Combinatorics, Probability and Computing (2009) 18, 749–764. © 2009 Cambridge University Press
doi:10.1017/S096354830900995X Printed in the United Kingdom

On the Number of 2-SAT Functions

L. ILINCA and J. KAHN†

Department of Mathematics, Rutgers University, Piscataway, NJ 08854, USA
(e-mail: {ilinca|jkahn}@math.rutgers.edu)

Received 20 May 2008; revised 10 February 2009; first published online 1 June 2009

We give an alternative proof of a conjecture of Bollobás, Brightwell and Leader, first proved by Peter Allen, stating that the number of Boolean functions definable by 2-SAT formulae is $(1+o(1))2^{\binom{n+1}{2}}$. One step in the proof determines the asymptotics of the number of 'odd-blue-triangle-free' graphs on n vertices.

1. Introduction

Let $\{x_1,\ldots,x_n\}$ be a collection of Boolean variables. Each variable x is associated with a *positive* literal, x, and a *negative* literal, $\bar{x}$. Recall that a *k-SAT formula* is an expression of the form

$$C_1 \wedge \cdots \wedge C_t, \tag{1.1}$$

with t a positive integer and each C_i a *k-clause*; that is, an expression $y_1 \vee \cdots \vee y_k$, with $y_1,\ldots,y_k$ literals corresponding to different variables. A formula (1.1) defines a Boolean function of $x_1,\ldots,x_n$ in the obvious way; any such function is a *k-SAT function.* Here we will be concerned almost exclusively with the case $k=2$, and henceforth write 'clause' for '2-clause'.

We are interested in the number of 2-SAT functions of n variables, which, following [2], we denote by $G(n)$. Of course $G(n)$ is at most $\exp_2[4\binom{n}{2}]$, the number of 2-SAT formulae; on the other hand it is easy to see that

$$G(n) > 2^n\left(2^{\binom{n}{2}} - n2^{\binom{n-1}{2}}\right) \sim 2^{\binom{n+1}{2}}. \tag{1.2}$$

(All formulas obtained by choosing $y_i \in \{x_i, \overline{x}_i\}$ for each i and a set of clauses using precisely the literals $y_1,\ldots,y_n$ give different functions.)

† Supported by NSF grant DMS0701175.

The problem of estimating $G(n)$ was suggested by Bollobás, Brightwell and Leader [2] (and also, according to [2], by U. Martin). They showed that

$$G(n) = \exp_2[(1 + o(1))n^2/2], \tag{1.3}$$

and made the natural conjecture that (1.2) gives the asymptotic value of $G(n)$; this was proved in [1].

Theorem 1.1.

$$G(n) = (1 + o(1))2^{\binom{n+1}{2}}. \tag{1.4}$$

Here we give an alternative proof. An interesting feature of our argument is that it follows the original coloured graph approach of [2], in the process determining (Theorem 1.2) the asymptotics of the number of 'odd-blue-triangle-free' graphs on n vertices; both [2] and [1] mention the seeming difficulty of proving Theorem 1.1 along these lines.

The argument of [2] reduces (1.3) to estimation of the number of 'odd-blue-triangle-free' (OBTF) graphs (defined below). In brief, with elaboration below, this goes as follows. Each 'elementary' 2-SAT function (non-elementary functions are easily disposed of) corresponds to an OBTF graph; this correspondence is not injective, but the number of functions mapping to a given graph is trivially $\exp[o(n^2)]$, so that a bound $\exp_2[(1 + o(1))n^2/2]$ on the number, say $F(n)$, of OBTF graphs on n vertices – proving which is the main occupation of [2] – gives (1.3).

The Bollobás, Brightwell and Leader reduction to OBTF graphs is also the starting point for the proof of Theorem 1.1, and a second main point here (Theorem 1.2) will be determination of the asymptotic behaviour of $F(n)$. Note, however, that derivation of Theorem 1.1 from this – in contrast to the corresponding step in [2] – is not at all straightforward, since we can no longer afford a crude bound on the number of 2-SAT functions corresponding to a given OBTF G.

It is natural to try to attack the problem of (approximately) enumerating OBTF graphs using ideas from the large literature on asymptotic enumeration in the spirit of [4], for instance [5] and [6]. This is suggested in [2]; but the authors say their attempts in this direction were not successful, and their eventual treatment of $F(n)$ is based instead – as is Allen's proof of Theorem 1.1 – on the Regularity Lemma of Endre Szemerédi [7]. Here our arguments will be very much in the spirit of the papers mentioned; [6] in particular was helpful in providing some initial inspiration. We now turn to more precise descriptions.

We consider *coloured* graphs, meaning graphs with edges coloured *red* (R) and *blue* (B). For such a graph G, a subset of $E(G)$ is *odd-blue* if it has an odd number of *blue* edges (and *even-blue* otherwise), and (of course) G is *odd-blue-triangle-free* (OBTF) if it contains no odd-blue triangle. We use $\mathcal{F}(n)$ for the set of (labelled) OBTF graphs on n vertices and set $|\mathcal{F}(n)| = F(n)$.

A graph G (coloured as above) is *blue-bipartite* (BB) if there is a partition $U \sqcup W$ of $V(G)$ such that each blue edge has one endpoint in each of U, W, while any red edge is contained in one of U, W. We use $\mathcal{B}(n)$ for the set of blue-bipartite graphs on n vertices. It is easy to see that

$$|\mathcal{B}(n)| = (1 - o(1))2^{\binom{n+1}{2}-1}. \tag{1.5}$$

(The term $\exp_2[\binom{n+1}{2} - 1] = \exp_2[n - 1 + \binom{n}{2}]$ counts ways of choosing the unordered pair $\{U, W\}$ and an *uncoloured* G, the colouring then being dictated by 'blue-biparticity'; that the right-hand side of (1.5) is a *lower* bound follows from the observation that almost all such choices will have G connected, in which case different $\{U, W\}$ give different colourings.)

As mentioned above, the main step in the proof of (1.3) in [2] was a bound $F(n) < \exp_2[(1 + o(1))n^2/2]$; here we prove the natural conjecture that most OBTF graphs are blue-bipartite.

Theorem 1.2.

$$F(n) = (1 + o(1))2^{\binom{n+1}{2}-1}. \tag{1.6}$$

The bound here corresponds to that in Theorem 1.1, in that (as explained below) one expects a typical OBTF G to correspond to exactly two 2-SAT functions. Proving that this is indeed the case, and controlling the contributions of those OBTF Gs for which the number is larger, are the main concerns of Sections 4 and 5 (which handle blue-bipartite and non-blue-bipartite G respectively). These are preceded by a review, in Section 2, of the reduction from 2-SAT functions to OBTF graphs, and, in Section 3, the proof of Theorem 1.2 in a form that gives some further limitations on graphs in $\mathcal{F}(n)\backslash\mathcal{B}(n)$. The end of the proof of Theorem 1.1 is given in Section 6, and Section 7 contains a few additional remarks and questions.

