RANDOM GRAPHS FROM A BLOCK-STABLE CLASS

COLIN MCDIARMID AND ALEXANDER SCOTT

Abstract. A class of graphs is called block-stable when a graph is in the class if and only if each of its blocks is. We show that, as for trees, for most \( n \)-vertex graphs in such a class, each vertex is in at most \((1+o(1)) \log n / \log \log n\) blocks, and each path passes through at most \(5(n \log n)^{1/2}\) blocks. These results extend to ‘weakly block-stable’ classes of graphs.

1. Introduction

A block in a graph is a maximal 2-connected subgraph or the subgraph formed by a bridge or an isolated vertex. (A bridge is an edge the deletion of which increases the number of components.) Call a class of graphs (always assumed to be closed under isomorphism) block-stable when a graph \( G \) is in the class if and only if each block of \( G \) is in the class. For example, the class of all forests is block-stable and more generally so is any minor-closed class of graphs with 2-connected excluded minors. A different example is the class of all graphs in which each block is a triangle.

In this paper, we are interested in typical properties of graphs from such a class. Indeed we are interested in more general classes of graphs, namely ‘weakly block-stable’ classes. To define this notion, let us first introduce an equivalence relation on (finite) graphs, which is natural in this context. Given connected graphs \( G \) and \( H \), let \( G \sim H \) if they have the same vertex set and the same number of blocks of each kind (up to isomorphism). Given general graphs \( G \) and \( H \), let \( G \sim H \) if we can list the components as \( G_1, \ldots, G_k \) and \( H_1, \ldots, H_k \) (for some \( k \)) so that \( G_i \sim H_i \) for each \( i \). We say that a class \( \mathcal{A} \) of graphs is weakly block-stable if whenever \( G \in \mathcal{A} \) and \( H \sim G \) then \( H \in \mathcal{A} \). Clearly a block-stable class is weakly block-stable, but not conversely.

As mentioned above, we are most interested in typical properties of graphs from a block-stable class, but our results extend to weakly block-stable classes of graphs, and indeed that is the natural context for our investigations. In particular, we are interested in the maximum
number of blocks containing a given vertex, and the maximum number of blocks a path can pass through.

For a connected graph $G$, these are essentially properties of the block tree $\text{BT}(G)$ of $G$, which is the bipartite graph with a node $x_v$ for each vertex $v$ and a node $y_B$ for each block $B$, where $x_v$ and $y_B$ are adjacent if and only if $v \in B$. (There is an alternative slimmer version of the block tree, in which vertices which are not cut-vertices are ignored.) If $G$ is not necessarily connected, we let the block forest $\text{BF}(G)$ be the disjoint union of the block trees of the components.

Given a set $\mathcal{A}$ of graphs, for each positive integer $n$ let $\mathcal{A}_n$ denote the set of graphs in $\mathcal{A}$ on vertex set $[n] := \{1, \ldots, n\}$. Also, let $R_n \in_u \mathcal{A}$ mean that $R_n$ is sampled uniformly from $\mathcal{A}_n$. When we use this notation we implicitly consider only integers $n$ such that $\mathcal{A}_n$ is non-empty.

Now suppose that $\mathcal{A}$ is weakly block-stable and $\mathcal{P}$ is any graph property. Note that $\mathcal{A}_n$ may be partitioned into the distinct equivalence classes $[G]$ for $G \in \mathcal{A}_n$ (where the equivalence relation is graph isomorphism). Thus if we can show for each $G \in \mathcal{A}_n$ that $\mathbb{P}(R \in \mathcal{P}) \geq t$ when $R \in_u [G]$, then it will follow that $\mathbb{P}(R_n \in \mathcal{P}) \geq t$ when $R_n \in_u \mathcal{A}$. We say that a sequence $(E_n)$ of events holds with high probability (whp) if $\mathbb{P}(E_n) \to 1$ as $n \to \infty$.

Let $\mathcal{T}$ denote the class of trees, and let $T_n \in_u \mathcal{T}$. It will be natural for us to compare the block tree $\text{BT}(R_n)$ with $T_n$, and to compare the associated degree sequences. Given two random variables $X$ and $Y$, we say that $X$ is stochastically at most $Y$ if $\mathbb{P}(X \geq t) \leq \Pr(Y \geq t)$ for every real number $t$. More generally, for two sequences $X = (X_1, \ldots, X_n)$ and $Y = (Y_1, \ldots, Y_n)$ of random variables, we say that $X$ is stochastically at most $Y$ if $\mathbb{E}[f(X)] \leq \mathbb{E}[f(Y)]$ for each non-decreasing integrable real-valued function $f$ on $\mathbb{R}^n$.

We are interested in typical properties of $R_n$ when $\mathcal{A}$ is a (weakly) block-stable class, or is the set of connected graphs in such a class; and in particular we focus on degrees of nodes $x_v$ and on long paths in $\text{BT}(R_n)$ or $\text{BF}(R_n)$. We present our main results in the next two subsections.

Consider briefly a related but distinct setting, where there are results of a different nature. Suppose that our block-stable class is the class of all series-parallel graphs or another ‘subcritical’ graph class, or it is the class of planar graphs, or another such class where we know the corresponding generating functions suitably well. In such cases, we may be able to deduce precise asymptotic results, for example about vertex degrees or the numbers and sizes of blocks, by using analytic techniques or by analysing Boltzmann samplers: see for example [2], [6], [7], [8], [10], [11], [12], [13], [14], [22], [23], and for an authoritative
recent overview of related work on random planar graphs and beyond see the article [21] by Marc Noy.

