The extremal function for noncomplete minors

Joseph Samuel Myers* and Andrew Thomason

Department of Pure Mathematics and Mathematical Statistics
Centre for Mathematical Sciences
Wilberforce Road, Cambridge, CB3 0WB, United Kingdom

J.S.Myers@dpmms.cam.ac.uk
A.G.Thomason@dpmms.cam.ac.uk

14th October 2002; revised 2nd October 2003

Abstract

We investigate the maximum number of edges that a graph G can have if it does not contain a given graph H as a minor (subcontraction). Let

$$c(H) \,=\, \inf\{\, c \,:\, e(G) \geq c|G| \text{ implies } G \succ H \,\}\,.$$

We define a parameter $\gamma(H)$ of the graph H and show that, if H has t vertices, then

$$c(H) \,=\, (\alpha\,\gamma(H) + o(1))\, t\,\sqrt{\log t}$$

where $\alpha = 0.319\ldots$ is an explicit constant and $o(1)$ denotes a term tending to zero as $t \to \infty$. The extremal graphs are unions of pseudo-random graphs.

If H has $t^{1+\tau}$ edges then $\gamma(H) \leq \sqrt{\tau}$, equality holding for almost all H and for all regular H. We show how $\gamma(H)$ might be evaluated for other graphs H also, such as complete multi-partite graphs.

1 Introduction

Classical extremal graph theory is concerned with the maximum number of edges that a graph G can have if it does not contain a given graph H as a subgraph. The central result is the Erdős-Stone-Simonovits theorem [6], stating that if

$$\mathrm{ex}(H) \,=\, \lim_{n\to\infty} \inf\{\, c \,:\, |G| \geq n \text{ and } e(G) \geq c\binom{|G|}{2} \text{ implies } G \supset H \,\}$$

*Research supported by EPSRC studentship 99801140.

then $\mathrm{ex}(H) = 1 - (\chi(H) - 1)^{-1}$, where $\chi(H)$ is the chromatic number of H (in general, we use the standard terminology of [1].) Moreover the extremal graphs are complete $(\chi(H) - 1)$-partite graphs, with perhaps the addition of $o(n^2)$ edges.

In this paper we investigate the corresponding extremal problem where G is forbidden to have H as a *minor*, rather than just as a subgraph. We write $G \succ H$ to mean that the graph H is a minor, or subcontraction, of the graph G, in the usual sense; that is, there exist disjoint subsets $W_u \subset V(G)$ indexed by $u \in V(H)$, such that each $G[W_u]$ is connected and there is an edge of G between W_u and W_v whenever uv is an edge of H. The appropriate way to describe the maximum number of edges in a graph with no H minor is in terms of the function

$$c(H) \,=\, \inf\{\, c \,:\, e(G) \geq c|G| \text{ implies } G \succ H \,\}$$

introduced by Mader [9] for complete graphs K_t ; he showed that $c(K_t)$ is finite, and of course $c(H) \leq c(K_t)$ for all graphs H of order t. In other words, every graph of sufficiently large average degree, namely $2c(H)$, has an H minor. For $t \leq 7$, Mader [10] showed that $G = K_{t-2} + \overline{K}_{n-t+2}$ is an extremal graph with no K_t minor, but for large t this pattern is far from the truth. Kostochka [8], and also the second author [16], established the rate of growth of $c(t)$ as a function of t, showing in particular that random graphs give good, if not necessarily optimal, lower bounds.

Only recently has it been shown [18] that random graphs are indeed asymptotically optimal, so giving the following asymtotic value for $c(t)$:

$$c(K_t) \,=\, (\alpha + o(1))\, t\sqrt{\log t}$$

where $\alpha = 0.319086\ldots$ is explicitly defined. Moreover, the first author [11] has shown that all extremal graphs are pseudo-random graphs, or, more generally, essentially disjoint unions of such pseudo-random graphs.

Given that the function $c(H)$ is now well understood when H is complete, it has become feasible to look at its behaviour for more general H, which is our purpose here. Our hope is to identify some structural property of H that determines $c(H)$, analogously to the way in which the chromatic number $\chi(H)$ determines the classical extremal function for H. We already know that if $H = K_t$ then the extremal graphs are pseudo-random. The questions we ask for general H are therefore these:

(a) What does the function $c(H)$ look like?

(b) Can we describe some structural property of H that determines $c(H)$?

(c) Are the extremal graphs pseudo-random?

These questions would appear to be of increasing complexity but in fact we answer them in the reverse order to that in which they are posed; it turns out that we can show the extremal graphs are pseudo-random without inspecting the structure of H too closely, leaving it till afterwards to decide precisely what properties of H determine the value of $c(H)$.

Even without knowing the value of $c(H)$, the fact that the extremal graphs are pseudo-random means that the situation is quite different from the classical theory. For one thing, we do not have exact expressions for $c(H)$, but only asymptotic ones as t gets large (here, and throughout the paper, t denotes the number of vertices of H); the classical theory, on the other hand, gives exact information about any fixed H (although it may be necessary to require the order of the graph G containing H to be large). Another consequence of the nature of the extremal graphs is that the value of $c(H)$ is not affected significantly by the addition of one or two edges to H, because these edges can easily be found in a random graph; this is in marked contrast to the classical case, where the addition of a single edge can change the chromatic number and so also the extremal function.

2 Main results

We now define a parameter $\gamma(H)$ of H that determines $c(H)$. A consequence of the discussion above is that there is some latitude in how $\gamma(H)$ might be defined; if $\gamma'(H)$ were another parameter with $\gamma'(H) = \gamma(H) + o(1)$, where $o(1)$ denotes something tending to zero as $t \to \infty$, then $\gamma'(H)$ could be used just as well as $\gamma(H)$ in all our results. What this means in practice is that we have a choice of definitions, and different definitions are suited to different parts of the proofs. The definition given here is chosen because it is relatively clean. It is more suited to showing that the extremal graphs are pseudo-random than it is to allowing $\gamma(H)$ to be calculated for a specific H, but once the main theorems have been proven we devote some time (§4) to the evaluation of $\gamma(H)$.

Definition 2.1 *Let H be a graph of order t. We define*

$$\gamma(H) \,=\, \min_{w} \frac{1}{t} \sum_{u \in H} w(u) \qquad \textit{such that} \qquad \sum_{uv \in E(H)} t^{-w(u)w(v)} \,\leq\, t\,,$$

where the minimum is over all assignments $w : V(H) \to \mathbf{R}^+$ of non-negative weights to the vertices of H.

Remark. Observe that, roughly speaking, $\gamma(H)$ is the smallest average vertex weight that is achievable whilst keeping the products of weights across edges large. Also, by grouping together vertices of similar weight, we could think of $\gamma(H)$ as a partition function of some kind.

Remark. If H has $t^{1+\tau}$ edges then we can take $w(u) = \sqrt{\tau}$ for all $u \in H$, and so $\gamma(H) \leq \sqrt{\tau}$. In particular, $\gamma(H) \leq 1$ for every graph H.

Remark. We shall show in §4 that $\gamma(H) = \sqrt{\tau} + o(1)$ for almost all graphs H with $t^{1+\tau}$ edges and for all regular graphs with $t^{1+\tau}$ edges. Indeed, in order for $\gamma(H)$ to be strictly smaller than $\sqrt{\tau}$ it is necessary for H to have a very restrictive structural property that we call having a *tail*. Note in particular that $\gamma(H) = 1 + o(1)$ for almost all graphs H and for all regular graphs H of positive density.

We now state how $c(H)$ depends on $\gamma(H)$. The expression involves a constant α defined by $\alpha = (1-\lambda)/2\sqrt{\log(1/\lambda)}$, where $\lambda < 1$ is the root of the equation $1 - \lambda + 2\lambda \log \lambda = 0$. Numerically, $\lambda = 0.284668\ldots$ and $\alpha = 0.319086\ldots$.

Theorem 2.2 *Let H be a graph of order t. Then*

$$c(H) = (\alpha\,\gamma(H) + o(1))\, t\sqrt{\log t}$$

where $\alpha = 0.319086\ldots$ and the $o(1)$ term denotes a quantity tending to zero as $t \to \infty$.

Remark. It follows from the previous remark that $c(H) = (1 + o(1))c(K_t)$ for almost all H, so most graphs are essentially no easier to find as a minor than is a complete graph of the same order. In this respect the situation here is quite different from the classical theory.

Though Theorem 2.2 answers our initial question about the value of $c(H)$, our main theorems are really the following two, from which Theorem 2.2 is derived and which show that random graphs provide extremal graphs for $c(H)$. We use the simplest model of random graphs, in which $G(n,p)$ denotes a graph with n labelled vertices, its edges being chosen independently at random, each with probability p.

Theorem 2.3 *Given $\epsilon > 0$ there exists $T = T(\epsilon)$ with the following property.*

Let H be a graph with $t > T$ vertices and with $\gamma(H) \geq \epsilon$. Let $\epsilon \leq p \leq 1 - \epsilon$, let $q = 1 - p$ and let $n \leq \lfloor \gamma(H)\, t\sqrt{\log_{1/q} t} \rfloor$. Then H is a minor of a random graph $G(n, p - \epsilon)$ with probability less than ϵ.

The next definition is self-explanatory but we state it nonetheless.

Definition 2.4 *The* density *of the graph G of order n is the ratio $|E(G)|/\binom{n}{2}$.*

Much of the work in the paper is devoted to proving the next theorem, which is the central one.

Theorem 2.5 *Given $\epsilon > 0$ there exists $T = T(\epsilon)$ with the following property.*

Let H be a graph with $t > T$ vertices and with $\gamma(H) \geq \epsilon$. Let $\epsilon \leq p \leq 1 - \epsilon$, let $q = 1 - p$ and let $n \geq \lfloor \gamma(H)\, t\sqrt{\log_{1/q} t} \rfloor$. Let G be a graph of order n, density $p + \epsilon$ and connectivity $\kappa(G) \geq n(\log\log\log n)/(\log\log n)$. Then H is a minor of G.

Remark. Theorems 2.3 and 2.5 show that the threshold probability p at which an H minor appears in $G(n,p)$ is the threshold density at which H minors appear in every reasonably connected graph of density p. This fact is at the heart of why Theorem 2.2 is true.

The remainder of the paper proceeds in the following way. In the next section we give the proofs of the main theorems. We then consider how the parameter $\gamma(H)$ can be evaluated, or estimated, for particular graphs. Finally we discuss briefly some related matters: in §5 the extremal graphs for $c(H)$ are described; in §6 the extremal problem for a set of graphs H, rather than just a single graph, is examined; and in §7 we discuss what might happen when H is very sparse or very asymmetric, and the main theorems do not apply.

3 Proofs of the main results

In this section we give the proofs of the theorems described above.