Numerical usage. We use $[n]$ for $\{1, \ldots, n\}$ and $\binom{n}{<k}$ for $\sum_{i=0}^{k-1} \binom{n}{i}$. All logarithms and the entropy $H(\cdot)$ are binary. We pretend throughout that large numbers are integers.

Graph-theoretic usage. We use Γ_x or $\Gamma(x)$ for the neighbourhood of a vertex x, preferring the former but occasionally resorting to the latter for typographical reasons (to avoid double subscripts or because we need the subscript to specify the graph). For a set of vertices Q, $\Gamma(Q)$ is $\cup_{x\in Q}\Gamma_x\backslash Q$. We use $\nabla(X, Y)$ for the set of edges having one end in X and the other in Y (X and Y will usually be disjoint, but we do not require this).

2. Reduction to OBTF graphs

In this section we recall what we need of the reduction from 2-SAT functions to OBTF graphs, usually referring to [2] for details.

The *spine* of a non-trivial 2-SAT function S is the set of variables that take only one value (True or False) in satisfying assignments for S. For a 2-SAT function S with empty spine, we say that variables x, y are *associated* if either $x \Leftrightarrow y$ is True in all satisfying assignments for S, or $x \Leftrightarrow \bar{y}$ has this property. A 2-SAT function with empty spine and no associated pairs is *elementary*. As shown in [2], the number, $H(n)$, of elementary, n-variable 2-SAT function satisfies

$$H(n) \leqslant G(n) \leqslant 1 + \sum_{k=0}^{n} H(n-k)\binom{n}{k}(2n-2k+2)^k,$$

and it follows that for Theorem 1.1 it is enough to show

$$H(n) = (1+o(1))2^{\binom{n+1}{2}}. \tag{2.1}$$

Given a 2-SAT formula F giving rise to an elementary function S_F, we construct a partial order P_F on $\{x_1, \bar{x}_1, \ldots, x_n, \bar{x}_n\}$, by setting $x < y$ if $\bar{x} \vee y$ appears in F (so $x \Rightarrow y$ is True in any satisfying assignment for F; note that x, y can be positive or negative literals), and taking the transitive closure of this relation. Then P_F is indeed a poset and satisfies:

(a) P_F depends only on the function S_F,
(b) each pair x, $\bar{x}$ is incomparable, and
(c) $x < y$ if and only if $\bar{y} < \bar{x}$.

This construction turns out to give a bijection between the set of elementary 2-SAT functions and the set $\mathcal{P}(n)$ of posets on $\{x_1, \bar{x}_1, \ldots, x_n, \bar{x}_n\}$ satisfying (b) and (c), and in proving (2.1) we work with the interpretation $H(n) = |\mathcal{P}(n)|$.

For $P \in \mathcal{P}(n)$ we construct a coloured graph $G(P)$ on (say) vertex set $\{w_1, \ldots, w_n\}$ by including a *red* edge w_iw_j whenever $x_i \lessdot \bar{x}_j$ or $\bar{x}_i \lessdot x_j$ in P (where, as usual, $x \lessdot y$ means $x < y$ and there is no z with $x < z < y$), and a *blue* edge w_iw_j whenever $x_i \lessdot x_j$ or $\bar{x}_i \lessdot \bar{x}_j$. Then:

(d) no edge of $G(P)$ is coloured both red and blue,
(e) $G(P)$ determines the *cover graph* of P (the set of pairs $\{x, y\}$ for which $x \lessdot y$ or $y \lessdot x$), and
(f) $G(P) \in \mathcal{F}(n)$.

Of course (e) is not enough to get us from Theorem 1.2 to the desired bound (2.1) on $H(n)$ $(= |\mathcal{P}(n)|)$, since it may be that a given cover graph corresponds to many Ps. It turns out that a typical blue-bipartite G does give rise to exactly two Ps; but bounding the contributions of general Gs is not so easy, and, *inter alia*, will require a somewhat stronger version of Theorem 1.2 (Theorem 3.1). If we only wanted Theorem 1.2, then Section 3 could be simplified, though the basic argument would not change.

3. Nearly blue-bipartite

Fix C and $\varepsilon > 0$. We will not bother giving these numerical values. We choose ε so that the expression on the right-hand side of (3.7) is less than 2, let $c < 1 - 0.6 \log_2 3$ be some positive constant satisfying (3.7), and choose (say)

$$C > 12/c. \tag{3.1}$$

Set $s = s(n) = C \log n$.

Throughout the following discussion, G is assumed to lie in $\mathcal{F}(n)$ and we use V for $[n]$, the common vertex set of these Gs. Set $\kappa(G) = \min\{|K| : K \subseteq V, G - K \text{ is blue-bipartite}\}$. Our main technical result is as follows.

Theorem 3.1. *There is a constant $c > 0$ such that, for sufficiently large n and any $t \leqslant s$,*

$$|\{G : \kappa(G) \geqslant t\}| < 2^{(1-c)sn}F(n-s) + 2^{(1-c)n\lceil t/3\rceil}F(n - \lceil t/3\rceil). \tag{3.2}$$

Notice that, according to (1.6), we expect $F(n) \approx 2^{na}F(n-a)$ (for a not too large); so (3.2) says that non-BB graphs contribute little to this growth. The easy derivation of Theorem 1.2 from Theorem 3.1 is given near the end of this section.

Very roughly, the proof of Theorem 3.1 proceeds by identifying several possible ways in which a graph might be anomalously sparse (see Lemmas 3.2–3.4 and 3.6), and showing that graphs with many anomalies are rare, while for those with few, κ is small. Central to our argument will be our ability to say that, for most G and most vertices x, there is a small (size about $\log n$) subset of Γ_x whose neighbourhood is most of G. The next lemma is a first step in this direction.

Let

$$\mathcal{X}_0(n,t) = \{G \in \mathcal{F}(n) : \exists\, Q \subseteq V(G) \text{ with } |Q| = t \text{ and } |\Gamma(Q)| < 0.6n\}$$

and $\mathcal{X}_0 = \mathcal{X}_0(n) = \mathcal{X}_0(n,s)$.

Lemma 3.2. *For sufficiently large n and $s \geqslant t > \omega(1)$,*

$$|\mathcal{X}_0(n,t)| < 2^{(0.6\log_2 3+o(1))tn}F(n-t).$$

Remark. The statement is actually valid as long as $t < o(n)$, but we will only use it with $t = s$. In place of 0.6 we could use any constant α with $\alpha\log_2 3 < 1$ and $\alpha > 1/2$, the latter being crucial for Lemma 3.4.

