}\epsilon_{\mathscr{B}} x\frac{m}{n}} \cdot \binom{\frac{1}{4}n^2}{m} \cdot 2^n \cdot 2^{\xi(n,m)+\frac{1700}{\log b}x\frac{m}{n}}.$$

Using the definitions from (2.1) we again easily check that (2.8) is an upper bound for $|\mathscr{B}_2(n,m)|$.

The set $\mathscr{C}(n,m)$. All graphs in $\mathscr{C}(n,m)$ can be constructed as follows. First choose the set Q' (less than n^q ways). Let e_q denote the number of edges inside Q' and let e_r denote the number of edges between Q' and $\Gamma(Q')$. Note that we may assume that $e_r \geqslant q \cdot (\epsilon_{\mathscr{A}}\frac{m}{n} - q) \geqslant \frac{1}{2}\epsilon_{\mathscr{A}} q\frac{m}{n}$ and $e_q + e_r \leqslant 70q\frac{m}{n}$. For all possible values choose a triangle-free graph with $m - e_q - e_r$ edges on the remaining vertices ($|\mathscr{T}(n-q, m-e_q-e_r)|$ ways). Next choose an appropriate vertex v and an appropriate set $Q \subseteq \Gamma(v)$ of size q (less than $n \cdot n^q = n^{q+1}$ ways).

The number of ways to choose the edges inside the set Q' is bounded by 2^{q^2}. In order to bound the number of ways to connect Q' to $\Gamma(Q')$ we first choose the set $\Gamma(Q')$ (in at most 2^n ways). Note that by the definition of $\mathscr{B}_1(n,m)$ and $\mathscr{B}_2(n,m)$ we only have to consider sets $\Gamma(Q')$ such that

$$|\Gamma(Q')| \leqslant \left(\frac{1}{2} + \epsilon_{\mathscr{B}}\right)n \qquad \text{and} \qquad |\Gamma(Q) \setminus \Gamma(Q')| \leqslant 2\epsilon_{\mathscr{B}} n.$$

For every vertex $w \in Q'$ we first choose the degree d_w and the set Q_w (in at most n^{q+1} ways) such that $|\Gamma(Q_w) \cap \Gamma(Q)| \geqslant \epsilon_{\mathscr{C}} n$ and hence $|\Gamma(Q') \setminus \Gamma(Q_w)| \leqslant (\frac{1}{2} - \epsilon_{\mathscr{C}} + 3\epsilon_{\mathscr{B}})n$. Next we connect w to $\Gamma(Q')$. Note that w may not be connected to a vertex in $\Gamma(Q_w)$. Hence, the number of choices is bounded by

$$\binom{|\Gamma(Q') \setminus \Gamma(Q_w)|}{d_w} \leqslant \binom{(\frac{1}{2} - \epsilon_{\mathscr{C}} + 3\epsilon_{\mathscr{B}})n}{d_w}.$$

Hence, the number of ways to connect Q' to $\Gamma(Q')$ is bounded by

$$\prod_{w \in Q'} n^{q+1}\binom{(\frac{1}{2} - \epsilon_{\mathscr{C}} + 3\epsilon_{\mathscr{B}})n}{d_w} \leqslant n^{q(q+1)} \cdot \binom{(\frac{1}{2} - \epsilon_{\mathscr{C}} + 3\epsilon_{\mathscr{B}})qn}{e_r}$$
$$\leqslant n^{q(q+1)} \cdot e^{-2(\epsilon_{\mathscr{C}} - 3\epsilon_{\mathscr{B}})e_r} \cdot \binom{\frac{1}{2}qn}{e_r}$$
$$\leqslant n^{q(q+1)} \cdot e^{-\epsilon_{\mathscr{A}}(\epsilon_{\mathscr{C}} - 3\epsilon_{\mathscr{B}})q\frac{m}{n}} \cdot 2^{q^2 \log n}\binom{\frac{1}{2}qn}{e_q + e_r},$$