3.1 The proof of Theorem 2.3

Bollobás, Catlin and Erdős [3] investigated the largest value of t for which K_t is almost certainly a minor of $G(n,p)$. The part of the argument showing that K_t is not a minor of $G(n,p)$ for large t is straightforward, and a modification of it gives Theorem 2.3. First we state a very simple lemma.

Lemma 3.1 *Let* $0 < \epsilon \le x \le 1-\epsilon$. *Then* $\sqrt{(\log(x+\epsilon))/\log x} \le 1-\epsilon$.

Proof. We must show that $f(x,\epsilon) = (1-\epsilon)^2 \log x - \log(x+\epsilon) \le 0$ for $0 < \epsilon \le x \le 1-\epsilon$. Now the partial derivative of f with respect to x satisfies $f_x(x,\epsilon) = \epsilon[(1-\epsilon)^2 - (2-\epsilon)x]/x(x+\epsilon)$, which is positive or negative according as $x < x_0$ or $x > x_0$, where $x_0(\epsilon) = (1-\epsilon)^2/(2-\epsilon)$. It might be that $x_0 < \epsilon$ but even so, for fixed ϵ, the maximum value of $f(x,\epsilon)$ as x varies is at most $f(x_0,\epsilon)$, and so it suffices to show that $g(\epsilon) = f(x_0,\epsilon) = 2(1-\epsilon)^2\log(1-\epsilon) + \epsilon(2-\epsilon)\log(2-\epsilon) \le 0$ for $0 \le \epsilon \le 1/2$. The second and third derivatives of $g(\epsilon)$ are $g'' = 4\log(1-\epsilon) - 2\log(2-\epsilon) + 3 + 2/(2-\epsilon)$ and $g''' = -2(\epsilon^2 - 4\epsilon + 5)/(1-\epsilon)(2-\epsilon)^2$. Thus $g''' < 0$ and so $g''(\epsilon) \ge g''(1/2) > 0$ for $\epsilon \le 1/2$. Consequently the maximum value of $g(\epsilon)$ for $0 \le \epsilon \le 1/2$ is $\max\{g(0), g(1/2)\} = 0$. This completes the proof. □

Proof of Theorem 2.3. We shall derive a contradiction from the assumption that the probability is at least ϵ of H being a minor of $G(n,p)$. Now there are at most t^n partitions of the vertex set of $G(n,p)$ into t parts W_u, $u \in V(H)$, and subject to our assumption there is some partition for which the probability is at least ϵt^{-n} of there being an edge between W_u and W_v whenever $uv \in E(H)$. Fix such a partition. Then, putting $q_\epsilon = q + \epsilon$, we have

$$\epsilon t^{-n} \le \prod_{uv \in E(H)} (1 - q_\epsilon^{|W_u||W_v|}) \le \exp\Big(-\sum_{uv \in E(H)} q_\epsilon^{|W_u||W_v|}\Big),$$

so

$$\sum_{uv \in E(H)} q_\epsilon^{|W_u||W_v|} \le n\log t + \log\frac{1}{\epsilon} \le \frac{t}{\epsilon}(\log t)^{3/2} + \log\frac{1}{\epsilon} < t(\log t)^2$$

if t is large compared to ϵ.

Let $w(u) = \delta + |W_u|(\log_{1/q_\epsilon} t)^{-1/2}$ where $\delta = \epsilon^2/2$. Then

$$\sum_{uv \in E(H)} t^{-w(u)w(v)} \le \sum_{uv \in E(H)} t^{-\delta^2} q_\epsilon^{|W_u||W_v|} \le t^{-\delta^2} t(\log t)^2 \le t$$

if t is large compared to ϵ. Moreover, since $\sum_{u \in H} |W_u| = n$, we have

$$\frac{1}{t} \sum_{u \in H} w(u) = \delta + \sum_{u \in H} \frac{|W_u|}{t\sqrt{\log_{1/q_\epsilon} t}} \le \delta + \gamma(H) \sqrt{\frac{\log q_\epsilon}{\log q}} \le \delta + \gamma(H)(1-\epsilon)$$

by Lemma 3.1, and because $\gamma(H) \ge \epsilon$ this means $(1/t)\sum_{u \in H} w(u) \le \gamma(H) - \delta$. The weighting w thus provides a contradiction to the definition of $\gamma(H)$ and so completes the proof. □

3.2 The proof of Theorem 2.5

The main work of the paper is to prove the obverse to Theorem 2.3, namely Theorem 2.5. Before we begin that work, though, we need some preparatory results. The first is a simple Chernoff-type bound on the tails of the binomial and hypergeometric distributions (see for example [2]).

Lemma 3.2 *Let X be a random variable, distributed either binomially with parameters (n, p) or hypergeometrically with parameters (n, N, pN). Then, for $0 < \epsilon < 1$,*

$$\Pr\{\, |X - np| \ge \epsilon np \,\} \le 2\exp(-\epsilon^2 np/4)\,.$$

The next two very simple lemmas are taken from [18].

Lemma 3.3 ([18, Lemma 4.1]) *Given a bipartite graph with vertex classes A and B, wherein each vertex of A has at least $\delta|B|$ neighbours in B ($\delta > 0$), then there exists a set $M \subset B$ such that every vertex in A has a neighbour in M, and $|M| \le \lfloor \log_{1/(1-\delta)} |A| \rfloor + 1$.*

Lemma 3.4 ([18, Lemma 4.2]) *Let G be a connected graph and let $u, v \in V(G)$. Then u and v are joined in G by at least $\kappa^2(G)/4|G|$ internally disjoint paths of length at most $2|G|/\kappa(G)$.*

Using these lemmas we can show that, when looking for an H minor in a graph, the connectivity of the parts $G[W_u]$ can easily be taken care of. The lemma and its proof are adapted from [18, Theorem 4.1].

Lemma 3.5 *There is a number T with the following property. Let $t > T$. Let G be a graph of order $n \ge t(\log t)^{1/4}$ whose connectivity $\kappa(G)$ is at least $n(\log\log\log n)/(\log\log n)$. Then G contains a set C with $|C| \le n/\log\log\log n$ having the property that, for any partition W_i, $1 \le i \le t$, of $V(G) - C$ there are subsets W_i', $1 \le i \le t$ of C with $G[W_i \cup W_i']$ connected, $1 \le i \le t$.*

Proof. Let $k = (\log\log n)/(\log\log\log n)$, so $\kappa(G) \ge n/k$. By Lemma 3.4 we can find, for each pair u, v of vertices of G, a set $Q(u,v)$ of at least $n/4k^2$ internally disjoint u–v paths of length at most $2k$.

Let $r = 4\log\log\log n$ and choose a subset $C_1 \subset V(G)$ by selecting vertices independently at random with probability $1/r$. Then $|C_1| \le 2n/r$ holds with probability greater than $1/2$. Given two vertices u, v of G, the probability that the internal vertices of a path in $Q(u,v)$ all lie in C_1 is at least r^{-2k}, and these probabilities are independent for the different paths. By Lemma 3.2, the probability of fewer than $n/8k^2r^{2k}$ paths lying entirely within C_1 is at most $2\exp(-n/64k^2r^{2k}) < n^{-3}$ if t is sufficiently large (which from now on we assume without comment). So the probability that this happens for some pair u, v is less than $1/n$. Similarly, since every vertex has degree at least $\kappa(G)$, the probability that some vertex has more than half its neighbours inside C_1 is at most $2n\exp(-\kappa/16r) < 1/n$. So there is some choice of C_1 with $|C_1| \le 2n/r$, such that for every pair u, v of vertices there are at least $n/8k^2r^{2k}$ u–v paths whose internal vertices are disjoint and lie within C_1, and moreover C_1 contains no more than half the neighbours of any vertex. Make such a choice and fix it.

Now choose within $V(G) - C_1$ another subset C_2, again picking vertices with probability $1/r$, so that $|C_2| \le 2n/r$ with probability more than $1/2$. Each vertex $v \in G$ has at least $\kappa(G)/2 \ge n/2k$ neighbours in $V(G)-C_1$, and, similarly to hefore, the probability that some vertex has less than $n/4kr$ neighbours in C_2 is less than $1/2$. So there is a choice of C_2 that contains at least $n/4kr$ neighbours of every vertex, with $|C_2| \le 2n/r$. Make such a choice and fix it.

Let $C = C_1 \cup C_2$. We shall show that C has the property claimed in the lemma. Clearly $|C| \le 4n/r = n/\log\log\log n$. Let $n = tl$, where $l \ge (\log n)^{1/5}$, and let W_i, $1 \le i \le t$, be a partition of $V(G) - C$. We first find disjoint subsets $M_i \subset C_2$ so that every vertex in W_i has a neighbour in M_i and such that $|M_i| \le 8kr\log|W_i| + 1$; to see that this can be done, observe that if we have so far found $M_1, \dots, M_j$ then, because of the concavity of the log function, we have

$$\sum_{i=1}^{j} |M_i| \;\le\; \sum_{i=1}^{t} |M_i| \;\le\; t + 8kr\sum_{i=1}^{t}\log|W_i| \;\le\; t + 8krt\log l \;\le\; n/8kr\,,$$

so every vertex of W_{j+1} has at least $n/8kr$ neighbours in $C_2 - M_1 - \ldots - M_j$, and we can find M_{j+1} by applying Lemma 3.3 with $A = W_{j+1}$, $B = C_2 - M_1 - \ldots - M_j$ and $\delta = 1/8kr$.

Finally, for each i, $1 \le i \le t$, pick $u_i \in M_i$ and, for each $v \in M_i - \{u_i\}$, find a u_i–v path of length at most $2k$ whose internal vertices lie inside C_1. Let N_i be the collection of the internal vertices of the u_i–v paths. To see that the sets N_i can be chosen disjointly, notice that to begin with there are at least $n/8k^2r^{2k}$ suitable and disjoint choices for the u_i–v path, and that the number of these paths intersecting any previously chosen paths is less than

$$2k\sum_{i=1}^{t}(|M_i| - 1) \;\le\; 2k(8krt\log l) \;\le\; n/8k^2r^{2k}\,,$$

so we can always pick a path that meets none of the previous paths.

Hence we obtain disjoint sets $W_i' = (M_i \cup N_i) \subset C$, $1 \le i \le t$, such that $G[W_i']$ is connected and every vertex of W_i has a neighbour in W_i'. This means that $G[W_i \cup W_i']$ is connected, as claimed. □

Now we are ready to do the main work of the paper.

Proof of Theorem 2.5. We shall assume throughout that t is sufficiently large (in terms of ϵ) for the various assertions that we make to be true.