Proof. All $G \in \mathcal{X}_0(n,t)$ can be constructed by choosing: Q; $G - Q$; $\Gamma(Q)$; and the restriction of G (including colours) to edges meeting Q (where we require $|Q| = t$ and $|\Gamma(Q)| < 0.6n$). We may bound the numbers of choices for these steps by (respectively): $\binom{n}{t}$; $F(n-t)$; 2^{n-t}; and $\exp_3[\binom{t}{2} + t(0.6n)]$. The lemma follows. □

Set $K_1(G) = \{x \in V(G) : d_G(x) < \varepsilon n\}$, $\kappa_1(G) = |K_1(G)|$, and for $t \leqslant n$,

$$\mathcal{X}_1(n,t) = \{G \in \mathcal{F}(n) : \kappa_1(G) \geqslant t\}.$$

Set $\mathcal{X}_1 = \mathcal{X}_1(n) = \mathcal{X}_1(n,s)$.

Lemma 3.3. *For sufficiently large n and any t,*

$$|\mathcal{X}_1(n,t)| < 2^{(H(\varepsilon)+\varepsilon+o(1))nt}F(n-t),$$

Proof. All $G \in \mathcal{X}_1(n,t)$ can be constructed by choosing: some t-subset K of $K_1(G)$; $G - K$; and Γ_x and colours for $\nabla(x, \Gamma_x)$ for each $x \in K$. (Of course redundancies here and in similar arguments later only help us.) The numbers of choices for these steps are bounded by:

$$\tbinom{n}{t},\quad F(n-t),\quad \text{and}\quad \Big(\tbinom{n}{<\varepsilon n}2^{\varepsilon n}\Big)^t.$$

The lemma follows. □

For each $G \in \mathcal{F}(n)$, let $K_2(G) = \{x_1,\dots,x_l,y_1,\dots,y_l\}$ be a largest possible collection of (distinct) vertices of $V\backslash K_1(G)$ such that $|\Gamma(x_i)\cap\Gamma(y_i)| < \varepsilon n\ \forall\ i \in [l]$, and set $\kappa_2(G) = l$. Let

$$\mathcal{X}_2(n,t) = \{G \in \mathcal{F}(n)\backslash(\mathcal{X}_0(n)\cup\mathcal{X}_1(n)) : \kappa_2(G) \geqslant t\}$$

and $\mathcal{X}_2 = \mathcal{X}_2(n) = \mathcal{X}_2(n,s)$. The next lemma is perhaps our central one.

Lemma 3.4. *For sufficiently large n and $t < o(n)$,*

$$|\mathcal{X}_2(n,t)| < 2^{2H(\varepsilon)nt}(3/4)^{0.2nt}4^{nt}F(n-2t).$$

(In fact we only use this with $t < O(\log n)$.)

Proof. All $G \in \mathcal{X}_2(n,t)$ can be constructed by choosing:

(i) $K = \{x_1,\dots,x_t,y_1,\dots,y_t\} \subseteq V\backslash K_1(G)$ (with $x_1,\dots,y_t$ distinct);
(ii) for each $i \in [t]$, $T_i := \Gamma(x_i)\cap\Gamma(y_i)$ of size at most εn;
(iii) $G' := G[V']$ (including colours), where $V' = V\backslash K$;
(iv) for each $v \in K$, some $Q_v \subseteq \Gamma(v)\backslash K$ of size s and colours for $\nabla(v,Q_v)$;
(v) the remaining edges of G meeting K (those not in $\cup_{v\in K}\nabla(v,Q_v)$) and colours for these edges.

(The point of (iv) is that, since G is OBTF, the colours for $\nabla(v,Q_v)$ together with those for edges of G' meeting Q_v limit our choices for the remaining edges at v.)

We may bound the numbers of choices in steps (i)–(iv) by n^{2t}, $2^{H(\varepsilon)nt}$, $F(n-2t)$, and $(\binom{n}{s}2^s)^{2t} < n^{2st}$ respectively, and the number of choices for $G[K]$ (in (v)) by $\exp_3[\binom{2t}{2}]$.

Given these choices (and aiming to bound the number of possibilities for $\nabla(K,[n]\backslash K)$), we write Γ' for $\Gamma_{G'}$, and, for $i \in [t]$, define $P_i = Q_{x_i}$, $Q_i = Q_{y_i}$; $R_i = (\Gamma'(P_i)\cap\Gamma'(Q_i))\backslash T_i$, $R_i' = \Gamma'(P_i)\backslash\Gamma'(Q_i)\backslash T_i$, $R_i'' = \Gamma'(Q_i)\backslash\Gamma'(P_i)\backslash T_i$ and $\bar{R}_i = V'\backslash(R_i\cup R_i'\cup R_i''\cup T_i)$; and $\alpha_i = |R_i|$, $\alpha_i' = |R_i'|$, $\alpha_i'' = |R_i''|$, $\beta_i = |\bar{R}_i|$ and $\delta_i = |T_i|$ ($< \varepsilon n$).

We then consider (the interesting part) the number of possibilities for $\nabla(z,\{x_i,y_i\})$ (including colours) for $z \in V'$. With explanations to follow, this number is at most: (i) 5 if $z \in \bar{R}_i$; (ii) 4 if $z \in T_i\cup R_i'\cup R_i''$; and (iii) 3 if $z \in R_i$. This is because:

(i) $z \notin T_i$ excludes the four possibilities with z connected to both x_i and y_i;
(ii) for $z \in T_i$ this is obvious; for $z \in R_i'$, we already know the colours on some (x_i,z)-path of length two, so the condition OBTF leaves only one possible colour for an edge between x_i and z, thus excluding one of the five possibilities in (i) (and similarly for $z \in R_i''$);
(iii) here we have (as in (ii)) one excluded colour for each of x_iz, y_iz.

Thus, letting z vary and noting that $\alpha_i + \alpha_i' + \alpha_i'' + \beta_i + \delta_i = n - 2t$, we find that the number of possibilities for $\nabla(\{x_i, y_i\}, V')$ is at most

$$5^{\beta_i} 4^{\alpha_i' + \alpha_i'' + \delta_i} 3^{\alpha_i} < 4^n \left(\frac{15}{16}\right)^{\beta_i} \left(\frac{3}{4}\right)^{\alpha_i - \beta_i} \leqslant 4^n \left(\frac{3}{4}\right)^{\alpha_i - \beta_i}.$$

The crucial point in all this is that $G \notin \mathcal{X}_0$ guarantees that $\alpha_i - \beta_i$ is big: each of $\alpha_i + \alpha_i'$, $\alpha_i + \alpha_i''$ is at least $0.6n - 2t - \delta_i$, and hence

$$n - 2t - \delta_i - \beta_i = \alpha_i + \alpha_i' + \alpha_i'' > 1.2n - 4t - 2\delta_i - \alpha_i,$$

implying $\alpha_i - \beta_i > 0.2n - 2t - \delta_i > (0.2 - \varepsilon)n - 2t$. So, finally, applying this to each i and combining with our earlier bounds (for (i)–(iv) and the first part of (v)) bounds the total number of possibilities for G by

$$2^{(H(\varepsilon)+o(1))nt}(3/4)^{(0.2-\varepsilon-o(1))nt}4^{nt}F(n-2t),$$

which is less than the bound in the lemma. □

Lemma 3.5. *For any $G \in \mathcal{F}(n) \setminus \mathcal{X}_0(n)$, $x \in V \backslash (K_1(G) \cup K_2(G))$ and $S \subseteq V$ of size $6s$, there exists $Q_x \subset \Gamma_x \backslash S$ with*

$$|Q_x| = \log n \quad \textit{and} \quad |V \backslash (K_1(G) \cup K_2(G) \cup \Gamma(Q_x))| < 2n^{1-\varepsilon}. \tag{3.3}$$