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

as $e_q \leqslant \frac{1}{2}q^2$. That is, we obtain

$$|\mathscr{C}(n,m)| \leqslant \sum_{\frac{1}{2}\epsilon_{\mathscr{A}} q\frac{m}{n} \leqslant e_q+e_r \leqslant 70q\frac{m}{n}} n^{q^2+3q+1} \cdot 2^{q^2+n+q^2\log n} \cdot e^{-\epsilon_{\mathscr{A}}(\epsilon_{\mathscr{C}}-3\epsilon_{\mathscr{B}})q\frac{m}{n}} \cdot \binom{\frac{1}{2}qn}{e_q+e_r} \cdot |\mathscr{T}(n-q, m-e_q-e_r)|$$

$$\leqslant \max_{\frac{1}{2}\epsilon_{\mathscr{A}} q\frac{m}{n} \leqslant e_q+e_r \leqslant 70q\frac{m}{n}} \cdot 2^{8q^2\log n} \cdot e^{-\epsilon_{\mathscr{A}}(\epsilon_{\mathscr{C}}-3\epsilon_{\mathscr{B}})q\frac{m}{n}} \cdot \binom{\frac{1}{2}qn}{e_q+e_r} \cdot |\mathscr{T}(n-q, m-e_q-e_r)|.$$

(Here we again used that $q^2 \log n \geqslant \frac{1}{4}n$ for $m \geqslant \sqrt{b}n^{3/2}\sqrt{\log n}$.)

We again use the bound on $|\mathscr{T}(n-x, m-y)|$, this time for $x = q$ and $y = e_q + e_r$, to obtain

$$|\mathscr{C}(n,m)| \leqslant 2^{8x^2\log n} \cdot \max_{\frac{1}{2}\epsilon_{\mathscr{A}} x\frac{m}{n} \leqslant y < 70x\frac{m}{n}} e^{-\epsilon_{\mathscr{A}}(\epsilon_{\mathscr{C}}-3\epsilon_{\mathscr{B}})x\frac{m}{n}} \cdot \binom{\frac{1}{2}xn}{y}\binom{\frac{1}{4}n^2 - \frac{1}{2}nx}{m-y} \cdot 2^n \cdot 2^{\xi(n,m)+\frac{1700x}{\log b}\frac{m}{n}}$$

$$\leqslant 2^{\frac{4}{\sqrt{b}}x\frac{m}{n}} \cdot \max_{\frac{1}{2}\epsilon_{\mathscr{A}} x\frac{m}{n} \leqslant y < 70x\frac{m}{n}} e^{-\epsilon_{\mathscr{A}}(\epsilon_{\mathscr{C}}-3\epsilon_{\mathscr{B}})x\frac{m}{n}} \cdot \binom{\frac{1}{4}n^2}{m} \cdot 2^n \cdot 2^{\xi(n,m)+\frac{1700x}{\log b}\frac{m}{n}}.$$

Using the definitions (2.1) we again easily observe that (2.8) is an upper bound for $|\mathscr{C}(n,m)|$.

The set $\mathscr{D}(n,m)$. All graphs in $\mathscr{D}(n,m)$ can be constructed as follows. First we choose an appropriate bipartition $V = V_1 \cup V_2$ (less than 2^n ways). Let e_v denote the number of edges inside the sets V_i. For the moment we will consider e_v as fixed. At the end we will sum over all possible values $1 \leqslant e_v \leqslant nq$. There are at most

$$\binom{\binom{|V_1|}{2} + \binom{|V_2|}{2}}{e_v} \leqslant \binom{\frac{1}{2}n^2}{e_v} \leqslant e^{2e_v \ln n}$$

ways to choose these edges. For further use we also note that we may assume without loss of generality that at least $\frac{1}{2}e_v$ of these edges are contained inside the set V_1.

We now have to choose a bipartite graph with $m - e_v$ vertices on the vertex set $V_1 \cup V_2$, in such a way that no triangle is generated.