The main tools we use in our proofs are a tree-like graph $\tilde{G}$ related to the block-tree of a graph $G$, and a corresponding tree $T_G$, together with a slight extension of Prüfer coding: these are discussed in the next section. The proofs of Theorem 1.1 and Theorem 1.2 are completed in Sections 3 and 4 respectively, and we make some brief concluding remarks in Section 5.

1.1. Block-degrees of vertices. First consider the number of blocks in $G$ containing a vertex $v$, that is, the degree of the node $x_v$ in the block tree $BT(G)$: let us call this number the block-degree of $v$, and denote it by $\tilde{d}_G(v)$. Observe that if $G$ is a tree (with at least two vertices) then $\tilde{d}_G(v)$ is just the degree $d_G(v)$ of $v$ in $G$. Denote the maximum of the numbers $\tilde{d}_G(v)$ by $\tilde{\Delta}(G)$. Recall that, for $T_n \in \mathcal{T}$, the maximum degree $\Delta$ satisfies

$$\Delta(T_n) \sim \log n / \log \log n \ \text{whp},$$

see [20], [3]. Also, for any constant $c > 0$,

$$\mathbb{P}(\Delta(T_n) \geq c n / \log n) = e^{-(c+o(1))n}.$$

Both these results follow easily from considering Prüfer coding.

The following theorem says roughly that block degrees are no larger than those for a random tree $T_n$. In particular, if we sample $R_n$ uniformly from the connected graphs in a block-stable class, then the maximum block degree $\tilde{\Delta}(R_n)$ is stochastically at most $\Delta(T_n)$, and so whp it is no more than about $\log n / \log \log n$; and indeed we can improve the bound if there are few blocks.

**Theorem 1.1.** Let $\mathcal{A}$ be a weakly block-stable class of graphs and let $\mathcal{C}$ be the class of connected graphs in $\mathcal{A}$.

(a) For $R_n \in \mathcal{C}$, the list of block degrees $(\tilde{d}_{R_n}(v) : v \in [n])$ is stochastically at most $(d_{T_n}(v) : v \in [n])$, where $T_n \in \mathcal{U}$ is a uniformly random tree on $[n]$; and in particular the maximum block degree $\tilde{\Delta}(R_n)$ is stochastically at most $\Delta(T_n)$.

(b) For $R_n \in \mathcal{A}$,

$$\tilde{\Delta}(R_n) \leq (1 + \epsilon(n)) \log n / \log \log n \ \text{whp}$$

where $\epsilon(n) = o(1)$, and indeed we may take $\epsilon(n) = 2 \log \log n / \log \log n$, (whatever $\mathcal{A}$ is); and for any constant $c > 0$

$$\mathbb{P}(\tilde{\Delta}(R_n) \geq cn / \log n) \leq e^{-(1-o(1))cn},$$
where $\eta(n) = o(1)$, and indeed we may take $\eta(n) = 2\log\log n/\log n$. Further, if the number of blocks in graphs in $A_n$ is at most $k = k(n)$ where $k \to \infty$ as $n \to \infty$, then

\begin{equation}
\Delta(R_n) \leq (1 + \epsilon(k)) \log k/\log \log k \text{ whp}
\end{equation}

where the function $\epsilon$ is as above.

For $R_n \in_u A$ as in part (b), there is no detailed result on stochastic dominance by a tree like that for $R_n \in_u C$ in part (a) (see the comment following Lemma 3.2 below). Of course the inequality (5) implies the earlier inequality (3) since there can be at most $n$ blocks. Theorem 1.1 will be deduced from more precise non-asymptotic results, Lemmas 3.1 and 3.2 below.

Finally here, consider the class $\text{Ex}(C_4)$ of graphs with no minor the cycle $C_4$ on 4 vertices. For graphs in this class, each block is a vertex or an edge or a triangle. Thus, for $R_n \in_u \text{Ex}(C_4)$, by (3) we have

\[ \Delta(R_n) \leq (2 + o(1)) \log n/\log \log n \text{ whp}, \]

as in [11] Lemma 10. This inequality is tight, and we have

\[ \Delta(R_n) \log \log n/\log n \to 2 \text{ in probability as } n \to \infty. \]

For the lower bound, see Theorem 4.1 of [18] (suitably amended) or Theorem 3 part 2 of [11].

1.2. Block length of paths. Now we consider paths, and see that graphs in $A$ are unlikely to contain any path which passes through many blocks (that is, any path which has edges in many different blocks). The diameter of a graph is the maximum distance between any two vertices in the same component.

For $T_n \in_u T$, with probability near 1 the diameter of $T_n$ is of order $\sqrt{n}$ [25]: more exactly, for any $\epsilon > 0$ there are constants $0 < c_1 < c_2$ such that with probability at least $1 - \epsilon$ the diameter is between $c_1 \sqrt{n}$ and $c_2 \sqrt{n}$. See [9] for a precise result on the maximum length of a path from a root vertex to another vertex (see also Theorem 4.8 of [5]). For contrast, it was shown in [8] that whp the diameter of a random planar graph $R_n$ is $n^{1+o(1)}$, see also [4] for more precise information. Also, observe that for $n \geq 2$ the probability that $T_n$ has diameter $n-1$ (that is, $T_n$ is a path) is $n!/(2n^{n-2}) = e^{-n+O(\log n)}$.