By the definition of $\gamma(H)$ we may take a weighting w of H with $\gamma(H) = (1/t)\sum_{u\in H} w(u)$ and $\sum_{uv\in EH} t^{-w(u)w(v)} \le t$. We shall find the sets W_u that we need for the minor by choosing them randomly, such that $|W_u|$ is proportional to $w(u)$. The difficulty with this approach is that there will not always be an edge between W_u and W_v whenever $uv \in E(H)$. However, by arguing carefully (as we describe later) it can be shown that the "missing" edges are incident with few vertices, say fewer than βt where β is some small constant. So, if instead of looking for an H minor we look in the first place for an $H + K_{\beta t}$ minor, we shall find an $H+K_{\beta t}$ minor with fewer than βt vertices missing, and so we shall have our H minor.

Let $\beta = \epsilon^4/32$ and let $J = H + K_{\beta t}$. First, remove from G a set C of size at most $n/\log\log\log n$ as given by Lemma 3.5. It will be convenient to add a few vertices to C until $|V(G) - C|$ is a multiple of $m = \lfloor(\log t)^{1/8}\rfloor$; this can be done so that $|C| < 2n/\log\log\log n$, and of course C still retains the properties asserted by Lemma 3.5.

Now let

$$Z \;=\; \{\, u \in V(H) \,:\, w(u) + \epsilon\gamma(H)/4 > 8/\epsilon^2 \,\}\,.$$

Put $q_\epsilon = q - \epsilon/2$ and for each $u \in J$ let

$$l(u) \;=\; \begin{cases} \left\lceil (w(u) + \epsilon\gamma(H)/4)\sqrt{\log_{1/q_\epsilon} t}\;\right\rceil & \text{if } u \in V(H) - Z \\ \left\lceil (8/\epsilon^2)\sqrt{\log_{1/q_\epsilon} t}\;\right\rceil & \text{if } u \in Z \cup V(K_{\beta t})\,. \end{cases}$$

Observe that, since $\gamma(H) \ge \epsilon$, there are constants c_1 and c_2 depending only on ϵ, such that $c_1\sqrt{\log t} < l(u) < c_2\sqrt{\log t}$ for each $u \in J$; we abbreviate this to $l(u) = \Theta_\epsilon(\sqrt{\log t})$. Moreover, using $\gamma(H) \ge \epsilon$ and Lemma 3.1 with $x = q_\epsilon$, we obtain

$$\begin{aligned}
\sum_{u\in J} l(u) \;&\le\; \left\{\sum_{u\in H}\Big(w(u) + \frac{\epsilon}{4}\gamma(H)\Big) + \beta\frac{8}{\epsilon^2}t\right\}\sqrt{\log_{1/q_\epsilon} t} \;+\; |J| \\
&\le\; (1+\epsilon/2)\,\gamma(H)t\sqrt{\log_{1/q_\epsilon} t} + (1+\beta)t \\
&\le\; (1+\epsilon/2)(1-\epsilon/2)\,\gamma(H)t\sqrt{\log_{1/q} t} + (1+\beta)t \\
&\le\; n - |C|\,.
\end{aligned}$$

It is therefore possible to choose disjoint sets $W_u \subset V(G) - C$ with $|W_u| = l(u)$ for each $u \in J$. Let such a choice be made at random.

Given a vertex $x \in V(G) - C$ let

$$Q(x) = \{\, y \in V(G) - C \,:\, xy \notin E(G) \,\} \quad \text{and} \quad q(x) = |Q(x)|/(n - |C|) \;;$$

note that $x \in Q(x)$. For a subset $S \subset V(G) - C$ we write $Q(S) = \cap_{y \in S} Q(y)$, so $x \in Q(S)$ if and only if $S \subset Q(x)$. Then the average value of $q(x)$ satisfies

$$\begin{aligned} \frac{1}{n - |C|} \sum_{x \in V(G)-C} q(x) &= \frac{1}{(n - |C|)^2} \sum_{x \in V(G)-C} |Q(x)| \\ &\le \frac{1}{(n-|C|)^2} \left(n - |C| + (q - \epsilon)n(n-1)\right) \\ &\le (q - \epsilon)\left(1 + \frac{5}{\log\log\log t}\right) \le q_\epsilon \,. \end{aligned}$$

Now label the set

$$V(G) - C \,=\, \{x_1, \ldots, x_{n-|C|}\} \quad \text{where } q(x_i) \text{ increases with } i\,.$$

There are many sets W_u, each small, and some of the choices will almost inevitably be poor. The choices that work turn out to be those where both W and $Q(W)$ are well-distributed through the ordered set $V(G)-C$. To make this precise, split the vertices of $V(G)-C$ into m consecutive blocks B_j, $1 \le j \le m$, each of size $b = (n - |C|)/m$; so

$$B_j \,=\, \{\, x_i \,:\, (j-1)b \,<\, i \,\le\, jb \,\}\,.$$

We also define $q_j = \max\{q(x) : x \in B_j\} = q(x_{jb})$. We now say that a set W is *good* if

$$\begin{aligned} &\text{both} \quad \text{(a)} \qquad |W \cap B_j| \,\ge\, |W|/m - (\log t)^{1/3} \\ &\text{and} \quad \text{(b)} \qquad |Q(W) \cap B_j| \,\le\, (\log t)\, b\, q_j^{|W|} \end{aligned}$$

hold for all j, $1 \le j \le m$.

Let us check that most sets W_u, $u \in J$, are good. Firstly, by Lemma 3.2, the probability that $|W_u \cap B_j| \le l(u)/m - (\log t)^{1/3}$ for some given j is at most $\exp\{-(\log t)^{1/4}\} < \beta/4m$ since $l(u) = \Theta_\epsilon(\sqrt{\log t})$, so the probability that W_u fails condition (a) of goodness is at most $\beta/4$. Secondly, if $x \in B_j$ then

$$\Pr\{W_u \subset Q(x)\} = \binom{|Q(x)|}{l(u)}\binom{n-|C|}{l(u)}^{-1} \le \left(\frac{|Q(x)|}{n-|C|}\right)^{l(u)} = q(x)^{l(u)} \le q_j^{l(u)}$$

so the expected size of $|Q(W_u) \cap B_j|$ is at most $bq_j^{l(u)}$. Hence the probability that $|Q(W_u) \cap B_j|$ exceeds $(\log t)\, b\, q_j^{l(u)}$ is at most $1/\log t$, and so the probability of W_u failing condition (b) of goodness is at most $m/\log t < \beta/4$. We conclude that W_u fails to be good with probability at most $\beta/2$.

We now need to check that the probability of there being no edge between W_u and W_v is not much worse than what it would be in a random graph, namely

$q^{l(u)l(v)}$, provided both W_u and W_v are good. Let $P(u,v)$ be the probability that there is no W_u–W_v edge conditional on W_u satisfying condition (a) and W_v satisfying condition (b). Now there is an edge between W_u and W_v unless $W_u \cap B_j$ is contained in $Q(W_v) \cap B_j$ for $1 \le j \le m$. Let $P'(u,v)$ be the corresponding conditional probability in the experiment where W_u and W_v are chosen independently (not necessarily disjointly); then $P(u,v) \le P'(u,v)/P_{\mathrm{d}}$ where P_{d} is the probability that W_u and W_v are disjoint. Note that

$$P_{\mathrm{d}}^{-1} = \binom{n-|C|}{l(u)}\binom{n-|C|-l(v)}{l(u)}^{-1} \le \left(\frac{n-|C|-l(u)}{n-|C|-l(u)-l(v)}\right)^{l(u)} \le 2\,.$$

Given that W_u satisfies condition (a), choose at random a set $A_j \subset W_u \cap B_j$ of size $k = l(u)/m - (\log t)^{1/3}$; then the sets A_j are uniformly distributed in B_j, and $W_u \cap B_j$ is contained in $Q(W_v) \cap B_j$ only if $A_j \subset Q(W_v) \cap B_j$. Ignoring the constraint imposed here when $j = m$, we thus have

$$\begin{aligned}
P(u,v) &\le P'(u,v)\, P_{\mathrm{d}}^{-1} \\
&\le P_{\mathrm{d}}^{-1} \prod_{j=1}^{m-1} \binom{|Q(W_v)\cap B_j|}{k}\binom{b}{k}^{-1} \le P_{\mathrm{d}}^{-1} \prod_{j=1}^{m-1} \left(\frac{|Q(W_v)\cap B_j|}{b}\right)^{k} \\
&\le 2(\log t)^{m-1} \left(\prod_{j=1}^{m-1} q_j\right)^{kl(v)} \qquad \text{since } W_v \text{ satisfies condition (b)} \\
&\le (\log t)^m \left(\frac{1}{m-1}\sum_{j=1}^{m-1} q_j\right)^{(m-1)kl(v)} \qquad \text{by the AM-GM inequality} \\
&\le (\log t)^m \left(\frac{1}{b(m-1)}\sum_{i=1}^{n-|C|} q(x_i)\right)^{(m-1)kl(v)} \qquad \text{as } q() \text{ is monotone} \\
&\le (\log t)^m \left(\frac{n-|C|}{b(m-1)}\, q_\epsilon\right)^{(m-1)kl(v)} \qquad \text{by our earlier estimate} \\
&= (\log t)^m \left(\frac{m}{m-1}\, q_\epsilon\right)^{(m-1)kl(v)} \\
&\le (\log t)^m \left(\frac{m}{m-1}\, q_\epsilon\right)^{(l(u)-(\log t)^{1/24})l(v)} \qquad \text{as } l(u),\, l(v) = \Theta_\epsilon(\sqrt{\log t}) \\
&\le t^\beta\, q_\epsilon^{l(u)l(v)} \qquad \text{as } m = \lfloor (\log t)^{1/8} \rfloor \text{ and } l(u),\, l(v) = \Theta_\epsilon(\sqrt{\log t})
\end{aligned}$$

It follows from this that if both u and v are in $V(H) - Z$ then $P(u,v) \le t^\beta t^{-w(u)w(v)-\epsilon^4/16} \le (\beta/4)t^{-w(u)w(v)}$ whereas otherwise, since $\min\{l(u), l(v)\} \ge (\epsilon^2/4)\sqrt{\log_{1/q_\epsilon} t}$, we have $P(u,v) \le t^\beta t^{-2} \le (\beta/5)t^{-1}$. Therefore the expected number of edges $uv \in E(J)$, for which both W_u and W_v are good but there is no W_u–W_v edge, is at most

$$\sum_{uv \in E(J)\,:\,u,v \text{ good}} P(u,v) \;\le\; \sum_{uv \in H} \frac{\beta}{4}\, t^{-w(u)w(v)} \;+\; (1+\beta)^2\, t^2\, \frac{\beta}{5}\, t^{-1} \;\le\; \frac{\beta}{2} t\,.$$

Consequently there exist sets W_u, $u \in V(J)$, for which W_u is good for all but $\beta t/2$ vertices u, and amongst the pairs of vertices for which both W_u and W_v are good there is a W_u–W_v edge in G whenever $uv \in E(J)$ except in $\beta t/2$ cases. Remove from $V(J)$ the vertices u for which W_u is not good or for which some edge uv is not represented by a W_u–W_v edge. Since $J = H + K_{\beta t}$ and at most βt vertices were removed, the remaining graph contains a copy of H.