In order to do that we first consider a random bipartite graph on $V_1 \cup V_2$ with edge probability $p = (m - e_v)/(|V_1||V_2|)$. Note that the assumptions on the size of e_v and $|V_i|$ imply that $\frac{1}{5}\frac{m}{n^2} \leqslant p \leqslant \frac{1}{3}\frac{m}{n^2}$. For every combination of an edge $\{u,v\}$ inside V_1 and a vertex $w \in V_2$ (*i.e.*, for every possible triangle with an edge inside V_1) we construct an indicator random variable I_α, $\alpha = \{\{u,w\},\{v,w\}\}$, which indicates whether the triangle is present. Let $\mathscr{S}$ denote the set of all such events. Note that

$$\mathbb{E}[I_\alpha] = \mathbb{P}[I_\alpha = 1] = p^2 \qquad \text{for all } \alpha \in \mathscr{S}$$

and that

$$\mathbb{E}[I_\alpha I_{\alpha'}] = \mathbb{P}[I_\alpha = 1 \cap I_{\alpha'} = 1] = p^3 \qquad \text{for all } \alpha \neq \alpha' \in \mathscr{S} \text{ s.t. } \alpha \cap \alpha' \neq \emptyset.$$

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

We intend to bound the probability that $X = \sum_\alpha I_\alpha$ is equal to zero. Using the large deviation result from Janson, Łuczak and Ruciński [2] we obtain

$$\mathbb{P}[X=0] \leqslant \begin{cases} e^{-\frac{1}{2}\mu}, & \text{if } \mu \geqslant \Delta \\ e^{-\mu^2/(2\Delta)}, & \text{if } \mu \leqslant \Delta \end{cases} \leqslant e^{-\min\{\frac{1}{2}\mu, \frac{\mu^2}{2\Delta}\}}$$

where

$$\mu = \mathbb{E}[X] \qquad \text{and} \qquad \Delta = \sum_{\substack{\alpha \neq \alpha' \\ \alpha \cap \alpha' \neq \emptyset}} \mathbb{E}[I_\alpha I_{\alpha'}].$$

We now bound μ and Δ:

$$\mu = \mathbb{E}[X] = |\mathscr{S}| \cdot p^2 \geqslant \frac{1}{2} e_v \cdot \left(\frac{1}{2} - \frac{1}{100}\right) n \cdot p^2 \geqslant \frac{1}{200} e_v \cdot \frac{m^2}{n^3}$$

and

$$\begin{aligned} \Delta = \sum_{\substack{\alpha \neq \alpha' \\ \alpha \cap \alpha' \neq \emptyset}} \mathbb{E}[I_\alpha I_{\alpha'}] &\leqslant \sum_{v \in V_1} \binom{\deg_{V_1}(v)}{2} \cdot |V_2| \cdot p^3 \\ &\leqslant \frac{1}{2} \min\{e_v^2, e_v q\} \cdot n \cdot p^3 \leqslant \frac{1}{50} e_v \min\{e_v, q\} \cdot \frac{m^3}{n^5}. \end{aligned}$$

(Note that we used here that the degree $\deg_{V_1}(v)$ of every vertex $v \in V_1$ inside the set V_1 is at most q and that $\sum_{v \in V_1} \deg_{V_1}(v) \leqslant e_v$.) Hence we have that

$$\frac{1}{2}\mu \geqslant \frac{1}{400} e_v \frac{m^2}{n^3} \geqslant \frac{1}{800} e_v \frac{m^2}{n^3} + \frac{1}{800} \frac{m^2}{n^3} \geqslant 11 e_v \ln n + \frac{1}{800} \frac{m^2}{n^3}$$

(recall that $m \geqslant \sqrt{b} n^{3/2} \sqrt{\log n}$), and that

$$\frac{\mu^2}{2\Delta} \geqslant \frac{1}{1600} \frac{e_v}{\min\{e_v, q\}} \frac{m}{n} \geqslant \frac{2\sqrt{b}}{3200} e_v \log n + \frac{1}{3200} \frac{m}{n} \geqslant 11 e_v \ln n + \frac{1}{800} \frac{m^2}{n^3}$$