The following theorem shows in particular that, if we sample $R_n$ uniformly from the connected graphs in a block-stable class, then whp the block tree $\text{BT}(R_n)$ has diameter at most $5\sqrt{n} \log n$. We conjecture that the extra factor $\sqrt{\log n}$ (compared with the random tree $T_n$) could be replaced by any function tending to $\infty$. 
Theorem 1.2. Let $\mathcal{A}$ be a weakly block-stable class of graphs, and let $R_n \in \mathcal{A}$. Then whp the block forest $BF(R_n)$ has diameter at most $5\sqrt{n \log n}$; and for each $\epsilon > 0$ the probability that $BF(R_n)$ has diameter at least $\epsilon n$ is $e^{-\Omega(n)}$, where the function $\Omega(n)$ does not depend on the class $\mathcal{A}$. Further, if the number of blocks in graphs in $\mathcal{A}_n$ is at most $k = k(n)$ where $k \to \infty$ as $n \to \infty$, then whp the block forest $BF(R_n)$ has diameter at most $5\sqrt{k \log k}$.

Observe that if the blocks in the graphs considered are of bounded size (for example in the block class $\text{Ex}(C_4)$ of graphs with no minor $C_4$ each block has at most 3 vertices) then these results transfer easily from block trees or forests to $R_n$ itself.

To prove Theorem 1.2 we give a precise non-asymptotic lemma, Lemma 4.4 below, from which the theorem will follow easily.

2. Trees and coding following Prüfer

Let $\mathcal{C}$ be the class of connected graphs in a weakly block-stable class; or equivalently, let $\mathcal{C}$ be a weakly block-stable class of connected graphs. With respect to the equivalence relation we introduced earlier, $\mathcal{C}_n$ is naturally partitioned into equivalence classes $[G]$. We shall show that $\mathcal{C}_n$ may be partitioned more finely into parts $\mathcal{G}$, so that if $G \in \mathcal{C}$ has $k$ blocks and $\mathcal{G}$ is a part contained in $[G]$, then there is a bijection between $G$ and $[n]^{k-1}$, similar to that in Prüfer coding, see for example the book by van Lint and Wilson [26]. The encoding that we use here is essentially the same as that introduced by Kajimoto [15].

Given a connected graph $G$ on vertex set $V = [n]$ with $k$ blocks, we will ‘explode’ $G$ into a tree-like graph $\widetilde{G}$ rooted at vertex $n$. The graph $\widetilde{G}$ will contain vertex-disjoint copies of the blocks of $G$ (plus one additional root vertex), joined together in a tree structure that indicates how the blocks are joined together in $G$. See the graphs $H$ and $\tilde{H}$ in Figure 1.

Informally, $\widetilde{G}$ is constructed as follows: we begin by taking vertex-disjoint copies $B_1, \ldots, B_k$ of the $k$ blocks of $G$ and add an extra block containing the single vertex $n$. Thus we get one copy of a vertex for each block it belongs to, and an additional copy of $n$ (so a vertex appears more than once if and only if it is a cutvertex or $n$). For each vertex $j$ that occurs more than once, if $B$ is the block containing $j$ which is nearest to vertex $n$ in $G$, we give new labels to the copies of $j$ other than the copy in $B$, and refer to these as ‘ghost’ vertices; and finally, we join vertex $j$ to its corresponding ghost vertices.

More formally, we apply the following procedure:
• For each block $B$ of $G$, let $v_B$ be the vertex $n$ if $n$ is in $B$, and otherwise let $v_B$ be the cut-vertex in $B$ which separates $B$ from vertex $n$. Let $Q_B = V(B) \setminus \{v_B\}$. Note that every vertex other than $n$ appears in exactly one set $Q_B$, so the $k$ sets $Q_B$ partition $[n-1]$.

• Relabel the blocks as $B_1, \ldots, B_k$ in some canonical order (say in increasing order of the largest vertex in $Q_B$). For each $i = 1, \ldots, k$, denote $Q_{B_i}$ by $Q_i$ and $v_{B_i}$ by $v_i$ (the vertices $v_i$ need not be distinct). Additionally, set $Q_{k+1} = \{n\}$.

• For each $i = 1, \ldots, k$, add a new ‘ghost’ vertex $g_i$ to $Q_i$, and add edges from $g_i$ to the neighbours of $v_i$ in $B_i$. We set $P_i = Q_i \cup \{g_i\}$, and set $P_{k+1} = \{n\}$. Thus the $P_i$ partition $V(G) \cup \{g_1, \ldots, g_k\}$ and, for $i = 1, \ldots, k$, $P_i$ induces a copy of $B_i$.

• Finally, we delete all edges between the sets $P_i$, and then add edges $g_iv_i$ for each $i$.

Let $\tilde{G}$ be the resulting graph, with vertex partition $P_G = \{P_1, \ldots, P_{k+1}\}$.

The edges $g_iv_i$ join up the $P_i$ in a tree structure encoding the block structure of $G$. Observe that each edge $g_iv_i$ is a bridge in $\tilde{G}$, and if we contract each of the edges $g_iv_i$ we obtain the original graph $G$. If we start with $\tilde{G}$ and contract each set $P_i$ to a single node $i$ then we form a tree on $[k+1]$, which we denote by $T_G$. Note that if $G$ has a path with edges in $t+1$ distinct blocks then $T_G$ has a path of length $t$ (as edges in distinct blocks correspond to edges in distinct sets $P_i$, which are contracted into distinct vertices of $T_G$).

