On the purity of minor-closed classes of graphs

McDiarmid, Colin; Przykucki, Michal

DOI:
10.1016/j.jctb.2018.08.010

On the purity of minor-closed classes of graphs

Colin McDiarmid[1] and Michał Przykucki[2]

[1]Department of Statistics, University of Oxford[*]
[2]School of Mathematics, University of Birmingham[†]

September 1, 2018

Abstract

Given a graph H with at least one edge, let $\mathrm{gap}_H(n)$ denote the maximum difference between the numbers of edges in two n-vertex edge-maximal graphs with no minor H. We show that for exactly four connected graphs H (with at least two vertices), the class of graphs with no minor H is pure, that is, $\mathrm{gap}_H(n) = 0$ for all $n \geqslant 1$; and for each connected graph H (with at least two vertices) we have the dichotomy that either $\mathrm{gap}_H(n) = O(1)$ or $\mathrm{gap}_H(n) = \Theta(n)$. Further, if H is 2-connected and does not yield a pure class, then there is a constant $c > 0$ such that $\mathrm{gap}_H(n) \sim cn$. We also give some partial results when H is not connected or when there are two or more excluded minors.

1 Introduction

We say that a graph G contains a graph H as a *minor* if we can obtain a graph isomorphic to H from a subgraph of G by using edge contractions (discarding any loops and multiple edges, we are interested in simple graphs). A class of graphs $\mathcal{A}$ is *minor-closed* if for each $G \in \mathcal{A}$, each minor G' of G is also in $\mathcal{A}$. We say that H is an *excluded minor* for $\mathcal{A}$ if H is not in $\mathcal{A}$ but each minor of H (other than H itself) is in $\mathcal{A}$. Robertson and Seymour [17] showed that, for each minor-closed class $\mathcal{A}$ of graphs, the set $\mathcal{H}$ of excluded minors is finite. We say that G is *H-free* if it has no minor H; and given a set $\mathcal{H}$ of graphs, G is *$\mathcal{H}$-free* if it is H-free for all $H \in \mathcal{H}$. We denote the class of all $\mathcal{H}$-free graphs by $\mathrm{Ex}(\mathcal{H})$, and write $\mathrm{Ex}(H)$ when $\mathcal{H}$ consists of just the graph H.

Observe that $\mathrm{Ex}(\mathcal{H})$ contains n-vertex graphs for each n as long as each graph in $\mathcal{H}$ has at least one edge. Let us call such a set $\mathcal{H}$ of graphs *suitable* as long as it is non-empty. We shall restrict our attention to suitable classes $\mathcal{H}$, and be interested in the number of edges in edge-maximal $\mathcal{H}$-free graphs. Given a graph G, let $v(G)$ denote the number of vertices and $e(G)$ the number of edges. For all $n \geqslant 1$, let

$$E_{\mathcal{H}}(n) = \{e(G) : v(G) = n \text{ and } G \text{ is an edge-maximal } \mathcal{H}\text{-free graph}\}.$$

[*]Email: cmcd@stats.ox.ac.uk

[†]Email: m.j.przykucki@bham.ac.uk. During a large part of this project, the second author was affiliated with St Anne's College and the Mathematical Institute of the University of Oxford.

Also, let $M^+_{\mathcal{H}}(n) = \max E_{\mathcal{H}}(n)$, and $M^-_{\mathcal{H}}(n) = \min E_{\mathcal{H}}(n)$. Finally, let us define

$$\mathrm{gap}_{\mathcal{H}}(n) = M^+_{\mathcal{H}}(n) - M^-_{\mathcal{H}}(n).$$

As for $\mathrm{Ex}(H)$, we write $\mathrm{gap}_H(n)$ to denote $\mathrm{gap}_{\mathcal{H}}(n)$ when $\mathcal{H}$ consists of just the graph H. This is the case on which we focus.

The function $M^+_{\mathcal{H}}(n)$ (sometimes in the form of $2M^+_{\mathcal{H}}(n)/n$ to analyse the maximum average degree of graphs in $\mathcal{A} = \mathrm{Ex}(\mathcal{H})$) has been studied extensively for various suitable sets $\mathcal{H}$. Mader [9] showed that, given a graph H, there is a constant $c = c(H)$ such that $e(G) \leqslant c\, v(G)$ for each graph $G \in \mathrm{Ex}(H)$. Let us define $\beta_{\mathcal{H}}$ by setting

$$\beta_{\mathcal{H}} := \sup_{G\in\mathcal{A}} \frac{e(G)}{v(G)} = \sup_{n\geqslant 1} \frac{M^+_{\mathcal{H}}(n)}{n}, \tag{1}$$

noting that $\beta_{\mathcal{H}}$ is finite. Write β_H when $\mathcal{H}$ consists just of the graph H. Building on work of Mader [10], Kostochka [8] and Fernandez de la Vega [4], Thomason [18] showed that, for each positive integer r, we have $M^+_{K_r}(n) \sim \beta_{K_r} n$ as $n \to \infty$; and $\beta_{K_r} \sim \alpha\, r\sqrt{\log r}$ as $r \to \infty$, where $\alpha \approx 0.319$. The value of β_H for dense graphs H was studied by Myers and Thomason [14]. Reed and Wood [16] analysed this parameter for sparse forbidden minors H. Csóka, Lo, Norin, Wu and Yepremyan [2] focused on H being a union of disjoint cycles and, more generally, of disjoint 2-connected graphs.

Much less is known about the function $M^-_{\mathcal{H}}(n)$ and, consequently, about $\mathrm{gap}_{\mathcal{H}}(n)$, for a suitable set $\mathcal{H}$. From Mader's result it follows that we always have $\mathrm{gap}_{\mathcal{H}}(n) = O(n)$. We say that the class $\mathcal{A} = \mathrm{Ex}(\mathcal{H})$ is *pure* if we have $\mathrm{gap}_{\mathcal{H}}(n) = 0$ for each positive integer n. For example, the class $\mathrm{Ex}(K_3)$ of forests is pure, since the n-vertex edge-maximal forests are the trees, each with $n-1$ edges. Our first main theorem is:

Theorem 1. *The connected graphs H on at least two vertices such that the class $\mathrm{Ex}(H)$ is pure are precisely the complete graphs K_2, K_3, K_4 and the 3-vertex path P_3.*

We say that $\mathrm{Ex}(\mathcal{H})$ is *near-pure* if it is not pure, but we still have $\mathrm{gap}_{\mathcal{H}}(n) = O(1)$. Also, we define the 'linear impurity parameter'

$$\mathrm{limp}(\mathcal{H}) = \liminf_{n\to\infty} \frac{\mathrm{gap}_{\mathcal{H}}(n)}{n};$$

and we say that $\mathrm{Ex}(\mathcal{H})$ is *linearly impure* if $\mathrm{limp}(\mathcal{H}) > 0$. Our second main result shows that all connected graphs H fall into one of only three categories according to the purity of the class $\mathrm{Ex}(H)$.

Theorem 2. *For each connected graph H on at least two vertices, the class of H-free graphs is either pure, near-pure or linearly impure.*

In other words, Theorem 2 says that it is not possible for the impurity of a class of H-free graphs to be unbounded but not grow linearly fast in n. We have seen in Theorem 1 that if H is K_3 or K_4 then $\mathrm{gap}_H(n) = 0$ for each n, and $\mathrm{limp}(H) = 0$. More generally, whenever H is 2-connected, $\mathrm{gap}_H(n)/n$ tends to limit, so the 'liminf' in the definition of limp could be replaced by the more satisfactory 'lim' (see also Theorem 4).

Theorem 3. *Let H be a 2-connected graph other than K_3 or K_4. Then, as $n \to \infty$,*

$$\frac{\mathrm{gap}_H(n)}{n} \;\to\; \mathrm{limp}(H)\; > 0. \tag{2}$$

An important example of a pure minor-closed class is the class of planar graphs. Indeed, for each $n \geqslant 3$, all n-vertex edge-maximal graphs G embeddable in the plane are triangulations, satisfying $e(G) = 3n - 6$. However, somewhat surprisingly, it is not the case that a similar statement holds for graphs embeddable in the torus: it was shown in [6] that a complete graph on 8 vertices with the edges of a 5-cycle C_5 removed (thus containing 23 edges) is an edge-maximal graph embeddable in the torus, while each 8-vertex triangulation of the torus, by Euler's formula, contains 24 edges. However, for every surface S, the (minor-closed) class of graphs embeddable in S is pure or near-pure, as shown by McDiarmid and Wood [13].

At this point, let us check that the four connected graphs listed in Theorem 1, namely K_2, K_3, K_4 and P_3, give rise to pure H-free classes of graphs. The case of K_2 is trivial, as $\mathrm{Ex}(K_2)$ consists of the graphs without edges. We already noted that the class $\mathrm{Ex}(K_3)$ of forests is pure. If $H = P_3$, the path on 3 vertices, then the n-vertex edge-maximal H-free graphs are the maximal matchings, each with $\lfloor n/2 \rfloor$ edges. Finally, the class $\mathrm{Ex}(K_4)$ is the class of series-parallel graphs, which is also the class of graphs of treewidth at most 2. For each $n \geqslant 2$ each n-vertex edge-maximal such graph has exactly $2n - 3$ edges. In fact, for each fixed $k \geqslant 1$ the edge-maximal graphs of treewidth at most k are the k-trees, and each n-vertex k-tree has $kn - \binom{k+1}{2}$ edges for $n \geqslant k$ (and $\binom{n}{2}$ for $n < k$). Thus for each $k \geqslant 1$ the class of graphs of treewidth at most k is pure. We will have to work much harder to prove that the four graphs listed are the only connected graphs H for which $\mathrm{Ex}(H)$ is pure!