Proof. We have $|(\Gamma_x \cap \Gamma_y) \backslash S| \geqslant \varepsilon n - 6s$ for any $x, y \in V \backslash (K_1(G) \cup K_2(G))$. So, for any such x, y and Q a random (uniform) $(\log n)$-subset of $\Gamma_x \backslash S$,

$$\Pr(Q \cap \Gamma_y = \emptyset) < \left(1 - \frac{\log n}{n}\right)^{|\Gamma_x \cap \Gamma_y| - 6s} < \left(1 - \frac{\log n}{n}\right)^{\varepsilon n - 6s} < 2n^{-\varepsilon}.$$

Thus $\mathsf{E}|V \backslash (K_1(G) \cup K_2(G) \cup \Gamma(Q))| < 2n^{1-\varepsilon}$, and the lemma follows. □

For $x \in V$ and $Q_x \subseteq \Gamma_x$, say $z \in V$ is *inconsistent for* (x, Q_x) if there is an odd-blue cycle xx_1zx_2 with $x_1, x_2 \in Q_x$, and write $I(x, Q_x)$ for the set of such z. If in addition $y \sim x$ and $Q_y \subseteq \Gamma_y$, say $z \in V$ is *inconsistent for* (x, Q_x, y, Q_y) if $z \in I(x, Q_x) \cup I(y, Q_y)$ or there is an odd-blue cycle xx_1zy_1y with $x_1 \in Q_x$ and $y_1 \in Q_y$, and write $I(x, Q_x, y, Q_y)$ for the set of such z.

For $G \in \mathcal{F}(n)$, let $K_3(G) = \{x_1, \ldots, x_l, y_1, \ldots, y_l\}$ be a largest possible collection of (distinct) vertices of $V \backslash (K_1(G) \cup K_2(G))$ with $x_i \sim y_i$, and for which there exist $Q_v \subseteq \Gamma_v \backslash K_3$ for $v \in K_3$ satisfying (3.3) and

$$|I(x_i, Q_{x_i}, y_i, Q_{y_i})| > \varepsilon n \ \forall\ i \in [l]. \tag{3.4}$$

Set $\kappa_3(G) = l$. Let

$$\mathcal{X}_3(n, t) = \{G \in \mathcal{F}(n) \backslash (\mathcal{X}_0 \cup \mathcal{X}_1 \cup \mathcal{X}_2) : \kappa_3(G) \geqslant t\}$$

and $\mathcal{X}_3 = \mathcal{X}_3(n) = \mathcal{X}_3(n, s)$.

Now, for $G \in \mathcal{F}(n) \backslash \mathcal{X}_0$ and each $x \in V \backslash (K_1(G) \cup K_2(G))$, fix some $Q_x \subseteq \Gamma_x \backslash K_3(G)$ satisfying (3.3) and (3.4) if $x \in K_3(G)$ and (3.3) otherwise. Existence of such Q_xs is given by Lemma 3.5, and the maximality of $K_3(G)$ implies that for each $xy \in E(G-(K_1(G)\cup K_2(G)\cup K_3(G)))$ we have $I(x,Q_x,y,Q_y)\leqslant \varepsilon n$. Having fixed these Q_xs, we abbreviate $I(x,Q_x)=I(x)$ and $I(x,Q_x,y,Q_y)=I(x,y)$.

Lemma 3.6. *For sufficiently large n and $t\leqslant s$,*

$$|\mathcal{X}_3(n,t)| < (3/4)^{\varepsilon nt}2^{o(nt)}4^{nt}F(n-2t).$$

Proof. All $G\in\mathcal{X}_3(n,t)$ can be constructed by choosing:

(i) $K=\{x_1,\dots,x_t,y_1,\dots,y_t\}$ (with $x_1,\dots,y_t$ distinct);
(ii) $G[V']$ (including colours), where $V'=V\backslash K$;
(iii) for each $x\in K$, Q_x and colours for $\nabla(x,Q_x)$;
(iv) the remaining edges meeting K and colours for these edges.

We may bound the numbers of choices in steps (i)–(iii) by n^{2t}, $F(n-2t)$, and $n^{2t\log n}$ respectively, and the number of choices for $G[K]$ (in (iv)) by $3^{\binom{2t}{2}}$. Notice that the choices in (i)–(iii) determine the sets $I(x_i,y_i)$, which in particular are of size at least εn.

As in Lemma 3.4, the interesting point is the number of possibilities for $\nabla(z,\{x_i,y_i\})$ for $z\in V'$. In general, if $z\in\Gamma(Q_{x_i})\cap\Gamma(Q_{y_i})$ this number is at most 4, since (because G is to be OBTF) any path x_ixz with $x\in\Gamma(Q_{x_i})$ – so we already know the colours of x_ix and xz – excludes one possible colour for a (possible) edge x_iz, and similarly for y_i. Moreover, if $z\in I(x_i,y_i)$ then the number is at most 3: if $z\in I(x_i)$ then an edge x_iz of *either* colour gives an odd-blue triangle, and similarly if $z\in I(y_i)$; and otherwise, we cannot have z joined to both x_i and y_i without creating an odd-blue triangle (and we already know an edge x_iz or y_iz admits at most one possible colour). If $z\notin\Gamma(Q_{x_i})\cap\Gamma(Q_{y_i})$, then we just bound the number by 9, noting that the number of such z is $o(n)$ (since the Qs satisfy (3.3)).

Thus the number of possibilities for $\nabla(\{x_i,y_i\},V')$ is at most

$$4^{n-\varepsilon n}3^{\varepsilon n}9^{o(n)}=4^n(3/4)^{\varepsilon n}2^{o(n)};$$

so combining with our earlier bounds we find that the number of possibilities for G is less than $(3/4)^{\varepsilon nt}2^{o(nt)}4^{nt}F(n-2t)$. □

For $G\notin\mathcal{X}_0$ let $K(G)=K_1(G)\cup K_2(G)\cup K_3(G)$. As we will see, Theorem 3.1 is now an easy consequence of the following result.