Let $C$ be a weakly block-stable class of connected graphs. Let $G$ be a (fixed) graph in $C_n$, and suppose that $G$ has $k \geq 2$ blocks. Let $P_G = \{P_1, \ldots, P_{k+1}\}$. Let the explosion neighbourhood $G_G$ of $G$ be the set of all connected graphs $H$ on $[n]$ such that $P_H = P_G$, and the induced graphs $\tilde{H}[P_i] = \tilde{G}[P_i]$ for each $i = 1, \ldots, k$. (Note that the labelled graphs $\tilde{H}[P_i]$ and $\tilde{G}[P_i]$ are identical, not just isomorphic.) Then for each graph $H$ in $G_G$, the blocks of $G$ and $H$ are the same up to isomorphism (although they may be attached to each other differently); and thus $H \sim G$, $H$ is in $C_n$ and $G_G \subseteq [G] \subseteq C_n$. Thus $[G]$ is partitioned into disjoint explosion neighbourhoods. Also, notice that if $H$ is in $G_G$ and the trees $T_H$ and $T_G$ are the same, then the only differences between $\tilde{H}$ and $\tilde{G}$ are the choices of ‘external’ neighbours for the ghost vertices $g_i$. Recall that $v_i$ is the neighbour of $g_i$ outside $P_i$ in $\tilde{G}$. If $v_i$ is in $P_j$ (and so $v_i$ is in $Q_j$) then in $\tilde{H}$ we may have any vertex in $Q_j$ as neighbour of $g_i$. 
Further, suppose that we start from $P_G = \{P_1, \ldots, P_{k+1}\}$, and construct a graph $K$ as follows. For each $i = 1, \ldots, k$ we choose any neighbour $u_i$ for $g_i$ such that

- $u_i \in P_j$ for some $j \neq i$ and $u_i$ is not a ghost vertex (that is, $u_i \not\in Q_j$), and
- the graph obtained from $K$ by contracting each $P_i$ to a single node $i$ is a tree $T$ on $[k+1]$.

Then the graph $H$ obtained from $K$ by contracting each newly added edge $g_iu_i$ is in $G_G$, $K$ is the tree-like graph $\tilde{H}$ corresponding to $H$, and $T$ is the corresponding tree $T_H$.

The distinct parts $G_G$ for $G \in \mathcal{C}_n$ partition $\mathcal{C}_n$, and so it will suffice for us to fix one graph $G \in \mathcal{C}_n$ where $G$ has $k \geq 2$ blocks, and consider the part $G_G$. We shall see that there is a natural bijection between $G_G$ and $[n]^{k-1}$, obtained by a slight extension of Prüfer coding. Recall that the Prüfer coding of a labelled tree $T$ is obtained by repeatedly deleting the leaf with smallest label and recording the label of its neighbour, repeating until two vertices remain; this gives a bijection between trees on $[n]$ and elements of $[n]^{n-2}$. Given a tree $T$ on $[n]$ for some $n \geq 2$ let $t = t(T) \in [n]^{n-2}$ denote its Prüfer codeword; and given $t \in [n]^{n-2}$ let $T = T(t)$ be the corresponding tree.

For a graph $H \in G_G$, let us consider the tree-like graph $\tilde{H}$ and the tree $T_H$ on $[k+1]$. We construct a codeword $x_H = (x_1, \ldots, x_{k-1}) \in [n]^{k-1}$ as follows: if $i$ is the leaf of $T_H$ with smallest label, and $j$ is the neighbour of $i$ in $T_H$, then let $x_1$ be the neighbour of $g_i$ in $P_j$, record $x_1$, and delete vertex $i$; repeat to find $x_2$ (if $k \geq 3$), and continue until two
vertices remain. In the example in Figure 1, \( x_H = (x_1, x_2) = (2, 2) \). Further, let \( f : [n] \to [k + 1] \) be given by setting \( f(i) = j \) if vertex \( i \) is in \( P_j \). If \( x_H = (x_1, \ldots, x_{k-1}) \) then the Prüfer codeword \( t(T_H) \) is \( (f(x_1), \ldots, f(x_{k-1})) \).

Just as Prüfer coding gives a bijection between trees on \([k + 1]\) and vectors in \([k+1]^{k-1}\), so the map \( H \to x_H \) gives a bijection as required, between \( \mathcal{G}_G \) and \([n]^{k-1}\). Also, for each vertex \( j \) of \( H \), the number of blocks of \( H \) containing \( j \) is \( 1 + a(j, x_H) \), where \( a(j, x) \) is the number of appearances of \( j \) in the vector \( x \), that is, the number of co-ordinates of \( x \) which are equal to \( j \).

For each \( j = 1, \ldots, k \), let \( w_j = |f^{-1}(j)| = |Q_j| = |P_j| - 1 \), and let \( w_{k+1} = 1 \) (thus, for \( j = 1, \ldots, k + 1 \), \( w_j \) is the number of choices for the neighbour \( v_i \) of \( g_i \) in \( Q_j \), and \( \sum_j w_j = n \)). Let \( T \) be a tree on \([k + 1]\), with corresponding codeword \( t = (t_1, \ldots, t_{k-1}) \in [k + 1]^{k-1} \). Then by the above the number of graphs \( H \in \mathcal{G}_G \) with \( T_H = T \) is

\[
\prod_{i=1}^{k-1} w_{t_i} = \prod_{j=1}^{k+1} w_j^a(j, t) = \prod_{j=1}^{k+1} w_j^d_{t(j)-1}.
\]

Let \( n \geq 3 \). Consider a connected graph \( G \) with vertex set \( V = [n] \) and with \( k \geq 2 \) blocks, and with corresponding explosion neighbourhood \( \mathcal{G}_G \) as above. Let \( R \in_u \mathcal{G}_G \). Consider the corresponding tree-like graph \( \widetilde{R} \) and tree \( T_R \). We shall identify the distributions of the extended Prüfer codeword \( x_R \in [n]^{k-1} \) corresponding to \( \widetilde{R} \), and of the Prüfer codeword \( t(T_R) \in [k + 1]^{k-1} \) corresponding to \( T_R \).