The rest of the paper is organised as follows. In the next section we introduce addable graph classes, and prove a general limiting result, Theorem 4, which yields the 'limit part' of Theorem 3. We also sketch a useful consequence of purity or near-purity for such a class of graphs. In Section 3 we show that for each connected graph H with no leaf (that is, with minimum degree $\delta(H) \geqslant 2$), if H is not K_3 or K_4 then $\mathrm{Ex}(H)$ is linearly impure. This is a step towards proving both Theorems 1 and 2, and together with Theorem 4 proves Theorem 3, concerning a 2-connected graph H. In Section 4 we complete the proof of Theorem 2, showing that for a connected excluded minor there are only the three possibilities of purity, near-purity or linear impurity. In Section 5 we complete the proof of Theorem 1, showing that only four connected graphs H give rise to pure H-free classes. In Section 6 we give some extensions of our results to suitable sets $\mathcal{H}$ of two or more excluded graphs, and to forbidding disconnected graphs; and finally we propose some natural open problems.

2 Addable graph classes

In this section we introduce addable graph classes. We show that, for an addable minor-closed class $\mathcal{A}$ of graphs with suitable set $\mathcal{H}$ of excluded minors, $\mathrm{gap}_{\mathcal{H}}(n)/n$ tends to a limit, and we identify that limit as a difference of two terms (see (4)). Finally we describe a consequence of purity or near-purity for growth constants when we have a given average degree.

We say that a graph class $\mathcal{A}$ is *addable* when

1. $G \in \mathcal{A}$ if and only if every component of G is in $\mathcal{A}$ (following Kolchin [7], if $\mathcal{A}$ satisfies this property we call it *decomposable*), and
2. whenever $G \in \mathcal{A}$ and u, v belong to different components of G then the graph obtained from G by adding the edge $\{u, v\}$ is also in $\mathcal{A}$ (following [11], such a class $\mathcal{A}$ is called *bridge-addable*).

A minor-closed class is decomposable if and only if each excluded minor is connected, and it is addable if and only if each excluded minor is 2-connected. For example, the classes of forests $(\mathrm{Ex}(K_3))$, series-parallel graphs $(\mathrm{Ex}(K_4))$, and planar graphs $(\mathrm{Ex}(\{K_5, K_{3,3}\}))$ are each addable.

The following general limiting result shows that in the addable case, the 'liminf' in the definition of limp can be replaced by 'lim'.

Theorem 4. *Let $\mathcal{A}$ be an addable minor-closed class of graphs, with suitable set $\mathcal{H}$ of excluded minors. Then, as $n \to \infty$,*

$$\frac{\mathrm{gap}_{\mathcal{H}}(n)}{n} \to \mathrm{limp}(\mathcal{H}). \tag{3}$$

To prove this result, we use two lemmas, treating $M^+_{\mathcal{H}}(n)$ and $M^-_{\mathcal{H}}(n)$ separately. Recall that $\beta_{\mathcal{H}}$ was defined in (1). In the following lemma, it is easy to see that $\beta_{\mathcal{H}} \geqslant 1$, since $\mathcal{A}$ contains all the forests.

Lemma 5. *Let $\mathcal{A}$ be a decomposable minor-closed class of graphs, with suitable set $\mathcal{H}$ of excluded minors. Then*

$$\frac{1}{n} M^+_{\mathcal{H}}(n) \to \beta_{\mathcal{H}} \quad \text{as } n \to \infty.$$

Proof. Denote $M^+_{\mathcal{H}}(n)$ by $f(n)$. For $i = 1, 2$ let n_i be a positive integer and let $G_i \in \mathcal{A}_{n_i}$ satisfy $e(G_i) = f(n_i)$. Since the disjoint union $G_1 \cup G_2$ is in $\mathcal{A}_{n_1+n_2}$ we have

$$f(n_1 + n_2) \geqslant f(n_1) + f(n_2);$$

that is, f is superadditive. Hence by Fekete's Lemma (see for example van Lint and Wilson [19])

$$\frac{f(n)}{n} \to \sup_k \frac{f(k)}{k} = \beta_{\mathcal{H}} \quad \text{as } n \to \infty.$$

□

Lemma 6. *Let $\mathcal{A}$ be an addable minor-closed class of graphs, with suitable set $\mathcal{H}$ of excluded minors. Then there is a constant $\beta^-_{\mathcal{H}} \geqslant 1$ such that*

$$\frac{1}{n} M^-_{\mathcal{H}}(n) \to \beta^-_{\mathcal{H}} \quad \text{as } n \to \infty.$$

Proof. Let $h = \min\{v(H) : H \in \mathcal{H}\}$, and note that $h \geqslant 3$. Consider the function $f(n) = M^-_{\mathcal{H}}(n) + (h-2)^2$. Note that each edge-maximal graph in $\mathcal{A}$ is connected, so $f(n) \geqslant n$ for each n. Let $\beta^-_{\mathcal{H}} = \inf_k f(k)/k \geqslant 1$. We claim that $f(n)$ is subadditive (that is $f(a+b) \leqslant f(a) + f(b)$), so by Fekete's Lemma, as $n \to \infty$ we have $f(n)/n \to \beta^-_{\mathcal{H}}$ and thus also $M^-_{\mathcal{H}}(n)/n \to \beta^-_{\mathcal{H}}$.

It remains to establish the claim that f is subadditive. Let $n_1, n_2 \geqslant 1$ and let G_1, G_2 be edge-maximal $\mathcal{H}$-free graphs with $v(G_1) = n_1$, $v(G_2) = n_2$, and such that $e(G_1) = M^-_{\mathcal{H}}(n_1), e(G_2) = M^-_{\mathcal{H}}(n_2)$. Note that G_1 and G_2 are connected.

As in the proof of the last lemma, the disjoint union $G = G_1 \cup G_2$ is $\mathcal{H}$-free. It will be enough to show that we cannot add more than $(h-2)^2$ edges to G without creating an H-minor for some $H \in \mathcal{H}$. Indeed, let $u_1 \neq v_1$ be in $V(G_1)$ and let u_2, v_2 be in $V(G_2)$, and assume that we can (simultaneously) add the edges $\{u_1, u_2\}$ and $\{v_1, v_2\}$ to G without creating any H-minor. Then the edge $\{u_1, v_1\}$ must be present in G_1 since otherwise, after adding $\{u_1, u_2\}$ and $\{v_1, v_2\}$ to G, by the connectedness of G_2 there is a path between u_1 and v_1 that uses only vertices in G_2, and we may contract this path to an edge between u_1 and v_1: this would necessarily create an H-minor for some $H \in \mathcal{H}$ by the edge-maximality of G_1.

Hence if we add edges to G without creating any H-minor then the vertices in G_1 incident to the edges that we add must induce a clique in G_1, with an analogous statement holding for G_2. By the definition of h, these cliques can have size at most $h-2$ (if there were an $(h-1)$-clique in G_1 say, and we contracted G_2 to a single vertex, we would obtain an h-clique), hence we can add at most $(h-2)^2$ edges. Consequently,

$$\begin{aligned} f(n_1+n_2) &= M^-_{\mathcal{H}}(n_1+n_2) + (h-2)^2 \\ &\leqslant \left(M^-_{\mathcal{H}}(n_1) + M^-_{\mathcal{H}}(n_2) + (h-2)^2\right) + (h-2)^2 \\ &= f(n_1) + f(n_2). \end{aligned}$$

Thus $f(n)$ is subadditive, and the proof is complete. □

The last two lemmas show that, if $\mathcal{A}$ is an addable minor-closed class of graphs with suitable set $\mathcal{H}$ of excluded minors, then

$$\frac{\mathrm{gap}_{\mathcal{H}}(n)}{n} \to \beta_{\mathcal{H}} - \beta^-_{\mathcal{H}} \quad \text{as } n \to \infty. \tag{4}$$

Thus $\mathrm{limp}(\mathcal{H}) = \beta_{\mathcal{H}} - \beta^-_{\mathcal{H}}$, and $\frac{\mathrm{gap}_{\mathcal{H}}(n)}{n} \to \mathrm{limp}(\mathcal{H})$ as $n \to \infty$, which completes the proof of Theorem 4.

We close this section by sketching a useful consequence of purity or near-purity. Let $\mathcal{A}$ be a minor-closed class of graphs, with non-empty set $\mathcal{H}$ of excluded minors. Let $\mathcal{A}_n$ denote the set of graphs in $\mathcal{A}$ on vertex set $[n] = \{1, 2, \ldots, n\}$, let $a_n = |\mathcal{A}_n|$, and let

$$\gamma(\mathcal{A}) = \limsup_{n\to\infty} \left(\frac{a_n}{n!}\right)^{1/n}.$$

Norine, Seymour, Thomas and Wollan [15] (see also Dvořák and Norine [3]) showed that $\gamma(\mathcal{A}) < \infty$. Now suppose that $\mathcal{A}$ is addable, that is, the excluded minors are 2-connected. Then (see, for example [11]), $(a_n/n!)^{1/n}$ converges to $\gamma(\mathcal{A})$ and we say that $\mathcal{A}$ has *growth constant* $\gamma(\mathcal{A})$. Defining $a_{n,q} = |\mathcal{A}_{n,q}|$ to be the number of graphs in $\mathcal{A}_n$ with $\lfloor qn \rfloor$ edges, following the methods in Gerke, McDiarmid, Steger and Weißl [5] it can be shown that $(a_{n,q}/n!)^{1/n}$ tends to a limit $\gamma(\mathcal{A}, q)$. If $\mathcal{A}$ is pure or near-pure then, again following the analysis in [5], we may see that $\gamma(\mathcal{A}, q)$ as a function of q is log-concave, and hence continuous, for $q \in (1, \beta_{\mathcal{H}})$.