Therefore $V(G) - C$ contains subsets W_u, $u \in V(H)$, such that G has a W_u–W_v edge whenever $uv \in E(H)$. Finally, by Lemma 3.5, we can find disjoint subsets $W'_u \subset C$ such that $G[W_u \cup W'_u]$ is connected — in other words, G has an H minor. □

Remark. The statement and proof of Theorem 2.5 involve only fixed values of ϵ, with t being sufficiently large. It is possible to pursue the proof when both $\gamma(H)$ and ϵ are functions of t that tend to zero as $t \to \infty$, but not too quickly. One of the principal constraints arises from the removal of the set C of size $n/\log\log\log n$, which perturbs the value of q. We could improve matters by noticing that C was chosen randomly in the proof of Lemma 3.5, and so its removal would not perturb q too much. But the need to use Lemma 3.5, or some version of it, nevertheless means that $\gamma(H)$ can tend to zero only extremely slowly, and so we prefer, to avoid further complication, to give Theorem 2.5 only in the form stated.

3.3 The proof of Theorem 2.2

The arguments needed to derive Theorem 2.2 from Theorems 2.3 and 2.5 are almost identical to those used in the case $H = K_t$, and can be taken almost verbatim from [18]. We therefore give just an outline, indicating those points at which the constants need to be changed slightly.

Theorem 2.2 is equivalent to the following statement: given $\delta > 0$, there exists $T = T(\delta)$, such that if $|H| = t > T$ then $(\alpha\gamma(H) - \delta)t\sqrt{\log t} \le c(H) \le (\alpha\gamma(H)+\delta)t\sqrt{\log t}$. The lower bound follows at once if $\gamma(H) < \delta$, and otherwise it follows by applying Theorem 2.3 with $p = 1 - \lambda$ (as defined in §2) to obtain a graph with no H minor and (almost certainly) average degree at least $(1 - \delta)(1-\lambda)\gamma(H)t\sqrt{\log_{1/\lambda} t} = 2(1-\delta)\alpha\gamma(H)t\sqrt{\log t}$.

To get the upper bound we must show that any graph with $d|G|$ edges contains an H minor, where $d = \lfloor(\alpha\gamma(H) + \delta)t\sqrt{\log t}\rfloor$. We may assume that $\gamma(H) \ge \delta$ (if not, add edges to H until this holds). Putting $k = \lceil d/\log\log\log d\rceil$, we define the class of graphs $\mathcal{E}_{d,k}$ by

$$\mathcal{E}_{d,k} \;=\; \{\, G \,:\, |G| \ge d \text{ and } e(G) > d|G| - kd \,\}\;.$$

It suffices then to show that, if G is in $\mathcal{E}_{d,k}$ but no minor of G is, then G has an H minor. Such a graph G has the properties that $e(G) = d|G| - kd + 1$, $d+1 \le \delta(G) \le 2d-1$, every edge of G lies in at least d triangles, and $\kappa(G) > k$; the simple verification of this can be found in [18, §2].

At this point we split the proof into two cases according to whether or not $|G| > cd$, where $c = c(\delta)$ is a constant given by Proposition 3.7 below. This proposition is itself based on small preliminary proposition about minors in bipartite graphs. The case of the preliminary proposition where $\delta = 3/10$ and $\zeta = 400$ is Lemma 5.1 of [18]; the same proof works for any values of δ and ζ.

Proposition 3.6 *Given $\delta > 0$ and $\zeta > 0$ there exist constants t_0 and C with the following property. Let $t > t_0$ and let $d \ge \delta t\sqrt{\log t}$. Let G be a bipartite graph with vertex classes A and B, such that $|A| > Cd$, $|B| < \zeta d$ and every vertex of A is incident with at least $d/3$ edges. Then $G \succ K_{2t}$.*

Proposition 3.7 *Given $\delta > 0$, there are constants $c = c(\delta)$ and $T = T(\delta)$ such that the following holds.*

Let $t > T$ be an integer and let $d \ge \delta t\sqrt{\log t}$. Let G be a graph with $|G| \ge cd$ and $\kappa(G) \ge t \log\log t$. Suppose that $e(G) \le d|G|$ and that there are at least d triangles on every edge of G. Then $G \succ K_{2t}$.

Proof. (**Sketch only.**) The proof of the proposition is almost the same as that of [18, Theorem 5.1], except that there the bound on d is $d \ge (3/10)t\sqrt{\log t}$ and the bound on κ is the weaker $\kappa(G) \ge 23t$. The argument given in [18] is roughly this. We aim first to find $67 = 1 + \binom{12}{2}$ disjoint $K_{t/3}$ minors. Suppose we have so far found some of them. Since the average degree of G is at most $2d$ we can find lots of vertices of degree at most $3d$. If lots of these have more than $d/3$ neighbours amongst the previously chosen $K_{t/3}$ minors, we get a dense bipartite graph, which by [18, Lemma 5.1] has a K_{2t} minor (here we need c to be large, so that the lemma can be applied). Otherwise, we find a vertex of degree at most $3d$ with at most $d/3$ neighbours in the previously chosen minors, and since each edge of G lies in d triangles, the part of the neighbourhood lying outside the previous minors is sufficiently dense to provide a further $K_{t/3}$ minor (by Theorem 2.5). Hence we obtain our 67 $K_{t/3}$ minors. We now arrange the parts of 66 of the $K_{t/3}$ minors into $2t$ groups of 11, in such a way that, given any two groups, there is some $K_{t/3}$ minor from which they both have a part, and hence there is some edge between the two groups. To get our K_{2t} minor, all we have to do now is to find paths internally linking each group so that the 11 parts in the group form a connected subgraph. These paths can easily be found by taking a path from each of the $22t$ parts to the remaining $K_{t/3}$ minor that we didn't use yet, and using this minor as a router to connect the ends of the paths in the way we want.

To prove the present proposition, all that is needed is an adjustment to the constants. Because the constant $3/10$ of [18, Theorem 5.1] has been replaced by δ we can no longer find $K_{t/3}$ minors but only $K_{\beta t}$ minors for some constant $\beta = \beta(\delta)$. This means that we must look for $1 + \binom{2\lceil 1/\beta \rceil}{2}$ of these minors. The argument then proceeds much as before. Proposition 3.6 is applied instead of Lemma 5.1 of [18], and instead of $\zeta = 400$ we take $\zeta = 6 \times \binom{2\lceil 1/\beta \rceil}{2}$. The value of c needed is then proportional to the C required by Proposition 3.6. As before, we either find a K_{2t} minor at once or we keep finding small dense subsets yielding fresh $K_{\beta t}$ minors. The argument that uses one of the small minors and the large connectivity to link the other small minors to form a K_{2t} minor needs barely any adjustment, and in fact the higher connectivity in the present case simplifies some of the estimates.

These remarks should suffice to guide the keen reader wishing to fill in the details of the proof; the less determined but nevertheless interested reader can find the details in [13]. □

Resuming the proof of the upper bound for Theorem 2.2, we see that if our minimal graph G satisfies $|G| > cd$ then it satisfies the conditions of Proposition 3.7, so $G \succ K_{2t}$ and we are done. Otherwise, put $\epsilon = \min(\delta, 1/c)$ so that the density of G is $p = e(G)/\binom{|G|}{2} \sim 2d/|G| \geq 2\epsilon$. Then we have

$$|G| \sim \frac{2d}{p} = \frac{2d}{1-q} \geq (1+\delta)\,\frac{2\alpha}{1-q}\sqrt{\log(1/q)}\,\gamma(H)\,t\,\sqrt{\log_{1/q} t}$$

and, since the maximum value of the expression $(1-q)/2\sqrt{\log(1/q)}$ is α (when $q = \lambda$), it follows that $|G| \geq (\gamma(H) + \epsilon)t\sqrt{\log_{1/q} t}$. Theorem 2.5 then tells us that $G \succ H$.

4 Estimating $\gamma(H)$

We now address the question of how to compute, or to estimate, the parameter $\gamma(H)$.

4.1 Exact evaluation

It appears to be difficult in general to evaluate $\gamma(H)$ exactly. In the case when H is regular it seems likely that the optimal weighting is uniformly distributed on the vertices, but we cannot prove this; in fact, we can evaluate $\gamma(H)$ exactly only when H is complete or complete bipartite. When $H = K_t$ it is easily checked that an optimal weighting gives the same weight to each vertex, from which it follows that $\binom{t}{2}t^{-\gamma^2(K_t)} = t$, so $\gamma(K_t) = 1 - O\left(1/\log t\right)$.

In general a couple of simple observations might be made. First, note that if H' is a spanning subgraph of H then $\gamma(H') \leq \gamma(H)$, since a weighting that is acceptable for H is acceptable also for H'.

A second observation is that, if u and v are non-adjacent and have the same neighbours, then (by the AM-GM inequality) the quantity $\sum_{uv \in E(H)} t^{-w(u)w(v)}$ is not increased if we assign weight $(w(u) + w(v))/2$ to each of u and v. In other words, in an optimal weighting, non-adjacent vertices having the same neighbours receive the same weight. So, for example, when $H = K_{\beta t,(1-\beta)t}$ an optimal weighting has only two distinct weights w_1 and w_2, such that $\beta w_1 + (1-$

$\beta)w_2 = \gamma(H)$ and $\beta(1-\beta)t^2 t^{-w_1 w_2} = t$. The optimum is $\beta w_1 = (1-\beta)w_2$, giving $\beta(1-\beta)t^{1-\gamma(H)^2/4\beta(1-\beta)} = 1$, so $\gamma(K_{\beta t,(1-\beta)t}) = 2\sqrt{\beta(1-\beta)} + O\left(1/\log t\right)$.

4.2 Shapes

Since we cannot evaluate $\gamma(H)$ exactly we try to approximate it, and for this purpose the next definition will be useful.

Definition 4.1 *A* shape *is a pair (F, f), where F is a graph (in which loops, but not multiple edges, are allowed) and $f : V(F) \to \mathbf{R}^+$ is a function assigning non-negative numbers to the vertices such that $\sum_{a\in V(F)} f(a) = 1$. The parameter $m(F, f)$ is given by*

$$m(F,f) \;=\; \max_{x\cdot f=1}\; \min_{ab\in E(F)}\; x(a)x(b)\,.$$

Here the maximum is over all functions $x \in [0,\infty)^{V(F)}$ of $V(F)$, and $x\cdot f$ stands for the standard inner product $\sum_{a\in F} x(a)f(a)$.