Lemma 3.7. *For each $G\in\mathcal{F}(n)\backslash(\mathcal{X}_0\cup\cdots\cup\mathcal{X}_3)$, $G-K(G)$ is blue-bipartite.*

Proof. We first assert that:

$$G-K(G) \text{ contains no odd-blue cycle of length 4 or 5.} \qquad (3.5)$$

To see this, suppose $x_1, \ldots, x_q$ is a cycle in $G' := G - K(G)$ with $q \in \{4,5\}$, and (with subscripts taken mod q) let

$$z \in \bigcap_{i=1}^{q} \Gamma(Q_{x_i}) \setminus \left(\bigcup_{i=1}^{q} I(x_i, x_{i+1}) \cup \{x_1, \ldots, x_q\} \right).$$

(Note that there is such a z; in fact the size of the set on the right-hand side is at least $n - |K_1(G) \cup K_2(G) \cup K_3(G)| - q \log n - 2qn^{1-\varepsilon} - q\varepsilon n - q$, so essentially $(1 - q\varepsilon)n$.)

Let $w_i \in \Gamma(z) \cap Q_{x_i}$ ($i \in [q]$). Each of the closed walks $zw_ix_ix_{i+1}w_{i+1}z$ is even-blue, either (if it is a 5-cycle) because $z \notin I(x_i, x_{i+1})$, or (otherwise) because G is OBTF, where we use the easy result:

$$\text{Any non-simple closed walk of length at most 5 in an OBTF graph is even-blue.} \tag{3.6}$$

But since these walks together with the original cycle use each edge of G an even number of times, it follows that the original cycle is also even-blue.

We now define the blue-bipartition for G' in the natural way. Note that the diameter of G' is at most 2 (in fact any two vertices of G' have at least $\varepsilon n - o(n)$ common neighbours), and that (3.5) and (3.6) imply that for any two vertices x, y, all (x, y)-paths of length at most 2 have the same *blue-parity* (defined in the obvious way). We may thus fix some vertex x and let U consist of those vertices for which this common parity is even (so $x \in U$) and $W = V(G') \setminus U$. That this is indeed a blue-bipartition is again an easy consequence of (3.5) and (3.6). □

Proof of Theorem 3.1. For $G \in \mathcal{F}(n) \setminus \mathcal{X}_0$, Lemma 3.7 gives $\kappa(G) \leqslant \kappa_1(G) + 2(\kappa_2(G) + \kappa_3(G))$, so that $\kappa(G) \geqslant t$ implies that either $\kappa_1(G) \geqslant t/3$ or at least one of $\kappa_2(G)$, $\kappa_3(G)$ is at least $t/6$. Since, for $G \notin \mathcal{X}_0$,

$$\begin{aligned}
&\kappa_1(G) \geqslant t/3 \Rightarrow G \in \mathcal{X}_1(n, \lceil t/3 \rceil), \\
&\kappa_2(G) \geqslant t/6 \Rightarrow G \in \mathcal{X}_1 \cup \mathcal{X}_2(n, \lceil t/6 \rceil), \\
&\kappa_3(G) \geqslant t/6 \Rightarrow G \in \mathcal{X}_1 \cup \mathcal{X}_2 \cup \mathcal{X}_3(n, \lceil t/6 \rceil), \quad \text{and} \\
&\quad \mathcal{X}_1 \subseteq \mathcal{X}_1(n, \lceil t/3 \rceil), \quad \mathcal{X}_2 \subseteq \mathcal{X}_2(n, \lceil t/6 \rceil),
\end{aligned}$$

it follows that

$$\{G : \kappa(G) \geqslant t\} \subseteq \mathcal{X}_0 \cup \mathcal{X}_1(n, \lceil t/3 \rceil) \cup \mathcal{X}_2(n, \lceil t/6 \rceil) \cup \mathcal{X}_3(n, \lceil t/6 \rceil).$$

The theorem, with any (fixed, positive) $c < 1 - 0.6 \log_2 3$ satisfying

$$2^{-c} > 2^{H(\varepsilon)+\varepsilon-1} + 2^{H(\varepsilon)}(3/4)^{0.1} + (3/4)^{\varepsilon/2}, \tag{3.7}$$

now follows from Lemmas 3.2–3.4 and 3.6. □

From this point we set $b(n) = 2^{\binom{n+1}{2}-1}$ ($\sim |\mathcal{B}(n)|$).

Proof of Theorem 1.2. We prove Theorem 1.2 by showing by induction that, for some constant Δ, c as in Theorem 3.1 and all n,

$$F(n) \leqslant (1 + \Delta \cdot 2^{-cn})b(n). \tag{3.8}$$

To see this, choose n_0 large enough so that the previous results in this section are valid for $n \geqslant n_0$, and then choose $\Delta > 2$ (say) so that (3.8) holds for $n \leqslant n_0$. Assuming (3.8) holds for $n-1$, we have, using Theorem 3.1 for the first inequality,

$$\begin{aligned} |\mathcal{F}(n)\backslash\mathcal{B}(n)| &= |\{G : \kappa(G) > 0\}| \\ &< 2^{(1-c)sn}F(n-s) + 2^{n-cn}F(n-1) \\ &< 2^{(1-c)sn}(1+\Delta 2^{-c(n-s)})b(n-s) \\ &\qquad + 2^{n-cn}(1+\Delta 2^{-c(n-1)})b(n-1) \\ &= [2^{\binom{s}{2}-csn}(1+\Delta 2^{-c(n-s)}) \\ &\qquad + 2^{-cn}(1+\Delta 2^{-c(n-1)})]b(n). \end{aligned} \tag{3.9}$$

So, since $|\mathcal{B}(n)| < b(n)$ and the coefficient of $b(n)$ in (3.9) is less than $\Delta 2^{-cn}$, we have (3.8). □

Feeding this back into Theorem 3.1 we obtain a quantitative strengthening of Theorem 1.2 that will be useful below. (Recall we assume $G \in \mathcal{F}(n)$.)

Theorem 3.8. *For any constant $\delta < c/3$, $t \leqslant s$ and large enough n,*

$$|\{G : \kappa(G) \geqslant t\}| < 2^{-\delta nt}b(n). \tag{3.10}$$

Proof. We have (for large enough n)

$$\begin{aligned} |\{G : \kappa(G) \geqslant t\}| &< 2^{(1-c)sn}F(n-s) + 2^{(1-c)n\lceil t/3\rceil}F(n-\lceil t/3\rceil) \\ &< 2[2^{(1-c)sn}b(n-s) + 2^{(1-c)n\lceil t/3\rceil}\cdot b(n-\lceil t/3\rceil)] \\ &< 2^{-\delta nt}b(n). \end{aligned}$$

□

In what follows we will also need an analogue of κ for edge removals, say

$$\gamma(G) := \min\{|E'| : E' \subseteq E(G), G - E' \text{ BB}\}.$$

Lemma 3.9. *There is a constant C' such that, for sufficiently large n,*

$$|\{G : \gamma(G) > C'\sqrt{n}\log^{3/2} n\}| < n^{-3n}b(n).$$

Proof. Fix $A > ((12\log_2 3)/c + 3)^{1/2}$. The story here is that $\kappa(G)$ small implies $\gamma(G)$ small unless we encounter the following pathological situation. Let $\mathcal{Y}(n)$ consist of those $G \in \mathcal{F}(n)$ for which there is some $K \subseteq V$ of size at most $k := (12\log n)/c$ such that $G - K$ is BB and there are disjoint $S, T \subseteq V\backslash K$, each of size $A\sqrt{n\log n}$, with $\nabla_G(S,T) = \emptyset$. We assert that, for any constant $C'' < A^2 - (12\log_2 3)/c$ (and large n),

$$|\mathcal{Y}(n)| < \exp_2\left[\binom{n}{2} - C''n\log n\right]. \tag{3.11}$$

This is a routine calculation: the number of choices for $G \in \mathcal{Y}(n)$ is at most

$$3^n \exp_3\left[\binom{k}{2} + k(n-k)\right] \exp_2\left[\binom{n}{2} - A^2 n \log n\right],$$

where the first term corresponds to the choices of K, the blue-bipartition and S, T; the second to edges of G meeting K; and the third to the remaining edges (whose colours are determined by the blue-bipartition). This gives (3.11).