3 Purity and linear impurity: excluding a leafless graph

In this section we prove the following lemma, which shows linear impurity for some excluded minors H. It is a step towards proving both Theorems 1 and 2, and together with Theorem 4 immediately yields Theorem 3.

Lemma 7. *Let H be a connected graph with $\delta(H) \geqslant 2$, other than K_3 and K_4. Then H is linearly impure.*

We shall often use the following fact proved by Sylvester in 1884.

Fact 8. *Let a_1, a_2 be a pair of positive coprime integers. Then for every integer $N > a_1a_2 - a_1 - a_2$ there are some non-negative integers b_1, b_2 such that*

$$N = a_1b_1 + a_2b_2.$$

Let us call a vertex v in a connected h-vertex graph H a *strong separating vertex* if each component of $H - v$ has at most $h - 3$ vertices (so v is a separating vertex which does not just cut off a single leaf). In order to prove Lemma 7 we first consider complete graphs, and then non-complete graphs with no leaves. In the next lemma we deal with complete graphs.

Lemma 9. *For each $r \geqslant 5$ the class of K_r-free graphs satisfies* $\operatorname{limp}(K_r) \geqslant \frac{7}{6}$.

Proof. We prove the lemma by induction on r. First, let $r = 5$. Wagner [20] showed that any edge-maximal K_5-free graph on at least 4 vertices can be constructed recursively, by identifying edges or triangles, from edge-maximal planar graphs (i.e., triangulations) and copies of the Wagner graph (recall that the Wagner graph is formed from the cycle C_8 by joining the four opposite pairs of vertices, hence it has 8 vertices and 12 edges). If $n = 6k + 2$, we can take G_1 to be an arbitrary plane triangulation on n vertices with $e(G_1) = 3n - 6 = 18k$. We then take G_2 to be a clique-sum of k copies of the Wagner graph W_8 that all overlap in one common edge. Then $e(G_2) = 11k + 1$ and

$$\frac{e(G_1) - e(G_2)}{n} = \frac{7k - 1}{6k + 2} \to 7/6$$

as $k \to \infty$. For general n we can modify the construction of G_2 by taking a clique-sum of k copies of W_8 and a triangulation on $3 \leqslant m \leqslant 7$ vertices (in fact, by the above characterisation of the edge-maximal K_5-free graphs, it is easy to check that $\operatorname{limp}(K_5) = \frac{7}{6}$). Therefore the lemma holds for $r = 5$.

The statement for $r + 1$ follows from the statement for r by observing that if we take any edge-maximal K_r-free graph G, add to it one vertex and connect it to all vertices of G, then the resulting graph is edge-maximal K_{r+1}-free. $\square$

Remark 10. Recall from [18] that $M^+_{K_r}(n) \sim \alpha\, r\sqrt{\log r}\, n$ for $\alpha \approx 0.319$, while the constructions in Lemma 9 have both $e(G_1)$ and $e(G_2)$ that grow linearly with r. Thus we see that $\operatorname{limp}(K_r) \sim \alpha\, r\sqrt{\log r}$.

We next consider connected graphs that are not complete but do not have any leaves. We say that G has *connectivity* k if k is the minimum size of a vertex cut of G (except that, for $n \geqslant 2$, K_n has connectivity $n - 1$). Also, we say that G is j-connected if G has connectivity at least j. Recall that $\delta(G)$ denotes the minimum degree and, for $u \in V(G)$, let

$$N(u) = \{v \in V(G) : \{u, v\} \in E(G)\}$$

denote the neighbourhood of u in G. The following simple fact will be very useful to us.

Fact 11. *Let G be a non-complete graph on n vertices with $\delta(G) = \delta$. Then G has connectivity at least $2\delta - n + 2$.*

Proof. Let u and v be a non-adjacent pair of vertices. Then

$$2\delta \leqslant \deg(u) + \deg(v) = |N(u) \cup N(v)| + |N(u) \cap N(v)| \leqslant n - 2 + |N(u) \cap N(v)|,$$

so u and v have at least $2\delta - n + 2$ common neighbours, and any vertex cut separating u and v must contain all of these vertices. □

Lemma 12. *Let H be a connected non-complete graph on $h \geqslant 4$ vertices with $\delta := \delta(H) \geqslant 2$. Then the class of H-free graphs satisfies* $\operatorname{limp}(H) \geqslant \frac{1}{2h}$.

Proof. Since H is connected, H has connectivity k for some $k \geqslant \max\{2\delta - h + 2, 1\}$. We first show that for all $m \geqslant 1$ there exist two graphs G_1, G_2, both on

$$n = (h-k)(h-k+1)m + k - 1$$

vertices, that are edge-maximal H-free and such that

$$e(G_1) - e(G_2) \geqslant \frac{(h-k)m}{2} = (1 + o(1))\frac{n}{2(h-k+1)}.$$

We construct the "dense" graph G_1 as follows. We take $(h-k+1)m$ copies of K_{h-1} that all overlap in a fixed set of $k-1$ vertices. Clearly G_1 is H-free since H has connectivity k and trying to fit an H-minor in G_1 we would need to find it across more than one of the copies of K_{h-1}. Also, G_1 has $(h-k)(h-k+1)m + k - 1$ vertices and

$$\begin{aligned} e(G_1) &= (h-k+1)m\left(\binom{h-k}{2} + (h-k)(k-1)\right) + \binom{k-1}{2} \\ &= (h-k)m(h-k+1)\frac{h+k-3}{2} + \binom{k-1}{2}. \end{aligned}$$

We construct the "sparse" graph G_2 similarly. We start by taking $(h-k)m$ copies of K_{h-1} that all overlap in a fixed set I of $k-1$ vertices. The resulting graph G_2' has $(h-k)^2m + k - 1$ vertices, i.e., $(h-k)m$ fewer that G_1. We complete the construction of G_2 by adding these $(h-k)m$ missing vertices and joining each of them to $\delta - 1$ vertices in a distinct copy of K_{h-1} in such a way that the neighbourhood of each new vertex does not contain the whole of I (see Figure 1). Note that G_2 is H-free: for if G_2 had a minor H then so would G_2' (since vertices v of degree $< \delta(H)$ with $N(v)$ complete are redundant), and we may see as for G_1 that G_2' has no minor H. We have

$$\begin{aligned} e(G_2) &= (h-k)m\left(\binom{h-k}{2} + (h-k)(k-1)\right) + \binom{k-1}{2} + (h-k)m(\delta-1) \\ &= (h-k)m\left((h-k)\frac{h+k-3}{2} + \delta - 1\right) + \binom{k-1}{2}. \end{aligned}$$

Consequently,

$$e(G_1) - e(G_2) = (h-k)m\left(\frac{h+k-3}{2} - \delta + 1\right).$$

By Fact 11 we have $h + k - 3 \geqslant 2\delta - 1$ hence

$$e(G_1) - e(G_2) \geqslant \frac{(h-k)m}{2} = \frac{n-k+1}{2(h-k+1)} \sim \frac{n}{2(h-k+1)}.$$

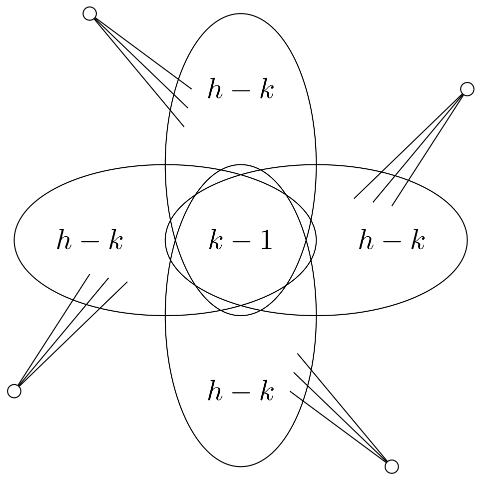

Figure 1: Graph G_2 as defined in Lemma 12.

To show that G_2 is edge-maximal H-free, assume that we add an edge e to G_2. If e connects vertices not in I in two distinct copies of K_{h-1}, then by contracting it we obtain two copies of K_{h-1} that overlap in k vertices and the resulting graph contains H as a subgraph because H has connectivity k. If e connects a vertex v of degree $\delta - 1$ to a vertex in the copy of K_{h-1} that contains the whole of $N(v)$ then this graph contains H as a subgraph because now $\deg(v) = \delta = \delta(H)$. If finally e connects a vertex v of degree $\delta - 1$ to another vertex u that either has degree $\delta - 1$ or is located in some other copy of K_{h-1} then we can contract the path between u and a vertex in $I \setminus N(v)$. The resulting graph again contains H as a subgraph, because now $\deg(v) = \delta = \delta(H)$.

To complete the proof of the lemma we observe that $h - k$ and $h - k + 1$ are coprime. Thus by Fact 8 for all n large enough we can build approximations G_1', G_2' of the above graphs G_1, G_2 using the building blocks described above (K_{h-1}, and K_{h-1} plus a vertex of degree $\delta - 1$), with $\frac{e(G_1')-e(G_2')}{n} \to \frac{1}{2(h-k+1)} \geqslant \frac{1}{2h}$. □

At this stage, we have seen by Lemmas 9 and 12 that, if the connected graph H has $\delta(H) \geqslant 2$ and H is not K_3 or K_4, then $\mathrm{limp}(H) > 0$; that is, we have proved Lemma 7. Now Theorem 3 follows from Theorem 4.