For \( x_R \) this is easy: we have already noted that there is a bijection between the graphs in \( \mathcal{G}_G \) and the codewords, and so \( x_R \) is uniformly distributed over \([n]^{k-1}\). For \( t(T_R) \), let the random variable \( X \) take values in \([k + 1]\), with \( \mathbb{P}(X = j) = w_j / \sum_{i=1}^{k+1} w_i \), and let \( X = (X_1, \ldots, X_{k-1}) \) where \( X_1, \ldots, X_{k-1} \) are independent, each distributed like \( X \). Then \( X \) and \( t(T_R) \) have the same distribution. For, by (6), given a vector \( t = (t_1, \ldots, t_{k-1}) \in [k + 1]^{k-1} \), both \( \mathbb{P}(X = t) \) and \( \mathbb{P}(t(T_R) = t) \) are proportional to \( \prod_{j=1}^{k+1} w_j^{a(j, t)} \), and they both take the same values \( t \) so the normalising constants must be the same.

3. **Number of blocks containing a vertex**

We begin by showing that, for any weakly block-stable class of connected graphs, the block degree sequence of a random graph in the class is stochastically dominated by the degree sequence of a random tree.
Lemma 3.1. Let $\mathcal{C}$ be a weakly block-stable class of connected graphs, and let $R_n \in u\mathcal{C}$. Then $(\tilde{d}_{R_n}(v) : v \in [n])$ is stochastically at most $(d_{T_n}(v) : v \in [n])$, where $T_n \in_u \mathcal{T}$ is a uniformly random tree on $[n]$.

Indeed, let $G$ be a fixed graph in $\mathcal{C}_n$ with $k$ blocks, let $G^* G$ be the explosion neighbourhood of $G$, and let $R \in_u G^* G$. Then $(\tilde{d}_R(v) : v \in [n])$ is stochastically at most $(d_{T_n}(v) : v \in [n])$. Further,

$$(7) \quad \mathbb{P}(\tilde{\Delta}(R) \geq s + 1) \leq n \left( \frac{ek}{ns} \right)^s \leq k(e/s)^s.$$ 

Proof. It suffices to prove the statements concerning $R \in_u G^* G$. Recall that $x_R \in_u [n]^{k-1}$. The block-degrees $\tilde{d}_R(j)$ of the vertices $j$ of $R$ satisfy

$$(8) \quad (\tilde{d}_R(1), \ldots, \tilde{d}_R(n)) = (a(1, x_R) + 1, \ldots, a(n, x_R) + 1).$$

Now let $T \in_u T_n$. Recall that $t = t(T) \in_u [n]^{n-2}$, and

$$(d_T(1), \ldots, d_T(n)) = (a(1, t) + 1, \ldots, a(n, t) + 1).$$

But $k - 1 \leq n - 2$, and so $(\tilde{d}_R(v) : v \in [n])$ is stochastically at most $(d_{T_n}(v) : v \in [n])$. Also, by (8), for each integer $s > 0$,

$$\mathbb{P}(\tilde{d}_R(1) \geq s + 1) = \mathbb{P}(a(1, x_R) \geq s) = \mathbb{P}(\text{Bin}(k - 1, n^{-1}) \geq s) \leq \binom{k - 1}{s} n^{-s} \leq \left( \frac{ek}{ns} \right)^s.$$ 

Thus for each integer $s \geq 1$

$$\mathbb{P}(\tilde{\Delta}(R) \geq s + 1) \leq n(ek/ns)^s \leq k(e/s)^s.$$ 

Finally, since $(e/x)^x$ is decreasing in $x$ for $x \geq 1$ we may drop the assumption that $s$ is integral, to obtain (7). $\square$

Lemma 3.1 proves part (a) of Theorem 1.1. The next lemma is a more detailed version of part (b) of that theorem, and will quickly yield that result.

Lemma 3.2. Let $\mathcal{A}$ be a weakly block-stable class of graphs. Fix a graph $G \in \mathcal{A}_n$, with a total of $k$ blocks. Let $R_n \in_u \mathcal{A}$. Then, for each real $s \geq 1$ we have

$$(9) \quad \mathbb{P}(\tilde{\Delta}(R_n) \geq s + 1 \mid R_n \in [G]) \leq k(e/s)^s.$$ 

Life would have been tidier if there had been a detailed stochastic dominance result here corresponding to that in Lemma 3.1 involving a random tree - but unfortunately that is not the case. For example, let $\mathcal{A}$ be the class of forests, let $n = 6$, let $G \in \mathcal{A}_6$ have two components both of which are paths of length 2, and let $R_n \in_u [G]$. Let $\mathcal{P}$ be the increasing set in $\{0, 1, \ldots\}^6$ where $x \in \mathcal{P}$ when we can partition the set
of co-ordinates into two 3-sets $I$ and $J$ such that both $\sum_{i \in I} x_i \geq 4$ and $\sum_{j \in J} x_j \geq 4$. Then the probability that the (block) degree sequence of $R_n$ is in $\mathcal{P}$ is 1, but the probability that the degree sequence of $T_n$ is in $\mathcal{P}$ is $< 1$, since for example $T_n$ can be a star. Thus here (with $n = 6$) it is not true that $(d_{R_n}(v) : v \in [n])$ is stochastically at most $(d_{T_n}(v) : v \in [n])$.