Remark. Notice that this definition allows $x(a) > 1$ even though we always have $f(a) \le 1$. Notice too that the constant function $x(a) = 1$ satisfies $x \cdot f = 1$ and so $m(F, f) \ge 1$. Also, if F consists of a single vertex a with a loop then $f(a) = 1$ and $m(F, f) = 1$.

We shall approximate $\gamma(H)$ by means of small shapes. On the face of it, this appears to be just substituting one weighted graph problem for another. But it turns out that we can restrict our attention to a small identifiable class of shapes, and for these shapes we can evaluate $m(F, f)$. Before developing this remark, though, we explain how $\gamma(H)$ can be estimated in terms of shapes.

Definition 4.2 *The graph H of order t is an ϵ-*fit *to shape (F, f) if there is a partition of $V(H)$ into sets V_a, $a \in V(F)$, such that $|V_a| = f(a)t$, and*

$$|\{\, uv \in E(H)\, : u \in V_a,\, v \in V_b \text{ and } ab \notin E(F)\,\}| \;\le\; t^{-\epsilon}\, |E(H)|\,.$$

So H is an ϵ-fit to (F, f) if there is a partition of H into classes indexed by $V(F)$ and of sizes proportional to f, so that all but a tiny fraction of the edges of H lie between classes corresponding to edges of F. The fact that F might have loops allows H to have edges within the corresponding classes.

The relevance of shapes to the calculation of $\gamma(H)$ can now be made explicit.

Theorem 4.3 *Let $\epsilon > 0$ and let H be a graph of order t with $t^{1+\tau}$ edges. Then H is an ϵ-fit to some shape (F, f) with*

$$|F| \le \frac{1}{\epsilon} \qquad \text{and} \qquad \gamma(H) \;\ge\; \frac{\sqrt{\tau}}{\sqrt{m(F,f)}} \;-\; 4\sqrt{\epsilon}\,.$$

Proof. The shape with one vertex satisfies the demands of the lemma if $\epsilon \geq 1/16$, so may assume $\epsilon < 1/16$. Let $\delta = 2\sqrt{\epsilon}$ and let w be a weighting of H satisfying $(1/t)\sum_{u\in H} w(u) = \gamma(H)$ and $\sum_{uv\in E(H)} t^{-w(u)w(v)} \leq t$. Define a new weighting w' of H by

$$w'(u) \;=\; \begin{cases} w(u)+\delta & \text{if } |H| \leq 1/\epsilon \\ \min\{w(u)+\delta,\, 3/\delta\} & \text{if } |H| > 1/\epsilon\,. \end{cases}$$

Then if $|H| \leq 1/\epsilon$ we have

$$\sum_{uv\in E(H)} t^{-w'(u)w'(v)} \;\leq\; \sum_{uv\in E(H)} t^{-w(u)w(v)-\delta^2} \;\leq\; t^{1-\delta^2} \;\leq\; t^{1-\epsilon}$$

whereas, since $w'(u)w'(v) \geq 3$ whenever $w'(u) = 3/\delta$, if $|H| > 1/\epsilon$ we have

$$\sum_{uv\in E(H)} t^{-w'(u)w'(v)} \;\leq\; \sum_{uv\in E(H)} t^{-w(u)w(v)-\delta^2} + t^{1+\tau}t^{-3} \;\leq\; t^{1-4\epsilon} + t^{-1} \;\leq\; t^{1-\epsilon}\,.$$

For each integer i let

$$A_i \;=\; \left\{\, u\in H \;:\; (1+\delta)^i \;<\; w'(u) \;\leq\; (1+\delta)^{i+1} \,\right\}$$

so the A_i form a partition of $V(H)$.

We define a shape (F,f) in the following way. Let the vertices of the graph F be $V(F) = \{A_i \,:\, A_i \neq \emptyset\}$. Certainly $|F| \leq 1/\epsilon$ if $|H| \leq 1/\epsilon$. On the other hand, if $|H| > 1/\epsilon$ then $\delta \leq w'(u) \leq 3/\delta$ for all $u \in H$ so $A_i = \emptyset$ for $i+1 < \log_{1+\delta}\delta$ or $i \geq \log_{1+\delta}(3/\delta)$. So in this case $|F| \leq 2 + \log_{1+\delta}\delta + \log_{1+\delta}(3/\delta) < 1/\epsilon$ also. Let $f(A_i) = |A_i|/t$, so $\sum_{a\in V(F)} f(a) = (1/t)\sum_{A_i\neq\emptyset}|A_i| = 1$. Finally, for $a = A_i \in V(F)$ put $y(a) = (1+\delta)^{1+i}$. Then the edges of F are defined by

$$E(F) \;=\; \{\, ab \,:\, a\in F,\, b\in F \;\text{ and }\; y(a)y(b) \geq \tau \,\}\;.$$

First we check that H is an ϵ-fit to (F,f). The partition V_a, $a \in F$ that we use is of course $V_a = a = A_i$, where by definition $|V_a| = |A_i| = f(a)t$. Let E be the set of edges $uv \in E(H)$ that do not correspond to edges of F; that is, if $uv \in E$ then there exist $a = A_i$ and $b = A_j$ in $V(F)$ with $u \in A_i$, $v \in A_j$ and $ab \notin E(F)$. This means that $w'(u)w'(v) \leq (1+\delta)^{i+1}(1+\delta)^{j+1} = y(a)y(b) < \tau$, so

$$|E|t^{-\tau} \;\leq\; \sum_{uv\in E} t^{-w'(u)w'(v)} \;\leq\; \sum_{uv\in E(H)} t^{-w'(u)w'(v)} \;\leq\; t^{1-\epsilon}$$

giving $|E| \leq t^{1+\tau-\epsilon} = t^{-\epsilon}|E(H)|$, as needed.

What remains is to verify the bound on $\gamma(H)$ claimed in the theorem. Writing $s = \sum_{a\in F} f(a)y(a)$ we have

$$s \;=\; \frac{1}{t}\sum_{A_i\neq\emptyset}|A_i|(1+\delta)^{1+i} \;\leq\; \frac{1}{t}\sum_{u\in H} w'(u)(1+\delta) \;\leq\; \frac{1+\delta}{t}\sum_{u\in H}(w(u)+\delta)$$

giving $s \le (1+\delta)(\gamma(H)+\delta)$. Let $x(a) = y(a)/s$; then we have

$$x \cdot f = \sum_{a\in F} f(a)x(a) = (1/s) \sum_{a \in F} f(a)y(a) = 1\,.$$

Thus the definition of $m(F,f)$ implies $\min_{ab\in E(F)} x(a)x(b) \le m(F,f)$. But the definition of $E(F)$ implies $\min_{ab\in E(F)} x(a)x(b) = (1/s^2)\min_{ab\in E(F)} y(a)y(b) \ge \tau/s^2$, and so

$$(1+\delta)(\gamma(H)+\delta) \;\ge\; s \;\ge\; \frac{\sqrt{\tau}}{\sqrt{m(F,f)}}\,,$$

from which we obtain

$$\gamma(H) \;\ge\; \frac{1}{1+\delta}\frac{\sqrt{\tau}}{\sqrt{m(F,f)}} - \delta \;\ge\; (1-\delta)\frac{\sqrt{\tau}}{\sqrt{m(F,f)}} - \delta \;\ge\; \frac{\sqrt{\tau}}{\sqrt{m(F,f)}} - 2\delta$$

which is the result claimed. □

4.3 Critical shapes and half-graphs

We make use of Theorem 4.3 to obtain a lower bound on $\gamma(H)$ in the following way. We know $\gamma(H) \le \sqrt{\tau}$. Suppose we wish to show $\gamma(H) \ge \sqrt{\tau/m}$ for some $m > 1$. Given $\epsilon > 0$ there is a bounded number of graphs F with $|F| \le 1/\epsilon$; if we can show that, whenever an f can be assigned to such an F so that H is an ϵ-fit to (F,f), then $m(F,f) < m$, we will have shown $\gamma(H) \ge \sqrt{\tau/m} - 4\sqrt{\epsilon}$, which is more or less what we wanted.

Observe that, if H is an ϵ-fit to (F,f) then it is also an ϵ-fit to any shape (F',f') where F' is obtained from F either by the addition of an edge or by the *merger* of two vertices of F ; by the merger of $a,b \in F$ we mean the replacement of a and b by a single vertex c joined to every vertex previously joined to either a or b (including a loop at c if F had an edge with both ends in $\{a,b\}$), with $f'(c) = f(a)+f(b)$ and $f' = f$ on the other vertices of F'.

Definition 4.4 *A shape (F,f) is said to be* critical *if the addition of any edge to F or the merger of any two vertices decreases the value of $m(F,f)$.*

Let us summarize the discussion above.

Corollary 4.5 *Let H be a graph of order t with $t^{1+\tau}$ edges. Let $\epsilon > 0$ and let m be the maximum value of $m(F,f)$ over critical shapes (F,f) with $|F| \le 1/\epsilon$ to which H is an ϵ-fit. Then $\gamma(H) \ge \sqrt{\tau/m} - 4\sqrt{\epsilon}$.*

What makes this corollary useful is that there are very few critical shapes, and we can describe them exactly.

Remark. It should be pointed out, however, that the corollary can sometimes give a bound much less than the correct value of $\gamma(H)$. For example, consider when H is the union of $K_{t/8,7t/8}$ and a $t^{1/2}$-regular graph H^* on the same vertex set. As shown earlier, $\gamma(K_{t/8,7t/8}) = \sqrt{7}/4 + o(1)$, and as shown later in §4.4,

$\gamma(H^*) = 1/\sqrt{2} + o(1)$. Thus $\gamma(H) \geq \max(\sqrt{7}/4, 1/\sqrt{2}) + o(1) = 1/\sqrt{2} + o(1)$. But, for every $\epsilon > 0$, if t is large this graph is an ϵ-fit to a two vertex shape with $f = (1/8, 7/8)$ and $m(F,f) = 16/7$ (see the next theorem). So Corollary 4.5 gives only $\gamma(H) \geq \sqrt{7}/4 + o(1)$.

In other words, Corollary 4.5 is blind to the presence of the very sparse subgraph H^*, though this subgraph is what is determining the actual value of $\gamma(H)$. This situation is, of course, analogous to the situation in the classical extremal theory, where the chromatic number of H is governed by $\chi(H^*)$ and not just by the chromatic number of a dense subgraph.

Critical shapes are described precisely by the next theorem.

Theorem 4.6 *A shape (F,f) with $|F| = k+1$ is critical if and only if F is the* half-graph *of order $k+1$ that is,*

$$V(F) = \{0,1,\ldots,k\} \qquad \text{and} \qquad E(F) = \{\, ij \,:\, i+j \geq k \,\},$$

and moreover f satisfies

$$\frac{f(k)}{f(0)} < \frac{f(k-1)}{f(1)} < \frac{f(k-\lfloor (k-1)/2 \rfloor)}{f(\lfloor (k-1)/2 \rfloor)} < 1\,.$$

For these shapes,

$$m(F,f) = \left\{ \sum_{i=0}^{k} \sqrt{f(i)f(k-i)} \right\}^{-2}.$$

Proof. First suppose that (F,f) is critical, and that x is such that $x \cdot f = 1$ and $m(F,f) = \min_{ab \in E(F)} x(a)x(b)$. Say that an edge ab of F is *critical* if $x(a)x(b) = m(F,f)$.