4 Purity, near-purity and linear impurity: excluding a graph with a leaf

In this section we complete the proof of Theorem 2, which says that for a connected excluded minor H there are only the three possibilities of purity, near-purity or linear impurity for $\mathrm{Ex}(H)$. We first deal quickly with graphs H which have a strong separating vertex, treating the claw graph $K_{1,3}$ separately in Observation 14; and then we consider graphs H with at least one leaf and no strong separating vertex.

Lemma 13. *Let H be a connected graph on $h \geqslant 5$ vertices which contains a strong separating vertex v. Then the class of H-free graphs satisfies* $\mathrm{limp}(H) \geqslant \frac{1}{2}$.

Proof. The construction here is very simple. For $m \geqslant 1$, let G_1 consist of $(h-2)m$ disjoint copies of K_{h-1} and let G_2 consist of $(h-1)m$ disjoint copies of K_{h-2}. Both graphs contain $n = (h-1)(h-2)m$ vertices and are trivially H-free. They are edge-maximal H-free because whenever we add an edge e to either G_1 or G_2, we can then contract it and identify the resulting common vertex of two cliques of size either $h-1$ or $h-2$ with v. The resulting graph contains H as a subgraph because $h \geqslant 5$ and consequently $h-2+h-3 \geqslant h$.

We clearly have $e(G_1) = (h-1)(h-2)^2m/2$ and $e(G_2) = (h-1)(h-2)(h-3)m/2$. Hence

$$e(G_1) - e(G_2) = \frac{(h-1)(h-2)m}{2} = \frac{n}{2}.$$

The construction for general n follows easily from Fact 8 since $h-1$ and $h-2$ are coprime. $\square$

Observation 14. *The only connected graph on $h = 4$ vertices with a strong separating vertex is the claw $K_{1,3}$. The class of $K_{1,3}$-free graphs is not pure, since for all $n \geqslant 4$ the cycle C_n and the union of a cycle C_{n-1} and an isolated vertex are edge-maximal $K_{1,3}$-free with n and $n-1$ edges respectively.*

However, this class is near-pure with $\mathrm{gap}_{K_{1,3}}(n) = 1$ for all $n \geqslant 4$. Indeed, note that any connected component of an edge-maximal $K_{1,3}$-free graph G on n vertices is either a cycle, an edge or an isolated vertex. Moreover, G can have at most one component of size less than 3 to preserve edge-maximality. Hence G must have either n or $n-1$ edges.

For the rest of this section we consider the case when the connected graph H on h vertices has at least one leaf and has no strong separating vertex. We say that a connected graph G is *leaf-and-edge-maximal H-free* if G is edge-maximal H-free and attaching a new leaf to an arbitrary vertex of G creates an H-minor.

Lemma 15. *Suppose that the connected graph H has a leaf, and the class of H-free graphs is not linearly impure. Then each leaf-and-edge-maximal H-free graph G satisfies $e(G)/v(G) = (h-2)/2$; and each H-free graph G satisfies $e(G)/v(G) \leqslant (h-2)/2$.*

Proof. Indeed, if there existed two leaf-and-edge-maximal H-free graphs G_1, G_2 with $\frac{e(G_1)}{v(G_1)} > \frac{e(G_2)}{v(G_2)}$ then we could trivially construct two arbitrarily large edge-maximal H-free graphs with the same number of vertices: G' consisting of disjoint copies of G_1, and G'' consisting of disjoint copies of G_2, such that

$$e(G') - e(G'') = \left(\frac{e(G_1)}{v(G_1)} - \frac{e(G_2)}{v(G_2)}\right) v(G').$$

Further, to handle general n, to both G' and G'' we could add a union of at most $\frac{\max\{v(G'),v(G'')\}}{h-1}$ disjoint copies of K_{h-1} and a K_i for some $1 \leqslant i \leqslant h-2$, keeping the graph edge-maximal H-free.

The claim now follows from the observation that, since H has a leaf, K_{h-1} is always a leaf-and-edge-maximal H-free graph. The second statement in the lemma follows similarly, by taking G_1 as G and G_2 as K_{h-1}. $\square$

Observation 16. *Suppose that H has no strong separating vertex. Then any edge-maximal H-free graph contains at most one component that is not leaf-and-edge-maximal H-free. Otherwise we could connect two such components by a suitably attached edge, and the resulting graph would still be H-free because H has no strong separating vertex and the components we started with were not leaf-and-edge-maximal H-free.*

The next lemma is the final step towards proving Theorem 2.

Lemma 17. *Let H be a graph on h vertices that is connected, has at least one leaf and has no strong separating vertex. If there exists $n > 0$ and two edge-maximal H-free graphs on n vertices G_1, G_2 such that*

$$e(G_1) - e(G_2) \geqslant M = \frac{h-2}{2} + 2\beta_H^2 + 1 \tag{5}$$

then the class of H-free graphs is linearly impure.

Proof. Assume for a contradiction that the class of H-free graphs is not linearly impure. Let G_1 and G_2 be edge-maximal H-free graphs on the same vertex set such that $e(G_1) - e(G_2) \geqslant M$, for M as in (5) (we observe that $\beta_H \leqslant \beta_{K_h}$). If all components of G_2 were leaf-and-edge-maximal H-free then by Lemma 15 we would have some component C of G_1 that was H-free and satisfied $e(C)/v(C) > (h-2)/2$, which contradicts Lemma 15.

Hence, by Observation 16, G_2 has exactly one component C, with $|C| = c$, that is edge-maximal H-free, but not leaf-and-edge-maximal H-free. By Lemma 15 we have

$$e(C) \leqslant \frac{h-2}{2}c - M.$$

Let A be the set of all vertices v in C such that attaching a leaf to v does not create an H minor, and let $a = |A| \geqslant 1$. Clearly the graph induced by A must be H-free so the set A induces at most $\beta_H a$ edges. Let v be a vertex in A with the minimum number of neighbours in A: clearly, $\deg_A(v) \leqslant 2\beta_H$.

Let $n = (c-1)m + 1 + t(h-1) + s$, where $t \geqslant 0$, $0 \leqslant s \leqslant h-2$, and $t(h-1) + s < c-1$. We take m isomorphic copies of C and turn them into one connected graph on $n' = (c-1)m + 1$ vertices and $me(C)$ edges by identifying the vertices v in all these copies into one vertex (still called v). Next, we add a copy of K_s and join all (if $s \leqslant h-3$) or one (if $s = h-2$) of its vertices to v by an edge. Finally, to this graph we add t disjoint copies of K_{h-1}. The resulting graph G on n vertices is H-free by the definition of A and by the fact that H has no strong separating vertex (this latter property is the reason why we can join all the vertices of K_s with v if $s \leqslant h-3$).

We do not know if this graph is edge-maximal H-free. However, observe that we can only add edges to G between distinct copies of the set A, or between one of the copies of A and the clique K_s, or between v and the clique K_s (if $s = h-2$). Moreover, we are not allowed to add edges incident to vertices in A that are not adjacent to v. Indeed, assume that we add an edge $\{u, w\}$ such that $u \notin N(v)$. Then by contracting a path from w to v (recall that C is connected) we "add" the edge $\{u, v\}$ to a copy of C which creates an H minor by the edge-maximality of C (see Figure 2). Hence there are at most $2\beta_H m + s$ vertices other than v between which we can add edges and keep the graph H-free.

Again, to avoid creating an H-minor we can add at most $2\beta_H^2 m + \beta_H s$ edges between these vertices. Since we can add at most $h-3$ edges incident to v, there exists an edge-maximal H-free graph G' on $n = (c-1)m + 1 + t(h-1) + s$ vertices with

$$e(G') \leqslant e(C)m + 2\beta_H^2 m + \beta_H s + h - 3 < m\left(\frac{h-2}{2}c - M + 2\beta_H^2\right) + h(\beta_H + 1).$$

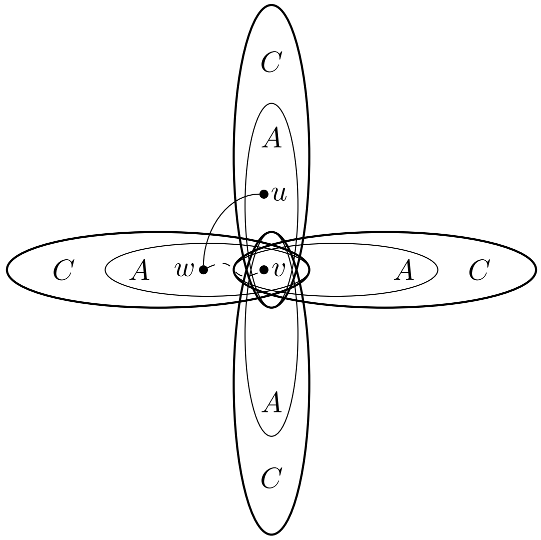

Figure 2: An H-free graph G with $t = s = 0$ as defined in Lemma 17.

We take G'' to be an edge-maximal H-free graph on n vertices consisting of $\lfloor n/(h-1) \rfloor$ disjoint copies of K_{h-1} and one copy of K_i for some $0 \leqslant i \leqslant h-2$. Hence we have

$$e(G'') > \frac{n(h-2)}{2} - \frac{(h-2)^2}{2}.$$

Therefore

$$\begin{aligned} e(G'') - e(G') &\geqslant ((c-1)m+1)\frac{h-2}{2} - \frac{(h-2)^2}{2} - m\left(\frac{h-2}{2}c - M + 2\beta_H^2\right) - h(\beta_H+1) \\ &\geqslant m\left(M - \frac{h-2}{2} - 2\beta_H^2\right) - \frac{(h-3)(h-2)}{2} - h(\beta_H+1) \\ &\geqslant m - \frac{(h-3)(h-2)}{2} - h(\beta_H+1), \end{aligned}$$

where the last inequality follows from (5). Consequently,

$$\frac{e(G'') - e(G')}{n} \geqslant \frac{m - \frac{(h-3)(h-2)}{2} - h(\beta_H+1)}{(c-1)m+1+t(h-1)+s} \to \frac{1}{c-1}$$

as $n \to \infty$. This completes the proof of Lemma 17, and thus of Theorem 2. $\square$

5 Purity with one forbidden connected minor

In this section we complete the proof of Theorem 1, showing that K_2, K_3, K_4 and P_3 are indeed the only connected graphs yielding pure H-free classes of graphs.