Thus, in view of Theorem 3.8 (noting $(12/c)\log n < s$; see (3.1)), Lemma 3.9 will follow from

$$G \in \mathcal{F}(n)\backslash\mathcal{Y}(n),\ \kappa(G) < (12/c)\log n \ \Rightarrow\ \gamma(G) < C'\sqrt{n}\log^{3/2} n \tag{3.12}$$

(for a suitable C'). To see this, suppose $G \notin \mathcal{Y}(n)$ and $G - K$ is BB with $|K| < (12/c)\log n$. Let $X \cup Y$ be a blue-bipartition of $G - K$, and write R and B for the sets of red and blue edges of G. Given $x \in K$, let $R_X = R_X(x) = \{v \in X : xv \in R\}$, and define B_X, R_Y, B_Y similarly. Then G OBTF implies

$$\nabla(R_X, B_X) = \nabla(R_X, R_Y) = \nabla(B_X, B_Y) = \nabla(R_Y, B_Y) = \emptyset,$$

and thus (since $G \notin \mathcal{Y}(n)$) we may assume that at most two of R_X, B_X, R_Y, B_Y have size at least $A\sqrt{n\log n}$, and if exactly two then these must be either R_X and B_Y, or B_X and R_Y. Thus there is a set $E'(x)$ of at most $2A\sqrt{n\log n}$ edges at x so that either $\nabla(x, X)\backslash E'(x) \subseteq R$ and $\nabla(x, Y)\backslash E'(x) \subseteq B$ or *vice versa*. Setting $E' = E(K) \cup \bigcup\{E'(x) : x \in K\}$, we find that $G - E'$ is BB with $|E'| < \binom{|K|}{2} + 2A|K|\sqrt{n\log n} < C'\sqrt{n}\log^{3/2} n$, for any $C' > 24A/c$ (and large n). □

4. Blue-bipartite graphs

We continue to assume $G \in \mathcal{F}(n)$ and now need some understanding of the sizes of the sets

$$\mathcal{P}(G) := \{P \in \mathcal{P}(n) : G(P) = G\}$$

(see following (2.1) for $\mathcal{P}(n)$ and $G(P)$). Recall (see property (e) of $G(P)$) that $G(P)$ determines the cover graph of P; thus, as observed in [2], we trivially have

$$|\mathcal{P}(G)| < (2n)! < n^{2n} \quad \forall G \in \mathcal{F}(n), \tag{4.1}$$

since a poset is determined by its cover graph and any one of its linear extensions.

If $P \in \mathcal{P}(G)$ then the cover graph of P is $C(G)$, which is defined to be the graph on $\{x_1, \ldots, x_n, \bar{x}_1, \ldots, \bar{x}_n\}$ with, for each $w_i w_j \in E(G)$, edges $x_i x_j$ and $\bar{x}_i \bar{x}_j$ if $w_i w_j$ is blue, and $x_i \bar{x}_j$ and $\bar{x}_i x_j$ if it is red. By property (c) in the definition of $\mathcal{P}(n)$, the orientation of either of the edges of $C(G)$ corresponding to a given edge of G determines the orientation of the other; so we speak, a little abusively, of orienting the edges of G.

A basic observation is that the orientations of the edges of any triangle $w_i w_j w_k$ of G are determined by the orientation of any one of them. Suppose for instance (other cases are similar) that the edges of $w_i w_j w_k$ are all *red*, and that $x_i < \bar{x}_j$ (so also $x_j < \bar{x}_i$). We must then have $\bar{x}_k > x_i, x_j$ (and $x_k < \bar{x}_i, \bar{x}_j$), since (*e.g.*) $\bar{x}_k < x_i$ would imply $x_k > \bar{x}_i$, and then $x_k > \bar{x_j}$ would give $x_k > \bar{x_k}$, while $x_k < \bar{x_j}$ would give $\bar{x_j} > x_j$, in either case a contradiction. It follows that the orientation of either of $e, f \in E(G)$ determines the orientation of the other whenever there is a sequence $T_0, \ldots, T_l$ of triangles with e (resp. f) an edge of T_1 (resp. T_l) and T_{i-1}, T_i sharing an edge for each $i \in [l]$. We then write $e \equiv f$, and call the classes of this equivalence relation *triangle-components* of G. If there is just one equivalence class, we say G is *triangle-connected*.

In general the preceding discussion bounds $|\mathcal{P}(G)|$ by $2^{\eta(G)}$ with $\eta(G)$ the number of triangle components of G; but all we need from this is the following result.

Lemma 4.1. *If $G \in \mathcal{B}(n)$ is triangle-connected then $|\mathcal{P}(G)| \leqslant 2$.*

(In fact it is easy to see that equality holds.) The last piece needed for the proof of Theorem 1.1 is as follows.

Lemma 4.2. *There are at most $2^{-\Omega(n)}b(n)$ $P \in \mathcal{P}(n)$ with $G(P)$ in $\mathcal{B}(n)$ and not triangle-connected.*

Proof. Fix $\delta > 0$ with $5(1 - H(\delta)) > 3\delta$, and for $G \in \mathcal{B}(n)$ let $X(G) = \{x \in V : d_G(x) < \delta n\}$. Set $D = 5/\delta$. We first dispose of some pathologies.

Proposition 4.3. *All but at most $n^{-3n}b(n)$ $G \in \mathcal{B}(n)$ satisfy:*

(i) $|X(G)| < D \log n$,

(ii) $\nexists$ *disjoint* $Y, Z \subseteq V$ *with* $|Y||Z| = Dn \log n$ *and* $\nabla(Y, Z) = \emptyset$,

(iii) $\forall x \in V \backslash X(G)$, *the size of the largest connected component of* $G[\Gamma_x]$ *is at least* $d_G(x) - D \log n$,

(iv) $\forall x \in V \backslash X(G)$, $|\{y \in V \backslash X(G) : |\Gamma_x \cap \Gamma_y| < D\sqrt{n \log n}\}| < D \log n$.