Every vertex must be adjacent to a critical edge; for if a were not, then $x(a)$ could be slightly decreased and $x(b)$ slightly increased for all vertices $b \neq a$ to keep $x \cdot f = 1$, and so $m(F,f)$ would increase. No distinct vertices a and b can have $x(a) = x(b)$, for such vertices could be merged without affecting $m(F,f)$. This means that every vertex is adjacent to exactly one critical edge (for, if ab and ac were critical, we would have $x(b) = x(c)$), and that there is at most one critical loop (for, if aa and bb were critical, we would have $x(a) = x(b)$). So the critical edges form a perfect matching (plus a loop if $|F|$ is odd).

Each critical edge ab must have one endpoint a with $x(a) < \sqrt{m(F,f)}$, and the other b with $x(b) > \sqrt{m(F,f)}$, except that a critical loop aa must have $x(a) = \sqrt{m(F,f)}$. Let $V(F) = \{0,1,\ldots,k\}$ and let $0,1,\ldots,\lfloor (k-1)/2 \rfloor$ be the vertices with $x(i) < \sqrt{m(F,f)}$, in increasing order of $x(i)$. Let $k/2$ be the vertex of the critical loop, if k is even. Let i and $k-i$ be the endpoints of a critical edge, for $0 \leq i \leq \lfloor (k-1)/2 \rfloor$. Since $x(k-i)x(i) = m(F,f)$, it follows that $x(0) < x(v_1) < \cdots < x(k)$. Criticality of (F,f) means that each i has as neighbours exactly those j for which $x(i)x(j) \geq m(F,f)$; that is, $i+j \geq k$. Thus F is the half-graph as described.

We can now evaluate $m(F, f)$. Since ij is a critical edge, where $j = k - i$, we have $m(F, f) = x(i)x(j) = f(i)x(i)f(j)x(j)/f(i)f(j)$. If we hold $x(a)$ fixed for all $a \neq i, j$, then $x(i)$ and $x(j)$ may be varied such that $f(i)x(i) + f(j)x(j)$ remains constant. If $f(i)x(i) \neq f(j)x(j)$, we can make a slight variation so that $f(i)x(i)f(j)x(j)$ increases without any new critical edges being created; but now there is no critical edge at i or j, contradicting the fact that every vertex is adjacent to a critical edge (else $m(F, f)$ could be increased). Therefore $f(i)x(i) = f(j)x(j)$. This means that

$$x(i)f(i) \;=\; \sqrt{f(i)x(i)f(j)x(j)} \;=\; \sqrt{m(F,f)f(i)f(j)}$$

so $x(i) = \sqrt{m(F,f)f(j)/f(i)}$. The condition the theorem gives on the values of $f()$ now follows from $x(0) < x(1) < \cdots < x(\lfloor(k-1)/2\rfloor) < \sqrt{m(F,f)}$. Moreover, since $1 = x \cdot f = \sum_{i=0}^{k} x(i)f(i)$, we obtain the value of $m(F, f)$ claimed in the theorem.

Conversely, suppose now that (F, f) has the described form. We need to show that (F, f) is critical. First, we show that there is a unique function x on $V(F)$ such that $m(F, f) = \min_{ij \in E(F)} x(i)x(j)$. For take any x satisfying this equation. As argued at the start of the proof, each vertex must be incident with a critical edge. Now if $i < j$ then every neighbour of i is also a neighbour of j, and since i is incident with a critical edge it must be that $x(j) \geq x(i)$. Thus $x()$ is increasing, and so $i(k - i)$ must be a critical edge for each i. In particular, if k is even, then $x(k/2) = \sqrt{m(F,f)}$. Moreover, if $x(i) = x(i + l)$ then $x(k - i) = x(k - i - l)$ and so $x(i) = \cdots = x(i + l)$ and $x(k - i) = \cdots = x(k - i - l)$.

If $x()$ is not strictly increasing there are two vertices $i < j \leq \lceil k/2 \rceil$ with $x(i) = x(j)$, in which case let $i < k/2$ be the smallest such vertex, so i is in exactly one critical edge; otherwise, let i be any vertex with $i < k/2$. Now $m(F, f) = x(i)x(k-i) = f(i)x(i)f(k-i)x(k-i)/f(i)f(k-i)$. If $x(i)f(i) \neq f(k-i)x(k-i)$, we may move these two quantities closer together whilst keeping their sum, and so $x \cdot f$, constant. This increases the product $f(i)x(i)f(k - i)x(k - i)$ so the edge $i(k - i)$ is no longer critical. If $x()$ is strictly increasing this means i and $k - i$ are no longer in any critical edge; the same is true even if $x()$ is not strictly increasing if the change meant a decrease for $x(i)$ and an increase for $x(k - i)$, that is, if $x(i)f(i) > f(k - i)x(k - i)$, because i is in only one critical edge. In either of these cases we may then reduce both $x(i)$ and $x(k-i)$ slightly, increasing the remaining x values, and so increase $m(F, f)$. It must therefore be that $x(i)f(i) = f(k - i)x(k - i)$ if $x()$ is strictly increasing, and otherwise $x(i)f(i) \leq f(k - i)x(k - i)$. In the latter case, since $f(k - i) < f(i)$, we cannot have $x(i) = x(k - i)$, and so if i' is the maximum vertex with $x(i') = x(i)$ then $i < i' < k/2$. We may then apply a similar argument to $x(i')$ and $x(k - i')$ to find, since $k-i'$ is in exactly one critical edge, that $x(i')f(i') \geq f(k-i')x(k-i')$. But then

$$\frac{f(k-i)}{f(i)} \;\geq\; \frac{x(i)}{x(k-i)} \;=\; \frac{x(i')}{x(k-i')} \;\geq\; \frac{f(k-i')}{f(i')}$$

which contradicts the assumption about (F, f). We conclude that $x()$ is strictly increasing and that $x(i)f(i) = f(k - i)x(k - i)$ for all i. As before, this means that $x(i) = \sqrt{m(F,f)f(k-i)/f(i)}$ for all i, and so there is a unique x realizing $m(F,f)$.

Suppose (F,f) is not critical, and let (F',f') be some shape, obtained by adding some edge ab or by merging two vertices a and b into a new vertex c, such that $m(F',f') \geq m(F,f)$. Let x' be a function on $V(F')$ such that $m(F',f') = \min_{ij\in E(F')} x'(i)x'(j)$. Define the function x on $V(F)$ by $x = x'$ if F' is obtained by adding an edge to F, and otherwise put $x(a) = x(b) = x'(c)$ and $x(i) = x'(i)$ for $i \in V(F) - \{a,b\}$. Then $x \cdot f = 1$. Moreover if $ij \in E(F)$ then, in almost every case, $ij \in E(F')$ and $x(i)x(j) = x'(i)x'(j):$ the only exceptions to this are if F' is obtained by merging, and if i or j equals a or b, but then either (say) $i = a$ and $j \notin \{a,b\}$, in which case $x(i)x(j) = x'(c)x'(j)$, or else $i = a$ and $j \in \{a,b\}$, in which case $x(i)x(j) = x'(c)x'(c)$. Thus in every case $x(i)x(j) \geq m(F',f')$, which means $m(F,f) \geq m(F',f')$.

Hence $m(F,f) = m(F',f')$ and x is such that $m(F,f) = \min_{ij\in E(F)} x(i)x(j)$. Since no two vertices have the same weight, F' was not obtained by merging, and so $x' = x$. But we know that $x(i) = \sqrt{m(F,f)f(k-i)/f(i)}$ for all i, and it is impossible to add an edge to F with this x without getting $\min_{ij\in E(F')} x'(i)x'(j) < m(F,f)$. This contradiction means that (F,f) is critical. □

4.4 Tails

As remarked in §2, most graphs H with $t^{1+\tau}$ edges have $\gamma(H) \sim \sqrt{\tau}$. We are now in a position to substantiate this remark. Roughly speaking, in a graph with $\gamma(H) < \sqrt{\tau}$, there must be an independent subset of the vertices whose neighbours lie entirely within a significantly smaller subset. We call such a configuration a *tail*. A more precise description is this.

Definition 4.7 *An* ϵ-tail *in a graph H of order t is a pair (T,S) such that $T,S \subset V(H)$, $T \cap S = \emptyset$, $|T| > |S| + \epsilon t$ and $|E(T, V(H) - S)| \leq t^{-\epsilon}|E(H)|$. (Here, $E(T, V(H) - S)$ includes all edges with both ends in T.)*

Notice that an ϵ-tail is also an η-tail for all $\eta \leq \epsilon$.

Theorem 4.8 *Let $\epsilon > 0$. Let H be a graph of order t with $t^{1+\tau}$ edges. If $\gamma(H) \leq \sqrt{\tau} - 5\sqrt{\epsilon}$ then H has an ϵ-tail.*

Proof. Theorem 4.3 states that H is an ϵ-fit to a shape (F,f) with $|F| \leq 1/\epsilon$ and $\sqrt{\tau}/\sqrt{m(F,f)} - 4\sqrt{\epsilon} \leq \sqrt{\tau} - 5\sqrt{\epsilon}$, so $\sqrt{\tau}/\sqrt{m(F,f)} \leq \sqrt{\tau} - \sqrt{\epsilon} \leq \sqrt{\tau}(1-\sqrt{\epsilon})$ whence $m(F,f) \geq (1-\sqrt{\epsilon})^{-2}$. By Theorem 4.6 we may assume that (F,f) is critical, so F is a half-graph of order $k+1$ for some k, and $m(F,f) = \{\sum_{i=0}^{k} \sqrt{f(i)f(k-i)}\}^{-2}$. Therefore $\sum_{i=0}^{k} \sqrt{f(i)f(k-i)} \leq 1 - \sqrt{\epsilon}$.

By the definition of an ϵ-fit, there is a partition of $V(H)$ into V_a, $a \in F$, with $|V_a| = f(a)t$. Let

$$T = \bigcup_{i=0}^{\lfloor (k-1)/2 \rfloor} V_i \qquad \text{and} \qquad S = \bigcup_{i=\lceil (k+1)/2 \rceil}^{k} V_i = \bigcup_{i=0}^{\lfloor (k-1)/2 \rfloor} V_{k-i}\,.$$

Since the neighbours in F of $\{0,\ldots,\lfloor(k-1)/2\rfloor\}$ lie entirely within $\{\lceil(k+1)/2\rceil,\ldots,k\}$, by the definition of an ϵ-fit we have $|E(T,V(H)-S)|\leq t^{-\epsilon}|E(H)|$.