Lemma 18. *Let $H \notin \{K_2, K_3, K_4, P_3\}$ be a connected graph. Then* $\mathrm{Ex}(H)$ *is not a pure class of graphs.*

By Lemma 7, $\mathrm{Ex}(H)$ is not pure if the minimum degree $\delta(H) \geqslant 2$. By Lemma 13 and Observation 14, $\mathrm{Ex}(H)$ is not pure if there is a strong separating vertex. Note that, in particular, Lemma 13 and Observation 14 cover all graphs H such that some vertex of H has at least two leaves attached to it. Hence in this section we focus on graphs H with at least one leaf and with no strong separating vertex.

Remark 19. In what follows, for various classes of excluded minors H with $v(H) = h$ we prove that there exists some $n \in \mathbb{N}$ such that $\mathrm{gap}_H(n) > 0$. Since we consider graphs H with $\delta(H) = 1$, this immediately implies that for all $k \geqslant 0$ we have $\mathrm{gap}_H(n + k(h-1)) > 0$. Indeed, a disjoint union of an edge-maximal H-free graph G and k copies of K_{h-1} is again an edge-maximal H-free graph.

Lemma 20. *Let H be a graph on $h > 5$ vertices with at least two leaves, and with no strong separating vertex. Then the class of H-free graphs is not pure.*

Proof. Let G_1 be the union of K_{h-1} and an isolated vertex. Clearly G_1 is H-free and is maximal since H has leaves. Also, $e(G_1) = \binom{h-1}{2}$.

Let G_2 be formed from a K_{h-2} and a K_3 that have one vertex in common. To see that G_2 is H-free notice that the removal of the common vertex would leave no component of size at least $h-2$. Also, G_2 is edge-maximal H-free since adding an extra edge would allow us to place two leaves of H in the initial K_3. Obviously, $e(G_2) = \binom{h-2}{2} + 3$ and $e(G_1) > e(G_2)$ for all $h > 5$. $\square$

Observation 21. *The only connected graph H on 4 vertices with at least two leaves and with no strong separating vertex is P_4, the path on 4 vertices. However, let us show that* $\mathrm{limp}(P_4) = \frac{1}{2}$. *Indeed, every edge-maximal P_4-free graph has at most one isolated vertex, thus we have $M^-_{P_4}(n) \geqslant \frac{n-1}{2}$. Also, a perfect matching for n even, or a triangle plus a perfect matching on the remaining $n-3$ vertices for n odd, is edge-maximal P_4-free, so $M^-_{P_4}(n) \leqslant \frac{n+3}{2}$.*

On the other hand, any component of a P_4-free graph must be acyclic or unicyclic, as otherwise it would contain a C_4 or a bowtie graph *(two triangles with one common vertex) as a minor, thus it would not be P_4-free. Hence $M^+_{P_4}(n) \leqslant n$. Since a star on n vertices guarantees $M^+_{P_4}(n) \geqslant n-1$, we have* $\mathrm{limp}(P_4) = 1 - \frac{1}{2} = \frac{1}{2}$.

Observation 22. *The only connected graph H on 5 vertices with at least two leaves and with no strong separating vertex consists of a triangle on $\{1,2,3\}$ and two additional edges $\{1,4\},\{2,5\}$ (it is the so called* bull graph*). Let us show that* $\mathrm{limp}(H) = \frac{1}{2}$. *Since every edge-maximal H-free graph has at most one acyclic component, we have $M^-_H(n) \geqslant n-1$. On the other hand, for all $n \geqslant 5$ the cycle C_n is edge-maximal H-free so $M^-_H(n) \leqslant n$ for all $n \geqslant 5$.*

Let us show that $M^+_H(n) \leqslant \frac{3n}{2}$. Let C be a component of size at least 5 of an edge-maximal H-free graph (components of size at most 4 trivially have the edge-to-vertex ratio at most $3/2$). Observe that for any cycle in C, at most one vertex of the cycle has degree higher than 2. Otherwise we immediately can find a bull graph in C, or C contains the diamond graph *(K_4 less an edge) as a subgraph (hence also the bull, since $|C| \geqslant 5$). Thus C is obtained from a tree by adding disjoint cycles, and then identifying one vertex of the cycle with one vertex of the tree. Hence, to maximise the ratio $e(C)/v(C)$ we should take all cycles to be triangles, and the tree to be just one vertex. This gives $e(C)/v(C) \leqslant 3(n-1)/2n$. Therefore $M^+_H(n) \leqslant \frac{3n}{2}$.*

On the other hand, a disjoint union of $\lfloor n/4 \rfloor$ copies of K_4 is H-free, so we have $M^+_H(n) \geqslant \frac{3(n-3)}{2}$. This gives $\mathrm{limp}(H) = \frac{3}{2} - 1 = \frac{1}{2}$.

We claim that the only graphs that remain to be checked are graphs H with exactly one leaf v and such that the graph $H' = H - v$ is 2-connected. Indeed, if $\delta(H') = 1$ then either H has two leaves or the unique vertex of degree 1 in H' is the neighbour u of v in H. In the latter case, let the unique neighbour of u in H' be w. Then either all the components of $H - w$ have size at most $v(H) - 3$ (so w is a strong separating vertex), or H is a P_4. Since neither of these is possible, we have $\delta(H') \geqslant 2$. But then, if H' has connectivity 1 then clearly H has a strong separating vertex. This establishes our claim.

Unfortunately, it will require several more steps to deal with the case in the claim.

Lemma 23. *Let H be a graph on $h \geqslant 5$ vertices consisting of a clique on $h-1$ vertices and one pendant edge. Then the class of H-free graphs satisfies* $\operatorname{limp}(H) \geqslant \frac{h-4}{2}$.

Proof. Let $n = m(h-1) + k$, $0 \leqslant k \leqslant h-2$. Let G_1 be the union of m disjoint copies of K_{h-1} and one copy of K_k. Clearly G_1 is edge-maximal H-free. Also, $v(G_1) = n = m(h-1) + k$ and

$$e(G_1) = m\binom{h-1}{2} + \binom{k}{2} \leqslant \frac{h-2}{2}n.$$

We construct a denser n-vertex graph G_2 as follows. We start with a clique on $h-4$ vertices and a cycle C_{n-h+4}. We then build a complete bipartite graph between the clique and the cycle (see Figure 3). To see that G_2 is H-free note that in order to obtain a clique on $h-1$ vertices we would need to contract the cycle C_{n-h+4} to a triangle, but then we would only have $h-1$ vertices left in the graph. But

$$e(G_2) = \binom{h-4}{2} + n - h + 4 + (h-4)(n-h+4) = (h-3)n - \frac{(h-1)(h-4)}{2}.$$

Hence

$$\frac{e(G_2) - e(G_1)}{n} \geqslant \frac{n\left(h - 3 - \frac{h-2}{2}\right) - \frac{(h-1)(h-4)}{2}}{n} \to \frac{h-4}{2}$$

as $n \to \infty$. $\square$

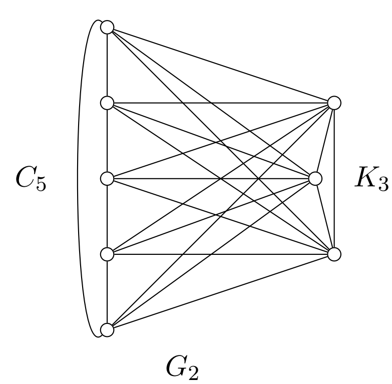

Figure 3: Graph G_2 as defined in Lemma 23 for $h = 7$ and $n = 8$.

Observation 24. *In Lemma 23 we prove linear impurity of* $\mathrm{Ex}(H)$ *when the clique in* H *contains at least* 4 *vertices. Indeed, when* H *is the* pan graph *on* 4 *vertices, consisting of a triangle and a pendant edge, then* $\mathrm{Ex}(H)$ *is near-pure with* $\mathrm{gap}_H(n) = 1$ *for all* $n \geqslant 4$*. To see this, observe that every connected component of an* H*-free graph is either a cycle or a tree, and an edge-maximal* H*-free graph has at most one acyclic component (in fact, this component can be any tree except a path* P_m *on* $m \geqslant 3$ *vertices which we could close to a cycle without creating an* H*-minor).*

Lemma 25. *Let the connected graph* H *have exactly one leaf* v*, with neighbour* u*. Let* $H' = H - v$ *satisfy* $\delta' := \delta(H') \geqslant 2$*, and suppose that there is a vertex* $w \neq u$ *in* H' *with* $\deg_{H'}(w) = \delta'$*. Then the class* $\mathrm{Ex}(H)$ *is not pure.*

Proof. By Lemma 23, we may assume that H' is not complete. Thus $2 \leqslant \delta' \leqslant h-3$, and so $h \geqslant 5$.

Let G_1 be the graph on vertex set $[h+1]$ constructed as follows. Start with a clique on $\{1, 2, \ldots, h-2\}$. Next, for $i = 1, 2, 3$, connect the vertex $h-2+i$ to $1, 2, \ldots, \delta'-2$, as well as to $\delta'-2+i$ (see Figure 4). Clearly, $e(G_1) = \binom{h-2}{2} + 3(\delta'-1)$. To see that G_1 is H-free, note that it has an independent set of 3 vertices each of degree $< \delta'$, so after one edge-contraction there must still be at least two vertices of degree $< \delta'$.