However, we can use the stochastic dominance in Lemma 3.1 ‘component by component’.

**Proof of Lemma 3.2.** Let $\mathcal{C}$ be the class of connected graphs in $\mathcal{A}$. Suppose that the graph $G \in \mathcal{A}_n$ has components $G_1, \ldots, G_j$ for some $j \geq 1$. For each $i = 1, \ldots, j$ let $W_i = V(G_i)$. Suppose that $G_i$ has $k_i$ blocks, and observe that $\sum_{j} k_i = k$. Let $\mathcal{E}$ be the set of graphs on $[n]$ with no edges between distinct sets $W_i$ and $W_i'$. Then for a graph $H$ on $[n]$, $H \in [G]$ iff $H \in \mathcal{E}$ and $H[W_i] \in [G_i]$ for each $i$.

For each $i$, let the random graph $S_i$ be uniformly distributed over the graphs in $[G_i]$. Then

$$\mathbb{P}(\tilde{\Delta}(R_n[W_i]) \geq s+1 \mid R_n \in [G])$$

$$= \mathbb{P}(\tilde{\Delta}(S_i) \geq s+1) \leq k_i(e/s)^s$$

by Lemma 3.1. Hence, by the union bound

$$\mathbb{P}(\tilde{\Delta}(R_n) \geq s+1 \mid R_n \in [G]) \leq \sum_{i=1}^{j} \mathbb{P}(\tilde{\Delta}(R_n[W_i]) \geq s+1 \mid R_n \in [G])$$

$$\leq \sum_{i=1}^{j} k_i(e/s)^s = k(e/s)^s$$

as required.

We may now complete the proof of Theorem 1.1.

**Proof of Theorem 1.1.** It remains to prove part (b) of the theorem. Let $G \in \mathcal{A}_n$ have $k$ blocks. It suffices to show that (5) (and thus (3)) and (4) hold for $R_n$ conditioned on $R_n \in [G]$. To see this, we use the last lemma: set $s + 1 = (1 + \epsilon) \log k / \log \log k$ to deduce (5), and $s + 1 = cn/ \log n$ to deduce (4).

4. **Path lengths**

Let $Q(t)$ denote the class of graphs $G$ which have a path containing edges in at least $t$ different blocks. Thus a forest is in $Q(t)$ if and only if it has a path of length at least $t$. We first consider connected graphs.
Lemma 4.1. Let \( \mathcal{C} \) be a weakly block-stable class of connected graphs, let \( G \) be a fixed graph in \( \mathcal{C}_n \) with \( k \) components, let \( \mathcal{G}_G \) be the explosion neighbourhood of \( G \), and let \( R \in u \mathcal{G}_G \). Then for each \( t \geq 0 \),

\[
\mathbb{P}(R \in Q(t+1)) \leq 2k^2e^{-\frac{t^2}{2(k+1)}}.
\]

In order to prove Lemma 4.1 we need two lemmas: the first preliminary lemma may well be known but we give a proof for completeness.

Lemma 4.2. Let \( 2 \leq j \leq n \), and let \( X_1, \ldots, X_j \) be iid random variables taking values in \( [n] \). Then (a) the probability that \( X_1 \) is not repeated is at most \((1 - 1/n)^{j-1}\); and (b) the probability that \( X_1, \ldots, X_j \) are all distinct is at most \((n/j)n^3\).

Proof. Denote \( \mathbb{P}(X_1 = i) \) by \( p_i \) for \( i = 1, \ldots, n \), and set \( p = (p_1, \ldots, p_n) \).

(a) The probability that \( X_1 \) is not repeated is \( g(p) = \sum_{i=1}^n p_i(1 - p_i)^{j-1} \). Suppose first that \( j = 2 \). Then \( g(p) = 1 - \sum_{i=1}^n p_i^2 \leq 1 - 1/n \) since, as is well known, \( \sum_{i=1}^n p_i^2 \) is minimised when each \( p_i = 1/n \).

Now suppose that \( j \geq 3 \). Let \( f(x) = x(1-x)^{j-1} \) for \( 0 \leq x \leq 1 \). Then \( g(p) = \sum_{i=1}^n f(p_i) \). Let \( m \) be the maximum value of this quantity, achieved at \( q = (q_1, \ldots, q_n) \). Now \( f'(x) = (1-x)^{j-2}(1-jx) \), which is \( > 0 \) for \( 0 < x < 1/j \), \( = 0 \) at \( x = 1/j \) and \( < 0 \) for \( 1/j < x < 1 \). Also \( f''(x) = (j-1)(1-x)^{j-3}(2-jx) \), which is \( > 0 \) for \( 0 < x < 2/j \).

Clearly each \( q_i \in [0,1] \). If \( q_i > 1/j \) for some \( i \) then there is \( k \) with \( q_k < 1/j \) (as \( \sum_k q_k = 1 \)); increasing \( q_i \) and decreasing \( q_k \) slightly would then increase \( g(q) \). We must therefore have \( \max_i q_i \leq 1/j \), and so by (strict) convexity \( g(q) \) is (uniquely) maximized when all the \( q_i \) take the same value, which must be \( 1/n \). Hence \( m = (1 - 1/n)^{j-1} \), which completes the proof of (a).