So the proof will be finished if we show that $|T|\geq|S|+\epsilon t$. Now

$$|T|-|S| \;=\; \sum_{i=0}^{\lfloor(k-1)/2\rfloor}|V_i| - \sum_{i=0}^{\lfloor(k-1)/2\rfloor}|V_{k-i}| \;=\; \sum_{i=0}^{\lfloor(k-1)/2\rfloor}(f(i)-f(k-i))t\,.$$

Since $\sum_{i=0}^{k}f(i)=1$ and $f(i)>f(k-i)$ for $0\leq i<k/2$, we also have (taking $f(k/2)=0$ if k is odd)

$$\begin{aligned}
1-\sqrt{\epsilon} \;&\geq\; \sum_{i=0}^{k}\sqrt{f(i)f(k-i)} \;=\; f(k/2)+2\sum_{i=0}^{\lfloor(k-1)/2\rfloor}\sqrt{f(i)f(k-i)}\\
&=\; 1-\sum_{i=0}^{\lfloor(k-1)/2\rfloor}\left[f(i)-2\sqrt{f(i)f(k-i)}+f(k-i)\right]\\
&=\; 1-\sum_{i=0}^{\lfloor(k-1)/2\rfloor}\left[\sqrt{f(i)}-\sqrt{f(k-i)}\right]^2\\
&\geq\; 1-\sum_{i=0}^{\lfloor(k-1)/2\rfloor}\left[\sqrt{f(i)}-\sqrt{f(k-i)}\right]\left[\sqrt{f(i)}+\sqrt{f(k-i)}\right]\\
&=\; 1-\sum_{i=0}^{\lfloor(k-1)/2\rfloor}f(i)-f(k-i)\,.
\end{aligned}$$

Therefore $|T|-|S|\geq\sqrt{\epsilon}t\geq\epsilon t$ as desired. □

Here is a simple necessary condition, in terms of the minimum and maximum degree, for a graph to have a tail.

Corollary 4.9 *Let $\epsilon>0$ and let H be a graph of order t with $t^{1+\tau}$ edges. If H has an ϵ-tail then $\Delta(H)\geq(1+\epsilon)\delta(H)-2t^{\tau-\epsilon}$.*

Proof. Let (T,S) be an ϵ-tail in H. Each vertex in S has degree at most $\Delta(H)$. Each vertex in T has degree at least $\delta(H)$, and there are at most $t^{1+\tau-\epsilon}$ edges meeting T but not meeting S. Therefore $|S|\Delta(H)\geq|T|\delta(H)-2t^{1+\tau-\epsilon}$. Since $|T|\geq|S|+\epsilon t$ this implies $|S|(\Delta(H)-\delta(H))\geq\epsilon t\delta(H)-2t^{1+\tau-\epsilon}$. Now $|S|\leq t$ so $\Delta(H)-\delta(H)\geq\epsilon\delta(H)-2t^{\tau-\epsilon}$, which is the result claimed. □

4.5 Some common examples

We describe here how $\gamma(H)$ can be evaluated asymtotically for a few common families of graphs. We express the results as limiting values of $\gamma(H)$ as $t\to\infty$.

Almost all graphs of order t have minimum and maximum degree in the range $t/2\pm t^{2/3}$ (see [2]). Likewise almost all graphs with $t^{1+\tau}$ edges have degrees in the range $t^{\tau}/2\pm t^{2\tau/3}$. So the next result follows at once from Theorem 4.8 and Corollary 4.9.

Corollary 4.10 *If H is a regular graph with $t^{1+\tau}$ edges then $\gamma(H) \sim \sqrt{\tau}$.*
For almost all graphs H with $t^{1+\tau}$ edges, $\gamma(H) \sim \sqrt{\tau}$.
For almost all graphs H, $\gamma(H) \sim 1$.

Consider now the complete r-partite graph $K_{\beta_1 t, \beta_2 t, \ldots, \beta_r t}$ of order t, where $\sum_{i=1}^r \beta_i = 1$. Suppose this graph has a tail (T, S). Since T is (apart from $t^{2-\epsilon}$ edges) an independent set, T must be contained almost entirely within one of the r vertex classes. But then S must be almost the whole of $V(H) - T$, else more than $t^{2-\epsilon}$ edges from T would avoid S. Since $|T| \geq |S| + \epsilon/2$ this means that $|T| > t/2$; hence $K_{\beta_1 t, \beta_2 t, \ldots, \beta_r t}$ does not have a hole unless $\beta = \max\{\beta_1, \ldots, \beta_r\} > 1/2$.

Suppose now that $\beta > 1/2$. Then $K_{\beta t, (1-\beta)t} \subset K_{\beta_1 t, \beta_2 t, \ldots, \beta_r t} \subset K_{(1-\beta)t} + \overline{K}_{\beta t}$. We saw in §4.1 that $\gamma(K_{\beta t, (1-\beta)t}) \sim 2\sqrt{\beta(1-\beta)}$. Moreover, if we weight the vertices of $H = K_{(1-\beta)t} + \overline{K}_{\beta t}$ by putting weight $\sqrt{\beta/(1-\beta)}$ on the vertices of $K_{(1-\beta)t}$ and weight $\sqrt{(1-\beta)/\beta}$ on the other vertices, we obtain $\sum_{uv \in E(H)} t^{-w(u)w(v)} \leq \beta(1-\beta)t^2 t^{-1} + \binom{(1-\beta)t}{2} t^{-\beta/(1-\beta)} < t$ if t is large, and so $\gamma(H) \leq 2\sqrt{\beta(1-\beta)}$.

We summarize these observations as follows.

Corollary 4.11 *Let $\beta = \max\{\beta_1, \ldots, \beta_r\}$.*
If $\beta \leq 1/2$ then $\gamma(K_{\beta_1 t, \beta_2 t, \ldots, \beta_r t}) \sim \gamma(K_t) \sim 1$.
If $\beta \geq 1/2$ then

$$\gamma(K_{\beta t, (1-\beta)t}) \sim \gamma(K_{\beta_1 t, \beta_2 t, \ldots, \beta_r t}) \sim \gamma(K_{(1-\beta)t} + \overline{K}_{\beta t}) \sim 2\sqrt{\beta(1-\beta)}.$$

4.6 Dense graphs

Finally in this section we give a very simple argument showing that $\gamma(H)$ is bounded below roughly by the density of H. We make use of this bound in §6.

Theorem 4.12 *Let $\epsilon > 0$ and let H be a graph of order $t \geq (1/\epsilon)^{1/\epsilon}$ and density p. Then $\gamma(H) \geq p - 5\sqrt{\epsilon}$.*

Proof. Let H have $t^{1+\tau} = p\binom{t}{2}$ edges. Let (F, f) be a shape given by Theorem 4.3 to which H is an ϵ-fit, such that $\gamma(H) \geq \sqrt{\tau}/\sqrt{m(F,f)} - 4\sqrt{\epsilon}$. By Theorem 4.6 we may assume that F is a half-graph of order $k+1$. The set E of edges of H lying between classes corresponding to edges of F satisfies (taking $f(k/2) = 0$ if k is odd)

$$(1 - t^{-\epsilon})p\binom{t}{2} \leq |E| \leq \sum_{i \geq k/2} \binom{f(i)t}{2} + \sum_{i<j\,:\,i+j \geq k} f(i)f(j)t^2 + (k+1)t\,,$$

and because $k + 1 \leq 1/\epsilon \leq t^{\epsilon}$ this means

$$\sum_{i+j \geq k} f(i)f(j) \geq p(1 - t^{-\epsilon})(1 - t^{-1}) - t^{-1+\epsilon} \geq p - 3t^{-\epsilon} \geq p - 3\epsilon\,.$$

Since $\sum_{i=0}^{k} f(i) = 1$ we have

$$\begin{aligned}
\sum_{i+j\geq k} f(i)f(j) &= \sum_{i=0}^{k} f(i) \sum_{j=k-i}^{k} f(j) \\
&= \sum_{i\geq k/2} f(i) \sum_{j=k-i}^{k} f(j) + \sum_{i<k/2} f(i) \sum_{j=k-i}^{k} f(j) \\
&= \sum_{i\geq k/2} f(i) \sum_{j=k-i}^{k} f(j) + \sum_{j>k/2} f(j) \sum_{i=k-j}^{k} f(i) \\
&= f(k/2) \sum_{j\geq k/2}^{k} f(j) + 2 \sum_{i>k/2} f(i) \sum_{j=k-i}^{k} f(j) \\
&\leq f(k/2) + 2 \sum_{i>k/2} f(i)\,,
\end{aligned}$$

and so $f(k/2)+2\sum_{i>k/2} f(i) \geq p-3\epsilon$. On the other hand, since $f(i) < f(k-i)$ for $i > k/2$, we have

$$\frac{1}{\sqrt{m(F,f)}} = \sum_{i=0}^{k} \sqrt{f(i)f(k-i)} \geq f(k/2) + 2 \sum_{i>k/2} f(i) \geq p - 3\epsilon\,.$$

Therefore $\gamma(H) \geq \sqrt{\tau}(p-3\epsilon) - 4\sqrt{\epsilon} \geq \sqrt{\tau}p - 3\epsilon - 4\sqrt{\epsilon}$. Now we may assume that $\epsilon \leq 1/25$ and that $p \geq 5\sqrt{\epsilon}$, and so $t^{\tau} = p(t-1)/2 \geq \sqrt{\epsilon}t$, implying $t^{1-\tau} \leq \epsilon^{-1/2} \leq t^{\epsilon/2}$. It follows that $\gamma(H) \geq (1-\epsilon/2)p-3\epsilon-4\sqrt{\epsilon} \geq p-4\epsilon-4\sqrt{\epsilon} \geq p-5\sqrt{\epsilon}$, as claimed. □

5 Extremal graphs

It is asserted in [18] that the extremal graphs for $c(K_t)$ are pseudo-random graphs of density $1-\lambda$ and order $(2\alpha/(1-\lambda))t\sqrt{\log t}$, or essentially disjoint unions of such graphs. (Here the term pseudo-random is to be interpreted in the natural sense of [17] or as quasi-random in Chung, Graham and Wilson [5].) A critical gap in the argument was plugged in [11]. The argument can easily be adapted to show that, for graphs H with $|H|$ large and $\gamma(H)$ bounded away from zero, the extremal graphs for $c(H)$ are also pseudo-random.