Next we show that G_1 is edge-maximal H-free. Suppose that we add an edge e to G_1, where wlog e is incident to vertex $h-1$. There are now two cases. (a) Suppose that e is incident to h or $h+1$, wlog to h. Contract e to form a new vertex called w, and place v at $h+1$. If $uw \in E(H)$ then place u at vertex 1; and if not then place u at vertex $\delta'+1$. (b) Suppose that e is incident to a vertex in $\{\delta', \ldots, h-2\}$. Then e is not incident to at least one of vertices $\delta', \delta'+1$, wlog the former. Place w at vertex $h-1$, place v at h, and delete vertex $h+1$. If $uw \in E(H)$ then place u at vertex 1; and if not then place u at vertex δ'.

We construct the graph G_2 as a disjoint union of K_{h-1} and the edge $\{h, h+1\}$. Clearly G_2 is edge-maximal H-free, and $e(G_2) = \binom{h-1}{2} + 1$.

We have $e(G_1) \neq e(G_2)$ unless $\delta' = (h+2)/3$. Note that the smallest value of h for which this could hold with both h and δ' being integers is $h = 7$ (which gives $\delta' = 3$).

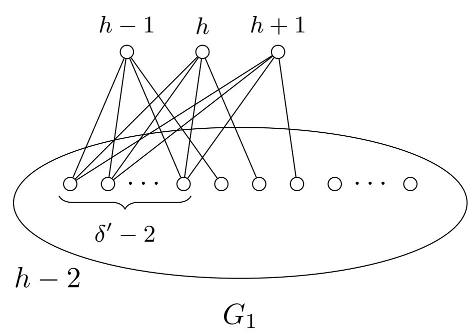

Figure 4: Graph G_1 as defined in Lemma 25.

If we have $\delta' = (h+2)/3$ and $h > 7$, which implies that $\delta' - 2 = (h-4)/3 < h-6$, then we alter our constructions of G_1 and G_2 as follows. We take the graph G_1' consisting of K_{h-2} together with four extra vertices $h-1, h, h+1, h+2$ such that $h-2+i$ is connected to $1, 2, \ldots, \delta'-2, \delta'-2+i$

for $1 \leqslant i \leqslant 4$ (observe that $\delta' - 2 + 4 < h - 2$). Then we have

$$e(G_1') = \binom{h-2}{2} + 4(\delta' - 1) = \binom{h-2}{2} + 4\frac{h-1}{3}.$$

We compare G_1' to G_2' formed of disjoint copies of K_{h-1} and K_3, which has $e(G_2') = \binom{h-1}{2} + 3$. We then obtain

$$e(G_1') - e(G_2') = \frac{h-7}{3} > 0.$$

In the last remaining case where $h = 7, \delta' = 3$, we alter the construction a little bit. We build G_1'' on 10 vertices, starting from a Hamiltonian cycle of edges $\{i, i+1\}$ (as usual, we identify vertex 11 with 1). Then we add edges to make the even vertices into a clique. Thus we have $e(G_1'') = 20$. Graph G_1'' is edge-maximal H-free for exactly the same reasons as our previous constructions: the odd vertices all have degree $2 < \delta'$ and form an independent set, while the union of the neighbourhoods of any two of them has size either 3 or 4. We take G_2'' to be a disjoint union of K_6 and K_4, so clearly $e(G_2'') = 15 + 6 = 21$. This completes the proof of the lemma. □

Lemma 26. *Let H be a graph on $h \geqslant 6$ vertices with exactly one leaf v and such that the graph $H' = H - v$ has connectivity k for some $2 \leqslant k \leqslant h - 4$. Also, let the unique neighbour of v in H be u. If $\deg_H(u) = \delta(H') + 1$ then the class $\mathrm{Ex}(H)$ is not pure.*

Proof. We use a similar construction as in Lemma 12. Let A and B be $(h-2)$-sets with $|A \cap B| = k - 1$. Let G_1 be the union of a clique on A and a clique on B. This graph clearly has $|A \cup B| = 2h - k - 3 \geqslant h + 1$ vertices and $2\binom{h-2}{2} - \binom{k-1}{2}$ edges. To see why G_1 is H-free we note that it is H'-free, since we cannot have a model of H' within A or within B.

On the other hand, adding a single edge to G_1 and contracting it gives us a graph on at least h vertices consisting of a union of two cliques on $h-2$ vertices each that overlap in k vertices. Let us show that this graph is not H-free. We can obviously find H' in this graph as a subgraph because H' has $h - 1$ vertices and connectivity (exactly) k. The only time we need to worry about being able to add the leaf v to our minor is when all but one of the vertices of H' are located in A and only one in $B \setminus A$ (or vice-versa). But then that one vertex (say vertex x) would have degree at most k in H', so $\deg_{H'}(x) = \delta(H') = k$, and now we can place vertex u at x.

We take G_2 to be a disjoint union of K_{h-1} and K_{h-k-2}, which is clearly seen to be edge-maximal H-free. We have $e(G_2) = \binom{h-1}{2} + \binom{h-k-2}{2}$. The only integer solutions to $e(G_1) = e(G_2)$ are $h = 1, k = 0$ and $h = 5, k = 2$. This completes the proof. □

The next lemma fills one of the gaps left by Lemma 26.

Lemma 27. *Let H be a graph on $h \geqslant 6$ vertices with exactly one leaf v and such that the graph $H' = H - v$ has connectivity $h - 3$. Then the class $\mathrm{Ex}(H)$ satisfies $\mathrm{limp}(H) \geqslant \frac{h-5}{2} > 0$.*

Proof. For each $m \geqslant 2$, let $n = h - 4 + 2m$ and let the n-vertex graph G_1 be the union of m cliques, each on $h - 2$ vertices, that overlap in a common set of $h - 4$ vertices (see Figure 5). As in Lemma 26, G_1 is H-free and has size $e(G_1) = \binom{h-4}{2} + m(2(h-4) + 1)$. Let G_2 be a disjoint union of (as many a possible) cliques on $h - 1$ vertices and possibly one smaller clique containing the remaining k vertices, where $0 \leqslant k \leqslant h - 2$. Then G_2 is edge-maximal H-free.

It is easy to see that, as $n \to \infty$, we have $e(G_1) = (1 + o(1))(2(h-4)+1)n/2$, while $e(G_2) = (1 + o(1))(h-2)n/2$. Thus

$$\frac{e(G_1) - e(G_2)}{n} \to \frac{h-5}{2},$$

as desired. Note that we do not need G_1 to be edge-maximal here, and so for $n = h - 4 + 2m + 1$ we can just take G_1 plus an isolated vertex. □

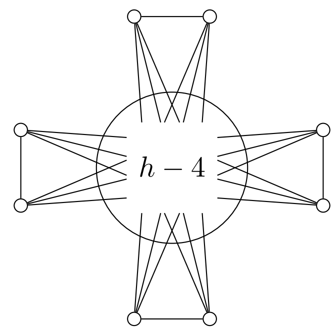

Figure 5: Graph G_1 with $m = 4$ as defined in Lemma 27.

Observation 28. *The last remaining graphs that we need to consider are the connected graphs H on 5 vertices that have exactly one leaf v and are such that $H - v$ is 2-connected but is not a complete graph. Up to isomorphism, there are exactly three such graphs H, and they each give rise to classes satisfying* $\text{limp}(H) \geqslant \frac{1}{2}$. *Consider $n > 4$. In each case, our 'dense' example is the disjoint union of $\lfloor n/4 \rfloor$ copies of K_4, together with a copy of K_t where $t = n - 4\lfloor n/4 \rfloor$ if $4 \nmid n$, which is an edge-maximal H-free graph with $3n/2 + O(1)$ edges.*

1. *Let H_1 be C_4 with an added leaf. Then C_n is an edge-maximal H_1-free graph with n edges. Hence* $\text{limp}(H_1) \geqslant \frac{1}{2}$.

2. *Let H_2 be a diamond (K_4 minus an edge), with an added leaf adjacent to a vertex of degree 2 of the diamond. Then the graph obtained from C_{n-1} by adding one vertex and joining it to two adjacent vertices on the cycle is an edge-maximal H_2-free graph with $n+1$ edges; and it follows that* $\text{limp}(H_2) \geqslant \frac{1}{2}$.

3. *Let H_3 be a diamond (K_4 minus an edge), with an added leaf adjacent to a vertex of degree 3 of the diamond. Then the graph obtained from K_4 by subdividing one edge $n-4$ times (or equivalently, from C_{n-1} by adding a vertex and joining it to three consecutive vertices on the cycle) is an edge-maximal H_3-free graph with $n+2$ edges; and it follows* $\text{limp}(H_3) \geqslant \frac{1}{2}$.

Remark 29. In fact, it can be shown that $\text{limp}(H_1) = 1/2$, $\text{limp}(H_2) = 1$ and $\text{limp}(H_3) = 2/3$ (see Appendix A in the arXiv version of this paper [12]).

This completes the proof of Lemma 18, and thus of Theorem 1.

6 Forbidding several minors or disconnected minors

We start this section by generalising Lemma 12 to a case where we may have more than one excluded minor, and the excluded minors need not be connected. For our proof to work, the forbidden set $\mathcal{H}$ needs to satisfy specific and rather strict conditions. Roughly, we require that one component of one excluded minor is 'smallest' in several senses. However, cases like $\mathcal{H} = \{mC_h\}$ (that is, m disjoint copies of the cycle C_h) for $h \geqslant 4$, or $\mathcal{H} = \{K_{2,3}, C_5\}$, can be dealt with using the following result, which shows that in these cases the classes $\mathrm{Ex}(\mathcal{H})$ are linearly impure.