Proof. (i) We may specify $G \in \mathcal{B}(n)$ violating (i) by choosing: a blue-bipartition $S \cup T$; $X' \subseteq X(G)$ of size $D \log n$; $E(X') \cup \nabla(X', V \backslash X')$; and $G - X'$. The numbers of ways to make these choices are at most: 2^n; $\binom{n}{D \log n}$; $(\sum_{i<\delta n} \binom{n}{i})^{D \log n}$; and $\exp_2[\binom{n - D \log n}{2}]$; and, in view of our restriction on δ, the product of these bounds is much less than $n^{-3n}b(n)$.

(ii) For use in (iii) we show a slightly stronger version, say (ii′), which is just (ii) with D replaced by 4. We may specify $G \in \mathcal{B}(n)$ violating (ii′) by choosing a blue-bipartition $S \cup T$ and Y, Z in at most (say) 5^n ways, and then the edges of G in at most $\exp_2[\binom{n}{2} - 4n \log n]$ ways.

(iii) Here we simply observe that any $G \in \mathcal{B}(n)$ satisfying (ii′) also satisfies (iii). To see this, notice that if $d_G(x) \geqslant \delta n$ and G satisfies (ii′), then there is no $K \subseteq \Gamma_x$ with $|K| \in (D \log n, d_G(x) - D \log n)$ and $\nabla(K, \Gamma_x \backslash K) = \emptyset$. But then if (iii) fails at x, it must be that all components of $G[\Gamma_x]$ have size less than $D \log n$, in which case we get the supposedly non-existent K as a union of components.

(iv) Here, with $k = D\log n$, we may specify a violator by choosing: a blue-bipartition $S \cup T$; $x \in V$; Γ_x of size at least δn; $y_1, \ldots, y_k \in V$; $\Gamma_{y_i} \cap \Gamma_x$ of size at most $r := D\sqrt{n\log n}$ (for $i \in [k]$); and $E(G)\backslash(\nabla(x) \cup \nabla(\{y_1, \ldots, y_k\}), \Gamma_x)$. The number of possibilities for this whole procedure is at most

$$2^n \cdot n \cdot 2^n \cdot n^k \cdot \max_{m \geqslant \delta n}\{\left(2\tbinom{m}{r}\right)^k \exp_2\left[\tbinom{n}{2} - km + k^2\right]\}.$$

(We used $\sum\{\binom{m}{i} : i \leqslant r\} < 2\binom{m}{r}$; the irrelevant k^2 allows for some y_is in Γ_x; of course we could have strengthened $D\sqrt{n\log n}$ to some $\Omega(n)$.) □

We now return to the proof of Lemma 4.2. Let $\mathcal{H}(n)$ consist of those $G \in \mathcal{B}(n)$ that are not triangle-connected and for which (i)–(iv) of Proposition 4.3 hold. The proposition and (4.1) imply that Lemma 4.2 will follow from

$$\sum\{|\mathcal{P}(G)| : G \in \mathcal{H}(n)\} < 2^{-\Omega(n)}b(n). \tag{4.2}$$

Temporarily fix $G \in \mathcal{H}(n)$ and set $X = X(G)$ and $W = W(G) = V\backslash X$. For $x \in W$ let $\Gamma'_x = \Gamma_x \cap W$. Let L_x be the intersection of (the vertex set of) the largest connected component of $G[\Gamma_x]$ with Γ'_x, $K_x = \Gamma'_x\backslash L_x$, and $E_x = \nabla(x, K_x)$, and observe that all edges contained in $\{x\} \cup L_x$ lie in the same triangle component of G, say $\mathcal{C}(x)$.

For $x, y \in W$, write $x \leftrightarrow y$ if $E(L_x \cap L_y) \neq \emptyset$, and note this implies $\mathcal{C}(x) = \mathcal{C}(y)$. By (ii) we have $x \leftrightarrow y$ whenever $|L_x \cap L_y| > 2\sqrt{Dn\log n}$, and thus, by (iv) and (i),

$$|\{y \in W : y \not\leftrightarrow x\}| < D\log n \quad \forall x \in W.$$

In particular, '$\leftrightarrow$' is the edge set of a connected graph on W, implying all triangle components $\mathcal{C}(x)$ are the same; that is, $E(W) - \cup\{E_x : x \in W\}$ is contained in a single triangle-component of G. Note also that $z \in K_x$ implies $|\Gamma'_x \cap \Gamma'_z| < |K_x| < D\log n$ (by (iii)), so that, again using (iv) and (i), we have

$$|\{x \in W : z \in K_x\}| < \min\{D\log n, \textstyle\sum_{x \in W}|K_x| + 1\} \quad \forall z \in W. \tag{4.3}$$

(The extra 1 in the trivial second bound will sometimes save us from dividing by zero.) In what follows we set $t = |X|$, $m = |W|$ $(= n - t)$, $k_x = |K_x|$ and $\underline{k} = (k_x : x \in W) \in [0, D\log n]^W$.

We now consider the sum in (4.2), *i.e.*, the number of ways to choose $G \in \mathcal{H}(n)$ and $P \in \mathcal{P}(G)$. As usual there are 2^{n-1} ways to choose the blue-bipartition. We then choose $X = X(G)$ and the edges meeting X, the number of ways to do this for a given t being at most $\binom{n}{t}\binom{n}{<\delta n}^t < \exp_2[(\log n + H(\delta)n)t]$, define W and Γ'_x as above, and let $H = G[W]$. *Vertices discussed from this point are assumed to lie in* W, and we set $d'_x = d_H(x)$.

We first consider a fixed $\underline{k}$, setting $g(\underline{k}) = \min\{D\log n, \sum k_x + 1\}$. There are at most $\prod\binom{m}{k_x} < \exp_2[\sum k_x \log n]$ ways to choose the sets K_x. Once these have been chosen, we write $\mathcal{H}$ for the set of possibilities remaining for H. For a particular $H \in \mathcal{H}$, let $\mathcal{U}_H = \{\{y, z\} : \exists x\ y \in K_x, z \in L_x\}$. By (4.3) we have

$$|\mathcal{U}_H| > \frac{1}{g(\underline{k})}\sum_x (d'_x - k_x)k_x > \frac{\delta n}{2g(\underline{k})}\sum k_x. \tag{4.4}$$

Given an ordering $\sigma = (x_1, \ldots, x_m)$ of W, we specify H by choosing, for $i = 1, \ldots, m-1$, $\nabla(x_i, \{x_{i+1}, \ldots, x_m\} \backslash K_{x_i})$. Note that if $i < j, l$, and exactly one of x_j, x_l belongs to each of K_{x_i}, L_{x_i}, then $x_j \not\sim x_l$ is established in the processing of x_i, so we never need to consider potential edge $x_j x_l$ directly. Thus the number of choices, say $f(\sigma, H)$, that we actually make in producing a specific H is at most

$$\binom{m}{2} - |\{(i, \{j, l\}) : i < j, l; x_j \in K_{x_i}, x_l \in L_{x_i}\}|.$$