(b) Consider any positive integer \( n \). The result is trivially true for \( j = 1 \). Let \( 2 \leq j \leq n \) and suppose that it holds for \( j-1 \). Let \( A_i \) be the event that none of \( X_2, \ldots, X_j \) are equal to \( i \). Then by conditioning on \( X_1 \) and using the induction hypothesis, we find

\[
\mathbb{P}(X_1, \ldots, X_j \text{ distinct}) = \sum_{i=1}^n \mathbb{P}(X_1 = i, A_i) \mathbb{P}(X_2, \ldots, X_j \text{ distinct}|A_i)
\]

\[
\leq \sum_{i=1}^n \mathbb{P}(X_1 = i, A_i) \frac{(n-1)_{j-1}}{(n-1)^{j-1}}
\]

\[
= \mathbb{P}(X_1 \text{ not repeated}) \frac{(n-1)_{j-1}}{(n-1)^{j-1}}
\]

\[
\leq \left( \frac{n-1}{n} \right)^{j-1} \frac{(n-1)_{j-1}}{(n-1)^{j-1}} = \frac{(n)_j}{n^j}
\]
Lemma 4.3. Let \( m \geq 3 \) and let \( w_1, \ldots, w_m > 0 \). Let the random variable \( X \) take values in \([m]\), with \( \mathbb{P}(X = j) = w_j / \sum_{i=1}^{m} w_i \). Let \( X = (X_1, \ldots, X_{m-2}) \) where \( X_1, \ldots, X_{m-2} \) are independent, each distributed like \( X \). Consider the random tree \( T(X) \) on \([m]\). For each integer \( t \geq 1 \), the expected number of paths of length at least \( t+1 \) is at most

\[
\left( \frac{m}{2} \right) e^{-\frac{(t^2)}{2m}} \leq 2(m - 1)^2 e^{-\frac{t^2}{2m}}.
\]

Before we prove this lemma, let us note that it will yield Lemma 4.1. To see this, let \( G \) have \( k \) blocks, and set \( m = k + 1 \) in the last lemma. Now recall from the end of Section 2 that, for a suitable choice of \( w_1, \ldots, w_m \), \( T_R \) has the same distribution as \( T(X) \). But if \( H \in Q(t+2) \) then \( T_H \) has a path of length \( t + 1 \).

Proof. We first consider the distance in \( T(x) \) between vertices \( m-1 \) and \( m \) using Prüfer coding, and then extend to all pairs of vertices. Given a tree \( T \in T_m \) and distinct vertices \( i, j \in [m] \), denote the distance between \( i \) and \( j \) in \( T \) by \( \text{dist}(i,j;T) \). We claim that

\[
\mathbb{P}(\text{dist}(m-1,m;T(X)) \geq t + 1) \leq e^{-\frac{(t)}{m}}.
\]

(11)

To prove this, consider any vector \( x \in [m]^{m-2} \). If the path between vertices \( m - 1 \) and \( m \) in \( T(x) \) has length at least \( t + 1 \) then the last \( t \) co-ordinates of \( x \) are distinct (this follows by considering the algorithm for Prüfer applied to a tree \( T \) on \([m]\): running the algorithm for as long as it removes leaves with labels from \([m-2]\), we are left with the path from \( m - 1 \) to \( m \) in \( T \); the remaining \( t \) steps of the algorithm run through the path, starting from the \( m - 1 \) end, and record the internal vertices of the path in order). Hence the probability that the path between vertices \( m - 1 \) and \( m \) in \( T(X) \) has length at least \( t + 1 \) is at most the probability that the last \( t \) of the \( X_i \) are distinct, which is at most \( (m)_t/m^t \) by Lemma 4.2. But

\[
(m)_t/m^t = \prod_{i=0}^{t-1} (1 - i/m) \leq \exp \left( - \sum_{i=0}^{t-1} i/m \right) = e^{-\frac{(t)}{m}}.
\]

which establishes the claim (11).

There is sufficient symmetry for us to be able to use (11) to show that the same bound holds for the distance between an arbitrary pair of vertices. We spell this out now.

Let \( \pi \) be a permutation of \([m]\). We denote the image of an element \( i \in [m] \) by \( i^\pi \). Given a vector \( z = (z_1, \ldots, z_m) \) let \( z^\pi \) denote the
permutated vector with \((z^\pi)_i = z_{\pi(i)}\). Given a tree \(T \in \mathcal{T}_m\), let \(T^\pi\) denote the tree in \(\mathcal{T}_m\) with an edge \(i^\pi j^\pi\) for each edge \(ij\) in \(T\), so that \(\pi\) is an isomorphism from \(T\) to \(T^\pi\). Also, given a tree \(T \in \mathcal{T}_m\), let \(d(T)\) be the degree sequence \((d_T(1), \ldots, d_T(m))\) of \(T\). Thus \(d(T^\pi) = d(T)^{\pi^{-1}}\).