Lemma 30. *Let $\mathcal{H} = \{H_1, H_2, \dots, H_m\}$ be a set of $m \geqslant 1$ excluded minors. Let $t_1, \dots, t_m$ be positive integers. For each $1 \leqslant i \leqslant m$, let $H_i = \bigcup_{j=1}^{t_i} H_i^j$; that is, let each graph H_i be a disjoint union of connected graphs H_i^j for $1 \leqslant j \leqslant t_i$. Assume that the following conditions hold:*

1. $v(H_1) = \min_{1 \leqslant i \leqslant m} v(H_i)$ *and* $v(H_1^1) = \min\left\{v\left(H_i^j\right) : 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant t_i\right\} := h$.
2. $\delta(H_1^1) = \min_{1 \leqslant i \leqslant m} \delta(H_i) := \delta$ *and δ satisfies* $2 \leqslant \delta \leqslant v\left(H_1^1\right) - 2$.
3. *Taking k_i^j to be the connectivity of H_i^j we have* $k_1^1 = \min\left\{k_i^j : 1 \leqslant i \leqslant m, 1 \leqslant j \leqslant t_i\right\} := k$.

Then we have $\mathrm{limp}(\mathcal{H}) \geqslant \frac{1}{2h}$.

Proof. The proof of this lemma is nearly identical to the proof of Lemma 12 when we take $H = H_1^1$. We amend the constructions of graphs G_1 and G_2 by adding to both of them a clique of size $v(H_1)-1$ and identifying $k-1$ vertices of that new clique with the "small cut" I consisting of the central $k-1$ vertices in the previously built graphs. By our assumptions on $\mathcal{H}$, these graphs are $\mathcal{H}$-free, and adding an arbitrary edge to the graph creates an H_1-minor: we trivially find the graphs $H_1^2, \dots, H_1^{t_1}$ in the "large" clique on $v(H_1) - 1$ vertices (if $t_1 \geqslant 2$), and H_1^1 is created like H was in Lemma 12. □

Remark 31. One interesting case that is not covered by Lemma 30 is $\mathrm{Ex}(mK_3)$, i.e., the class of graphs with m disjoint triangles excluded for some $m \geqslant 2$. However, building on the work of Corradi and Hajnal [1] on the number of disjoint cycles in graphs of given density, it was observed in [16] that every graph G with $e(G) \geqslant (2m-1)v(G)$ contains mK_3 as a minor. Moreover, this bound is asymptotically sharp as demonstrated by the complete bipartite graph $G = K_{2m-1,n-2m+1}$.

On the other hand, any maximal mK_3-free graph can have at most one acyclic component, so $M^-_{mK_3}(n) \geqslant n-1$, and by analysing G constructed from K_{3m-1} by adding $n-3m+1$ pendant edges we see that $M^-_{mK_3}(n) = n + O(1)$. Hence we can conclude that $\mathrm{limp}(mK_3) = 2m-2$ for all $m \geqslant 2$.

So far we have seen only two graphs H such that the class $\mathrm{Ex}(H)$ is near pure. Namely, this happens when H is the claw or the pan graph. However, in both cases $\mathrm{gap}_H(n) \leqslant 1$ and it is unclear whether there are more connected graphs H such that $\mathrm{Ex}(H)$ is near pure, and if the answer to that question is positive, whether $\mathrm{gap}_H(n)$ can take arbitrarily large values (or in fact, any value larger than 1). This is the case when we forbid more complex sets of graphs. In the following proposition we only take $t \geqslant 16$ to avoid complications in the statement that would make the conclusions more difficult to observe.

Proposition 32. *Let $t \geqslant 16$ be an integer, and let $\mathcal{H} = \{K_{1,t}, 2K_{1,3}\}$. Then the class* $\mathrm{Ex}(\mathcal{H})$ *is near-pure with* $t - 10 \leqslant \mathrm{gap}_{\mathcal{H}}(n) \leqslant t - 1$.

Proof. We first claim that for all $n \geqslant 1$, every edge-maximal $\mathcal{H}$-free graph G satisfies $e(G) \geqslant n-1$. Indeed, every $\mathcal{H}$-free graph must have at most one component that is not a cycle nor a path to avoid creating a $2K_{1,3}$-minor. Consequently, a maximal $\mathcal{H}$-free graph has at most one acyclic component because we could connect one of the endpoints of any path to a leaf of any other tree without creating any of the forbidden minors.

Now let G be an edge-maximal $\mathcal{H}$-free graph on $n \geqslant 4$ vertices. Let $\Delta = \Delta(G)$ denote the maximum degree of a vertex in G. Clearly $\Delta \geqslant 3$. We consider three cases – when $3 \leqslant \Delta \leqslant 5$, $\Delta = 6$ and $\Delta \geqslant 7$. In each case, let v be a vertex of degree Δ, let V_i denote the set of vertices at distance i from v in G, and let W_2 denote $\bigcup_{i \geqslant 2} V_i$, the set of vertices at distance at least 2 from v (recall that G might not be connected).

Suppose first that $3 \leqslant \Delta \leqslant 5$. Each vertex in V_1 can have at most 2 edges to V_2 (this is immediate if $\Delta = 3$, while for $\Delta \geqslant 4$ follows from the fact that otherwise we have a $2K_{1,3}$ minor), so there are at most 2Δ edges between V_1 and V_2. Similarly, each vertex in W_2 can have at most 2 edges to vertices in W_2. Hence the degree sum is at most

$$(\Delta + 1) \cdot \Delta + 2\Delta + (n - \Delta - 1) \cdot 2 = 2n + (\Delta + 1)\Delta - 2 \leqslant 2n + 28.$$

If $\Delta = 6$ then each vertex in W_2 has degree at most 2, so the degree sum is at most $7 \cdot 6 + (n-7) \cdot 2 = 2n + 28$. (Observe that the disjoint union of K_7 and C_{n-7} achieves this bound.) If $\Delta \geqslant 7$ then also each vertex in V_1 has degree at most 3, so the degree sum is at most

$$\Delta + \Delta \cdot 3 + (n - \Delta - 1) \cdot 2 = 2n + 2\Delta - 2;$$

and since also $\Delta \leqslant t - 1$ this is at most $2n + 2t - 4$. Thus for $\Delta \leqslant 6$ we have $e(G) \leqslant n + 14$, and for $\Delta \geqslant 7$ we have $e(G) \leqslant n + t - 2$. Hence, since $t \geqslant 16$, we always have $e(G) \leqslant n + t - 2$. It follows that $\mathrm{gap}_{\mathcal{H}}(n) \leqslant n + t - 2 - (n - 1) \leqslant t - 1$.

The upper bound $n + t - 2$ is achieved. Let G' be the disjoint union of the t-vertex wheel (a C_{t-1} plus a central vertex) and C_{n-t}. Then G' is (edge-maximal) $\mathcal{H}$-free, and $e(G') = n + t - 2$.

We now find a much smaller edge-maximal $\mathcal{H}$-free graph. Start with K_5, choose two vertices u and v in the K_5 and add two new vertices x and y, both adjacent to both of u and v (this gives the total of $10 + 4$ edges so far). Finally we add $n - 7$ vertices which form a path of $n - 6$ edges between x and y. The resulting graph is edge-maximal $\mathcal{H}$-free with $n + 8$ edges. Hence $\mathrm{gap}_{\mathcal{H}}(n) \geqslant n + t - 2 - (n + 8) = t - 10$. □

7 Concluding remarks and open problems

When the connected graph H satisfies $\delta(H) = 1$ then a natural example of a leaf-and-edge-maximal H-free graph is a union of disjoint copies of K_{h-1}, where $h = v(H)$. It often turns out to be a "dense" example of such a graph, though in some cases we can find denser H-free graphs (see, e.g., Lemmas 23 and 27). In general, it appears that the graphs with minimum degree 1 can cause us most trouble analysing their purity, as illustrated in the following example.

Example 33. Let the graph H on $h \geqslant 6$ vertices consist of a clique on $h-2$ vertices and two pendant, non-incident edges. Two obvious examples of edge-maximal H-free graphs are a union of disjoint cliques each on $h-1$ vertices, and a union of cliques each on $h-2$ vertices that share one common vertex. It is easy to check that in both cases the density of these graphs tends to $(h-2)/2$ as the number of cliques constituting them tends to infinity. As finding other edge-maximal H-free graphs appears non-trivial, this might suggest that $\mathrm{Ex}(H)$ is near-pure.

This is, however, not true and the following sparse construction, after comparing with the disjoint union of copies of K_{h-1} (together with a smaller clique if necessary), will show that we have

$$\mathrm{limp}(H) \geqslant \frac{h-4}{2} \tag{6}$$

for all $h \geqslant 6$. Let G' be a subdivision of K_{h-2}, obtained from K_{h-2} by subdividing every edge at least once. Let H^- be H less a leaf (that is, K_{h-2} plus one pendant edge). Clearly, adding an edge joining two vertices created through subdivisions of the same edge of K_{h-2} creates an H^--minor in G'. In fact, by case analysis, it is easy to check that G' is leaf-and-edge-maximal H^--free (it is enough to check it for $h = 6$ because the edge we add to G' can be "wrapped" in a K_4 containing it). Importantly, when we add an edge to G' then we always have at least two choices of an original vertex of K_{h-2} to which we can attach a leaf of the H^--minor (see Figure 6). For n large enough, it is then enough to take a union of two such (not necessarily identical) subdivisions of K_{h-2} of sizes that sum up to $n+1$, and connect them by picking an original vertex of K_{h-2} from each subdivided graph and identifying them. The resulting graph is edge-maximal H-free with density tending to 1 as n tends to infinity. This establishes (6), and completes the example.