(recall that $q = \frac{1}{2\sqrt{b}} \frac{m}{n \log n}$ and $m \leqslant n^2/4$). Hence we obtain

$$\mathbb{P}[X=0] \leqslant e^{-\frac{1}{800}\frac{m^2}{n^3} - 11 e_v \ln n}.$$

The probability that such a random bipartite graph G has exactly $m - e_v$ edges is certainly at least $1/n^2$. That is, we have

$$\mathbb{P}[X = 0 \mid E(G) = m - e_v] \leqslant \frac{\mathbb{P}[X=0]}{\mathbb{P}[E(G) = m - e_v]} \leqslant e^{-\frac{1}{800}\frac{m^2}{n^3} - 9 e_v \ln n}.$$

Together, we obtain

$$\begin{aligned} |\mathscr{D}(n,m)| &\leqslant 2^n \cdot \sum_{1 \leqslant e_v \leqslant nq} e^{2 e_v \ln n} \cdot e^{-\frac{1}{800}\frac{m^2}{n^3} - 9 e_v \ln n} \cdot \binom{\frac{1}{4}n^2}{m - e_v} \\ &\leqslant 2^n \cdot n^2 \cdot \max_{1 \leqslant e_v \leqslant nq} e^{-\frac{1}{800}\frac{m^2}{n^3} - 7 e_v \ln n} \cdot \binom{\frac{1}{4}n^2}{m} \cdot e^{2 e_v \ln n} \\ &\leqslant \max_{1 \leqslant e_v \leqslant nq} e^{-e_v \ln n - \frac{1}{800}\frac{m^2}{n^3}} \cdot |\mathscr{S}(n,m)| = e^{-\ln n - \frac{1}{800}\frac{m^2}{n^3}} \cdot |\mathscr{S}(n,m)|, \end{aligned}$$

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press

which easily implies that (2.7) is an upper bound for $|\mathscr{D}(n,m)|$. □

References

[1] Erdős, P., Kleitman, D. J. and Rothschild, B. L. (1976) Asymptotic enumeration of K_n-free graphs. In *International Colloquium on Combinatorial Theory*, *Atti dei Convegni Lincei* **17**, Vol. 2, Rome, pp. 19–27.
[2] Janson, S., Łuczak, T. and Ruciński, A. (1990) An exponential bound for the probability of nonexistence of a specified subgraph in a random graph. In *Random Graphs '87* (M. Karoński, J. Jaworski and A. Ruciński, eds), Wiley, Chichester, pp. 73–87.
[3] Kohayakawa, Y. (1997) Szemerédi's regularity lemma for sparse graphs. In *Foundations of Computational Mathematics*, Springer, Berlin/Heidelberg, pp. 216–230.
[4] Łuczak, T. (2000) On triangle-free random graphs. *Random Struct. Alg.* **16** 260–276.
[5] Osthus, D. Prömel, H. J. and Taraz, A. (2003) For which densities are random triangle-free graphs almost surely bipartite? *Combinatorica* **23** 201–250.
[6] Osthus, D. Prömel, H. J. and Taraz, A. (2001) Almost all graphs of high girth and suitable density have high chromatic number. *J. Graph Theory* **37** 220–226.
[7] Prömel, H. J. and Steger, A. (1995) Random l-colorable graphs. *Random Struct. Alg.* **6** 21–37.
[8] Prömel, H. J. and Steger, A. (1996) On the asymptotic structure of sparse triangle-free graphs. *J. Graph Theory* **21** 137–151.
[9] Wormald, N. (1996) The perturbation method and triangle-free graphs. *Random Struct. Alg.* **9** 253–270.

https://doi.org/10.1017/S0963548304006613 Published online by Cambridge University Press