For a fixed H and random (uniform) σ, the expectation of the subtracted expression is at least $|\mathcal{U}_H|/3$. This gives (using (4.4))

$$\begin{aligned} \frac{1}{m!} \sum_{\sigma} \sum_{H} f(\sigma, H) &= \sum_{H} \frac{1}{m!} \sum_{\sigma} f(\sigma, H) \\ &< \left(\binom{m}{2} - \frac{\delta n}{6g(\underline{k})} \sum k_x \right) |\mathcal{H}|. \end{aligned} \tag{4.5}$$

Thus there is some σ for which $\sum_H f(\sigma, H)$ is at most the right-hand side of (4.5), whence we assert

$$|\mathcal{H}| < \exp_2 \left[\binom{m}{2} - \frac{\delta n}{6g(\underline{k})} \sum k_x \right].$$

Proof. This is a standard observation: for a given σ we may think of the above procedure as a decision tree, with $f(\sigma, H)$ the length of the path leading to the leaf H; and we then have

$$1 \geqslant \sum_H 2^{-f(\sigma,H)} \geqslant |\mathcal{H}| \exp_2 \left[-|\mathcal{H}|^{-1} \sum_H f(\sigma, H) \right]. \qquad \square$$

Finally, we need to choose an orientation. By Lemma 4.1 there are just two ways to orient the edges of the triangle component of G containing $H - \cup\{E_x : x \in W\}$. We then extend to $\cup\{E_x : x \in W\}$ and the remaining edges meeting X in at most $\exp_2[\delta nt + \sum\{k_x : x \in W\}]$ ways. In summary the number of ways to choose the pair (G, P) is less than

$$2^n \sum_t \sum_{\underline{k}} \exp_2 \left[\binom{m}{2} + \big((H(\delta) + \delta)n + \log n\big)t + \left(1 + \log n - \frac{\delta n}{6g(\underline{k})}\right) \sum_{x \in W} k_x \right], \tag{4.6}$$

with the double sum over $t \in [0, D \log n]$ and $\underline{k} \in [0, D \log n]^m$, excluding the $(0, \underline{0})$-term, which counts only triangle-connected graphs. Noting that $\binom{m}{2} = \binom{n}{2} - t(n-t) - \binom{t}{2}$, we find that, for any constant $\gamma < 1 - H(\delta) - \delta$, the expression in (4.6) is (for large n) less than

$$2^{\binom{n+1}{2}} \sum_t \sum_{\underline{k}} \exp_2 \left[-\gamma nt - \frac{\delta n}{7g(\underline{k})} \sum k_x \right] < 2^{-\Omega(n)} b(n). \qquad \square$$

5. Proof of Theorem 1.1

This is now easy. We have

$$|\mathcal{P}(n)| = |\cup\{\mathcal{P}(G) : G \in \mathcal{B}(n)\}| + |\cup\{\mathcal{P}(G) : G \in \mathcal{F}(n)\backslash\mathcal{B}(n)\}|.$$

Here the first term on the right-hand side is asymptotic to $2^{\binom{n+1}{2}}$ by (1.5) and Lemmas 4.1 and 4.2; so we just need to show that the second is $o(b(n))$. Moreover, according to Lemma 3.9 and (4.1), it is enough to show this when we restrict to G with $\gamma(G) \leqslant C'\sqrt{n}\log^{3/2} n$ (C' as in Lemma 3.9). Thus the theorem will follow from

$$\sum\{|\mathcal{P}(G)| : G \in \mathcal{F}(n)\backslash\mathcal{B}(n), \gamma(G) \leqslant C'\sqrt{n}\log^{3/2} n\} < 2^{-\Omega(n)}b(n). \tag{5.1}$$

Proof. For G as in (5.1), let $E' = E'(G)$ be a subset of $E(G)$ of size at most $C'\sqrt{n}\log^{3/2} n$ with $G - E'$ BB. To bound the sum in (5.1) – *i.e.*, the number of possibilities for a pair (G, P) with G as in (5.1) and $P \in \mathcal{P}(G)$ – we consider two cases (in each of which we use the fact that if $P \in \mathcal{P}(G)$, then the poset generated by the restriction of P to $E(G)\backslash E'$ belongs to $\mathcal{P}(G - E')$).

For $G - E'$ not triangle-connected, we may think of choosing $G - E'$ and $P' \in \mathcal{P}(G - E')$, which by Lemma 4.2 can be done in at most $2^{-\Omega(n)}b(n)$ ways, and then choosing E' and extending P' to $P \in \mathcal{P}(G)$ (that is, choosing orientations for the edges of E'), which can be done in at most $2^{o(n)}$ ways.

For $G - E'$ triangle-connected we specify (G, P) by choosing: G; E'; $P' \in \mathcal{P}(G - E')$; and P extending P' to $E(G)$. The number of possibilities in the first step is at most $2^{-\Omega(n)}b(n)$ by Theorem 3.8; the numbers of possibilities in the second and fourth steps are $2^{o(n)}$; and there are (by Lemma 4.1) just two possibilities in step 3. □

6. Questions

One obvious question suggested in [2] is estimation of the number of k-SAT functions for other values of k. Here fixed k seems to us most interesting. It is conjectured in [2] that in this case the number of k-SAT functions is $\exp_2[(1 + o(1))\binom{n}{k}]$, and we see no reason not to expect the following to hold.

Conjecture 6.1. *For fixed k the number of k-SAT functions of n variables is asymptotically* $\exp_2[n + \binom{n}{k}]$.

We would also like to mention one natural question that, surprisingly, seems not to have been considered previously.

Question 6.2. *How many posets can have the same n-vertex cover graph?*

Although it is not even obvious that the answer here is $2^{\omega(n)}$, a construction of Graham Brightwell [3] gives a lower bound $(c\log n/\log\log n)^n$.

Acknowledgement

Thanks to the referee for a careful reading.

References

[1] Allen, P. (2007) Almost every 2-SAT function is unate. *Israel J. Math.* **161** 311–346.
[2] Bollobás, B., Brightwell, G. and Leader, I. (2003) The number of 2-SAT functions. *Israel J. Math.* **133** 45–60.
[3] Brightwell, G. (2008) Personal communication.
[4] Erdős, P., Kleitman, D. J. and Rothschild, B. L. (1976) Asymptotic enumeration of K_n-free graphs. In *Colloquio Internazionale sulle Teorie Combinatorie* (Rome 1973), Vol. II, Accad. Naz. Lincei, Rome, pp. 19–27.
[5] Kleitman, D. J. and Rothschild, B. L. (1975) Asymptotic enumeration of partial orders on a finite set. *Trans. Amer. Math. Soc.* **205** 205–220.
[6] Prömel, H. J., Schickinger, T. and Steger, A. (2002) A note on triangle-free and bipartite graphs. *Discrete Math.* **257** 531–540.
[7] Szemerédi, E. (1978) Regular partitions of graphs. In *Problèmes Combinatoires et Théorie des Graphes* (Colloq. Internat. CNRS, Orsay 1976), CNRS, Paris, pp. 399–401.