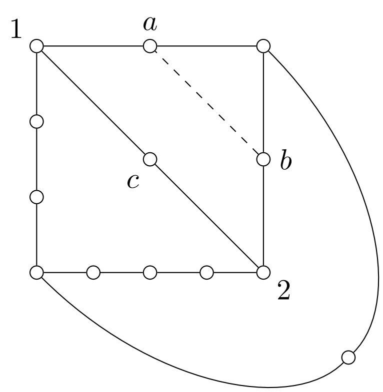

Figure 6: Graph G' for $h = 6$ which is a subdivision of K_4. When, for example, we add the edge $\{a, b\}$ to G', we can contract $\{1, a\}$ and $\{2, b\}$ and delete either $\{1, c\}$ or $\{2, c\}$, hence finding a minor consisting of a K_4 and a pendant edge attached to either 1 or 2.

Let us recall the definition of the set

$$E_{\mathcal{H}}(n) = \{e(G) : v(G) = n \text{ and } G \text{ is an edge-maximal } \mathcal{H}\text{-free graph}\}.$$

The main objects of study of this paper were the extreme values of the set $E_{\mathcal{H}}(n)$, i.e., $M^-_{\mathcal{H}}(n)$ and $M^+_{\mathcal{H}}(n)$. However, once we know that $\mathrm{Ex}(\mathcal{H})$ is not pure (i.e., that $M^-_{\mathcal{H}}(n) \neq M^+_{\mathcal{H}}(n)$ for some

n), we can ask additional questions about the structure of $E_{\mathcal{H}}(n)$. As a test case, let us consider $\mathcal{H} = \{K_5\}$.

Recall again the result of Wagner, who proved that edge-maximal K_5-free graphs are obtained as 2- or 3-clique-sums of planar graphs and of the Wagner graph W_8 (the sums must be "maximal", in particular, we only take a 2-clique-sum of two graphs along an edge if that edge in not in any triangle in at least one of those graphs).

Consequently, taking clique-sums of only planar graphs, always leads to building edge-maximal K_5-free graphs on n vertices and $3n-6$ edges. Therefore, the first interesting case is $n = 8$. The only possible edge-numbers of edge-maximal K_5-free graphs on 8 vertices are 12 (the Wagner graph W_8) and 18 (3-clique-sums of planar graphs). For $n = 9$ these edge-numbers are 14 (W_8 glued to a triangle) and 21, while for $n = 10$ it can be 16 (W_8 plus two triangles glued to different edges of W_8), 17 (W_8 and K_4 glued along an edge) or 24. Continuing this way, for $n = 14$, we can build edge-maximal K_5-free graphs with any number of edges between 23 and 29, as well as 36.

More generally, taking $0 \leqslant i \leqslant 5$ and $n = 6k+2+i$ large, we have $M^-_{K_5}(n) = \frac{11(n-2-i)}{6} + 1 + 2i$, and $E_{K_5}(n)$ contains all values between $M^-_{K_5}(n)$ and $3n-13$ (obtained, e.g., using one copy of W_8 glued along an edge with a triangulation on $n-6$ vertices), as well as $3n-6$. Hence in general we don't have $E_{\mathcal{H}}(n)$ forming an interval, but do we always have $\operatorname{gap}_{\mathcal{H}}(n) - |E_{\mathcal{H}}(n)| = O(1)$, or at least is it always the case that if $\operatorname{Ex}(\mathcal{H})$ is linearly impure then $|E_H(n)|/\operatorname{gap}_H(n) \to 1$ as n tends to infinity?

We have determined the complete list of four connected graphs H leading to pure minor-closed classes $\operatorname{Ex}(H)$. For connected H we also know that $\operatorname{Ex}(H)$ is linearly impure if

- $\delta(H) \geqslant 2$, see Lemma 12, or
- H has a strongly separating vertex (except for the claw $K_{1,3}$), see Lemma 13, or
- H is the path P_4 (Observation 21), the bull graph (Observation 22), a clique on at least four vertices with an additional one (Lemma 23) or two leaves (see the discussion at the beginning of this section), or
- H consists of a clique on at least five vertices minus a matching, plus a pendant edge, see Lemma 27, or
- H is one of the three graphs discussed in Observation 28.

Additionally, we know that $\operatorname{Ex}(H)$ is near-pure with $\operatorname{gap}_H(n) = 1$ if H is the claw (Observation 14) or the pan graph (Observation 24). What about the remaining connected graphs H which are not pure? Are there any more near-pure minor-closed classes $\operatorname{Ex}(H)$ for some connected graph H? Can we find an example such that $\operatorname{gap}_H(n) \geqslant 2$ for some n?

We defined $\operatorname{limp}(\mathcal{H}) = \liminf_{n\to\infty} \operatorname{gap}_{\mathcal{H}}(n)/n$. Theorem 4 says that $\operatorname{gap}_{\mathcal{H}}(n)/n$ tends to a limit if all graphs in $\mathcal{H}$ are 2-connected, so that in this case we could define $\operatorname{limp}(\mathcal{H})$ as the limit of $\operatorname{gap}_{\mathcal{H}}(n)/n$. Do we always have $\operatorname{gap}_{\mathcal{H}}(n)/n \to \operatorname{limp}(\mathcal{H})$?

Finally, what about minor-closed classes with two or more connected excluded minors, whose analysis we started in Section 6: which are the pure classes, and are all such classes pure or near-pure or linearly impure? For example, the classes $\operatorname{Ex}(K_5, K_{3,3})$ of planar graphs, $\operatorname{Ex}(K_3, K_{1,3})$ of 'forests of paths', $\operatorname{Ex}(2K_2, K_3)$ of a star and isolated vertices, $\operatorname{Ex}(\text{Diamond}, \text{Bowtie})$ of graphs consisting of unicyclic and acyclic components, and $\operatorname{Ex}(K_4, K_{2,3})$ of outerplanar graphs are all pure; while for all $t \geqslant 5$, the class $\mathrm{Ex}(C_t, K_{1,3})$ where each component is a path or a short cycle, is near-pure with $\mathrm{gap}(n) = 1$ for all $n \geqslant \max\{t, 6\}$ (two examples of $\{C_t, K_{1,3}\}$-free edge-maximal graphs are a path on n vertices and $n-1$ edges, and a union of disjoint copies of C_3 and C_4 with total of n vertices and n edges, which exists for all $n \geqslant 6$ by Fact 8). Note that $\mathrm{Ex}(C_4, K_{1,3})$ is an interesting case with $\mathrm{gap}(3k) = 1$ for all $k \geqslant 2$, and $\mathrm{gap}(n) = 0$ otherwise. Obviously, similar questions could be asked about excluding disconnected minors.

Acknowledgements We would like to thank Andrius Vaicenavicius for stimulating discussions during the course of this work. We would also like to thank the referee for careful comments.

References

[1] K. Corradi, A. Hajnal, On the maximal number of independent circuits in a graph, Acta Mathematica Academiae Scientiarum Hungarica 14 (1963) 423–443.

[2] E. Csóka, I. Lo, S. Norin, H. Wu, L. Yepremyan, The extremal function for disconnected minors, Journal of Combinatorial Theory Series B 126 (2017) 162–174.

[3] Z. Dvořák, S. Norine, Small graph classes and bounded expansion, Journal of Combinatorial Theory Series B 100 (2010) 171–175.

[4] W. Fernandez de la Vega, On the maximum density of graphs which have no subcontraction to K^s, Discrete Mathematics 46 (1983) 109–110.

[5] S. Gerke, C. McDiarmid, A. Steger, A. Weißl, Random planar graphs with n nodes and a fixed number of edges, in: Proceedings of the ACM–SIAM Symposium on Discrete Algorithms (SODA 2005), 2005, pp. 999–1007.

[6] F. Harary, P. Kainen, S. A.J., A. White, A maximal toroidal graph which is not a triangulation, Mathematica Scandinavica 33 (1973) 108–112.

[7] V. Kolchin, Random Mappings, Optimization Software Inc., New York, 1986.

[8] A. Kostochka, Lower bound on the Hadwiger number of graphs by their average degree, Combinatorica 4 (1984) 307–316.

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

[10] W. Mader, Homomorphiesätze für Graphen, Mathematische Annalen 178 (1968) 154–168.

[11] C. McDiarmid, Random graphs from a minor-closed class, Combinatorics, Probability and Computing 18 (2009) 583–599.

[12] C. McDiarmid, M. Przykucki, On the purity of minor-closed classes of graphs, preprint, `http://arxiv.org/abs/1608.08623`.

[13] C. McDiarmid, D. Wood, Edge-maximal graphs on surfaces, Canadian Journal of Mathematics 70 (2018) 925–942.

[14] J. Myers, A. Thomason, The extremal function for non-complete minors, Combinatorica 25 (2005) 723–753.

[15] S. Norine, P. Seymour, R. Thomas, P. Wollan, Proper minor-closed families are small, Journal of Combinatorial Theory Series B 96 (2006) 754–757.

[16] B. Reed, D. Wood, Forcing a sparse minor, Combinatorics, Probability and Computing 25 (2016) 300–322.

[17] N. Robertson, P. Seymour, Graph minors I-XX, Journal of Combinatorial Theory Series B.

[18] A. Thomason, The extremal function for complete minors, Journal of Combinatorial Theory Series B 81 (2001) 318–338.

[19] J. van Lint, R. Wilson, A Course in Combinatorics, Cambridge University Press, 1992.

[20] K. Wagner, Über eine Eigenschaft der ebenen Komplexe, Mathematische Annalen 114 (1937) 570–590.