A crucial part of the argument is as follows. Suppose we have a graph G of order n and density $p = 1-q$. Let X be a subset of the vertices and let $Y = V(G) - X$. Define the three densities

$$p_X = \frac{e(X)}{\binom{|X|}{2}}\,, \qquad p_{XY} = \frac{e(X,Y)}{|X||Y|}\,, \qquad p_Y = \frac{e(Y)}{\binom{|Y|}{2}}$$

where $e(X)$, $e(Y)$ and $e(X,Y)$ are the numbers of edges of G spanned by X, spanned by Y and joining X to Y. Likewise define $q_X = 1-p_X$, $q_{XY} = 1-p_{XY}$ and $q_Y = 1-p_Y$. It is the principal feature of pseudo-random graphs that G is pseudo-random if and only if $p_{X'}$ differs little from p_X for every X' with $|X'| = |X|$, which of course implies that each of p_X, p_{XY} and p_Y are close to p, the density of G. Note that, whether or not G is quasi-random, the density of G satisfies

$$q \;=\; x^2 q_X + 2x(1-x)q_{XY} + (1-x)^2 q_Y$$

if G is large, where $q = 1-p$ and $x = |X|/|G|$. Now put

$$q^* \;=\; q_X^{x^2}\, q_{XY}^{2x(1-x)}\, q_Y^{(1-x)^2} \,.$$

By taking logarithms and applying Jensen's inequality it can be seen that $q \;\geq\; q^*$ with equality if and only if $q_X = q_{XY} = q_Y = q$.

The following proposition was proved in the case $H = K_t$ by the first author in [11] in response to a question of Sós.

Proposition 5.1 *Given $\epsilon > 0$ there exists $T = T(\epsilon)$ with the following property.*

Let H be a graph with $t > T$ vertices and with $\gamma(H) \geq \epsilon$. Let G be a graph of order n and connectivity $\kappa(G) \geq n(\log\log\log n)/(\log\log n)$, having a vertex partition into X and Y as described above, where $\epsilon < q_X, q_{XY}, q_Y \;\leq\; 1$ and $q^* < 1-\epsilon$. Suppose $n \geq \lfloor (1+\epsilon)\,\gamma(H)\,t\,\sqrt{\log_{1/q^*} t}\rfloor$. Then $G \succ H$.

The proposition shows that, as far as minors are concerned, a graph with a partition as described behaves as well as a graph of density $1 - q^* \geq 1 - q$; in particular (using the arguments of §3.3) the extremal graphs for $c(H)$ must have the property that $q_X = \lambda$ for all X, and so must be pseudo-random.

Proposition 5.1 can be proved by adapting the proof of Theorem 2.5 using the ideas of [11]. We give only a sketch here. The essential difference from the proof of Theorem 2.5 is that the vertices of X and Y are ordered separately, each according to the value of the parameter $q(u,X)^x q(u,Y)^{1-x}$, where $q(u,X)$ is the proportion of the vertices of X not joined to the vertex u. The sets X and Y are then each partitioned into blocks according to their respective orderings. The parts W_u are now once again chosen at random, the condition for goodness being that each of W and $Q(W)$ should be well-distributed both with respect to the ordering of X and with respect to that of Y. The remainder of the proof then follows a similar line to that of Theorem 2.5, with appropriate modifications to the calculations.

6 Sets of forbidden minors and linking

The classical extremal theory takes into account not just a single forbidden graph but classes $\mathcal{H}$ of forbidden graphs. So we might define

$$\mathrm{ex}(\mathcal{H}) \;=\; \lim_{n\to\infty} \inf\{\, c \;:\; |G| \geq n, e(G) \geq c\binom{|G|}{2} \text{ implies } \exists H \in \mathcal{H}, G \succ H \,\}\,.$$

If we put $\chi(\mathcal{H}) = \min\{\chi(H) : H \in \mathcal{H}\}$ then, because $(\chi(\mathcal{H}) - 1)$-partite Turán graphs contain no member of $\mathcal{H}$, it follows at once from the extremal result for a single graph H that $\mathrm{ex}(\mathcal{H}) = \min\{\mathrm{ex}(H) : H \in \mathcal{H}\} = 1 - (\chi(\mathcal{H}) - 1)^{-1}$.

When considering a class $\mathcal{H}$ of forbidden minors, we define

$$c(\mathcal{H}) = \inf\{\, c : e(G) \geq c|G| \text{ implies } G \succ H \text{ for some } H \in \mathcal{H}\,\}.$$

Clearly $c(H) \leq \inf\{c(H) : H \in \mathcal{H}\}$. But the fact that the extremal graphs are pseudo-random means that equality does not always hold. If $|\mathcal{H}|$ is bounded then equality essentially does hold, because if for each $H \in \mathcal{H}$ the random graph $G(n,p)$ almost surely does not contain an H minor, then $G(n,p)$ almost surely contains no H minor for all $H \in \mathcal{H}$. However, this statement can fail if $|\mathcal{H}|$ is unbounded.

Indeed, a vivid instance of this phenomenon occurs if we take

$$\mathcal{H} = \{\, H : |H| \leq t \text{ and } 2\delta(H) \geq |H| + t/3\,\}.$$

Since each graph in H has density at least 2/3, it follows from Theorems 2.2 and 4.12 that (if t is large) $\inf\{c(H) : H \in \mathcal{H}\} \geq (2\alpha/3 + o(1))t\sqrt{\log t}$. But it is proved in [16], based on an idea of Mader [12], that $c(\mathcal{H}) \leq t$. The fact that the $\sqrt{\log t}$ factor is absent here is crucial to the proof in [4] that a graph G with vertex connectivity $\kappa(G) \geq 22k$ is k-linked; the weaker bound $c(\mathcal{H}) \leq (2\alpha/3 + o(1))t\sqrt{\log t}$ would require $\kappa(G) = \Omega(k\sqrt{\log k})$ as in Robertson and Seymour [15].

7 Very sparse minors

As stated, our main theorems above give information only for graphs with $\gamma(H)$ bounded away from zero. It has been observed, after the proof of Theorem 2.5, that the proof given for that theorem will not work for graphs with $\gamma(H) < (\log\log\log t)^{-1/4}$. Thus the methods here cannot tell us anything about very sparse graphs, say with $t(\log t)^{10}$ edges.

Finding very sparse graphs as minors generally involves difficulties that are different to those encountered when looking for somewhat denser graphs. For dense graphs, the problem is to find edges between the subsets W_u that are contracted to form the vertices $u \in H$. This is less of a problem for sparse graphs. Consider for example the graph K_m with $t - m$ further vertices each joined to one vertex of the K_m, where m is small, say $m = t^{2/3}$. (This graph has γ no larger than $t^{-1/3}$, as seen by putting all the weight uniformly on K_m.) The problem with finding this graph as a minor is just to find at least t vertices — there will usually then be enough edges around to find the K_m.

The question is open whether the extremal graphs for $c(H)$ are always random. A non-trivial example of a sparse H for which the extremal function is known exactly is $K_{2,t-2}$ and in this case the extremal graphs are essentially disjoint unions of K_{t-1}'s — see [12]. This means that, in order to get a positive answer to our question, we shall need to admit complete graphs as random graphs (with edge probability one).

As for Theorems 2.3 and 2.5, the question is whether the threshold probability for an H minor is always the same as the threshold density. In one particular instance we can show that this is not the case. A natural candidate of interest that lies in the borderland between graphs with t edges and graphs with $t^{1+\epsilon}$ edges is the k-cube $H = Q_k$, with $t = 2^k$ vertices and $(t/2)\log_2 t$ edges. Riordan [14] has proved that the threshold value of p at which Q_k appears as a *spanning subgraph* of $G(t, p)$ is $p = 1/4$. As he points out in discussion this implies that, for example, $G(2t, p)$ will almost surely have a Q_k minor if $p > 0{\cdot}07$ (since a 1-factor will almost surely appear once $p > 2/\log n$ and then $G(2t, p)$ will "have $G(t, 1/4)$ as a minor"). Consider, though, the graph G of order $t = 2^k$, comprising two copies of $K_{t/2}$ with $t/2 - 1$ independent edges in between, having density $p \sim 1/2$ and $\kappa(G) = t/2 - 1$. The graph G cannot contain Q_k, because the edge-isoperimetric inequality of Harper, Bernstein and Hart [7] says that in any partition of Q_k into two halves there must be at least $t/2$ edges between the halves. Therefore the extremal density for the appearance of Q_k is in graphs of order t not the same as the threshold probability. However, we do not know of similar constructions when, say, $n = 2t$.

References

[1] B. Bollobás, Extremal Graph Theory, Academic Press, London (1978) xx+488pp.

[2] B. Bollobás, Random graphs, *Cambridge Studies in Advanced Mathematics*, **73** (2001). Second edition. Cambridge University Press, Cambridge, xviii+498 pp.

[3] B. Bollobás, P. Catlin and P. Erdős, Hadwiger's conjecture is true for almost every graph, *Europ. J. Combinatorics*, **1** (1980) 195–199.

[4] B. Bollobás and A. Thomason, Highly linked graphs, *Combinatorica* **16** (1996) 313–320.

[5] F.R.K. Chung, R.L. Graham and R.M. Wilson, Quasi-random graphs, Combinatorica **9** (1989), 345–362.

[6] P. Erdős and M. Simonovits, A limit theorem in graph theory, *Studia Sci. Math. Hungar.* **1** (1966), 51–57.

[7] S. Hart, A note on the edges of the n-cube, *Discr. Math.* **14** (1976), 157–163.

[8] A.V. Kostochka, The minimum Hadwiger number for graphs with a given mean degree of vertices (in Russian), *Metody Diskret. Analiz.* **38** (1982) 37–58.

[9] W. Mader, Homomorphieeigenschaften und mittlere Kantendichte von Graphen, *Math. Ann.* **174** (1967) 265–268.

[10] W. Mader, Homomorphiesätze für Graphen, *Math. Ann.* **178** (1968) 154–168.

[11] J.S. Myers, Graphs without large complete minors are quasi-random, *Combinatorics, Probability and Computing* **11** (2002), 571–585.

[12] J.S. Myers, The extremal function for unbalanced bipartite minors, *Discrete Mathematics* **271** (2003) 209–222.

[13] J.S. Myers, Extremal theory of graph minors and directed graphs, Ph.D. dissertation, University of Cambridge (2003).

[14] O. Riordan, Spanning subgraphs of random graphs, *Combin. Probab. Comput.* **9** (2000), 125–148.

[15] Robertson, N. and Seymour, P.D., Graph Minors. XIII. The disjoint paths problem, *J. Combin. Theory Ser. B* **63** (1995) 65–110.

[16] A. Thomason, An extremal function for contractions of graphs, *Math. Proc. Cambridge Phil. Soc.* **95** (1984), 261–265.

[17] A. Thomason, Pseudo-random graphs, in "Random Graphs '85", (M. Karonski and Z. Palka, Eds), Annals of Discrete Math. **33** (1987), 307–331.

[18] A. Thomason, The extremal function for complete minors, *J. Combinatorial Theory Ser. B* **81** (2001) 318–338.