Consider distinct vertices \(i\) and \(j\) in \([m]\). Let \(\pi\) be a (fixed) permutation of \([m]\) with \(i^\pi = m - 1\) and \(j^\pi = m\). Let \(z = (z_1, \ldots, z_m)\) be a vector of positive integers with \(\sum_i z_i = 2m - 2\). The permutation \(\pi\) yields a bijection \(\phi\) from \(\{T \in \mathcal{T}_m : d(T) = z\}\) to \(\{T \in \mathcal{T}_m : d(T) = z^{\pi^{-1}}\}\) which takes \(\{T \in \mathcal{T}_m : d(T) = z, \text{dist}(i, j; T) = s\}\) to \(\{T \in \mathcal{T}_m : d(T) = z^{\pi^{-1}}, \text{dist}(m-1, m; T) = s\}\). Also, \(d(T(x)) = z\) iff the number \(a(v,x)\) of appearances of \(v\) in \(x\) is \(z_v - 1\) for each \(v \in [m]\), so conditional on \(d(T(X)) = z\) all trees \(T\) with \(d(T) = z\) are equally likely, with probability \((\{T \in \mathcal{T}_m : d(T) = z\})^{-1}\). Let \(Y_i = (X_i)^\pi\) for each \(i\), and let \(Y = (Y_1, \ldots, Y_m)\). Then

\[
\mathbb{P}(\text{dist}(i, j; T(X)) = s \mid d(T(X)) = z) = \frac{\mathbb{P}(\text{dist}(i, j; T(X)) = s | d(T) = z)}{\mathbb{P}(d(T) = z)} = \frac{|\{T \in \mathcal{T}_m : \text{dist}(i, j; T) = s\}|}{|\{T \in \mathcal{T}_m : d(T) = z^{\pi^{-1}}\}|}.
\]

Hence, summing over the possible degree sequences \(z\),

\[
\mathbb{P}(\text{dist}(i, j; T(X)) = s) = \sum_z \mathbb{P}(\text{dist}(i, j; T(X)) = s | d(T(X)) = z) \mathbb{P}(d(T(X)) = z)
= \sum_z \mathbb{P}(\text{dist}(m-1, m; T(Y)) = s | d(T(Y)) = z^{\pi^{-1}}) \mathbb{P}(d(T(Y)) = z^{\pi^{-1}})
= \mathbb{P}(\text{dist}(m-1, m; T(Y)) = s).
\]

Now, since \(Y\) has a distribution of the same form as \(X\), we may apply the claim (11) to see that

\[
\mathbb{P}(\text{dist}(i, j; T(X)) \geq t + 1) = \mathbb{P}(\text{dist}(m-1, m; T(Y)) \geq t + 1) \leq e^{-(t+1)/m}.
\]

It follows that the expected number of paths in \(T(X)\) of length at least \(t + 1\) is at most

\[
\left(\frac{m}{2}\right) e^{-(t+1)/m} \leq (m-1)^2 e^{-\frac{t(t-1)}{2m}} \leq 2(m-1)^2 e^{-\frac{t^2}{2m}},
\]

since we may assume that \(t \leq m\) and then \(e^{\frac{t}{2m}} \leq e^{\frac{1}{2}} < 2\).
At this point we have completed the proof of Lemma 4.1. The next lemma is a more detailed version of Theorem 1.2, and will quickly yield that result. It may be deduced from Lemma 4.1 just as Lemma 3.2 was deduced from Lemma 3.1.

**Lemma 4.4.** Let $\mathcal{A}$ be a weakly block-stable class of graphs. Fix a graph $G \in \mathcal{A}_n$, with a total of $k$ blocks. Let $R_n \in \mathcal{A}$. Then for each $t \geq 0$,

$$
\mathbb{P}(R_n \in Q(t+2) \mid R_n \in [G]) \leq 2k^2e^{-\frac{t^2}{(k+1)^2}}.
$$

We may now complete the proof of Theorem 1.2, much as we did for Theorem 1.1.

**Proof of Theorem 1.2.** Let $G \in \mathcal{A}_n$ have $k$ blocks. It suffices to prove the theorem for $R_n$ conditioned on $R_n \in [G]$. If BF($H$) has a path of length $t$ then $H \in Q(t/2)$. Thus by the last lemma

$$
\mathbb{P}(\text{BF}(R_n) \text{ has diameter } \geq a((k+1)\log k)^{1/2} + 4 \mid R_n \in [G])
\leq \mathbb{P}(R_n \in Q((a/2)((k+1)\log k)^{1/2} + 2) \mid R_n \in [G])
\leq 2k^2e^{-(a^2/8)\log k} = o(1) \text{ if } a > 4.
$$

Further, since $k + 1 \leq n$,

$$
\mathbb{P}(\text{BF}(R_n) \text{ has diameter } \geq \epsilon n \mid R_n \in [G])
\leq \mathbb{P}(R_n \in Q(\frac{\epsilon n}{3} + 2) \mid R_n \in [G])
\leq 2n^2e^{-\frac{n^2}{18}}
$$

for $n \geq 12/\epsilon$ (so that $2(\frac{\epsilon n}{3} + 2) \leq \epsilon n$). \qed

5. **Concluding remarks**

We have seen that for a random graph $R_n$ from a block-stable class, (or from the connected graphs in such a class), the maximum number of blocks containing a vertex is roughly no more than for a random tree $T_n$, and the maximum number of blocks through which a path may pass is at most a factor $O(\sqrt{\log n})$ times the maximum length of a path in $T_n$.

Let us briefly consider connectedness. A minor-closed class is block-stable if and only if it is addable; that is, each excluded minor is 2-connected, see [17]. Indeed, any block-stable class $\mathcal{A}$ containing the single edge $K_2$ is bridge-addable, and so by [19] the probability that $R_n \in \mathcal{A}$ is connected is at least $1/e$, and indeed $\lim \inf \mathbb{P}(R_n \text{ is connected } \geq$
see [1, 16] and see also [24] for a recent more general result. However, consider the block-stable class \( \mathcal{A} \) in which the only allowed block is the triangle \( C_3 \): the set \( \mathcal{A}_n \) is non-empty for each \( n \geq 5 \), but for each even \( n \) each graph in \( \mathcal{A}_n \) is disconnected.

**Acknowledgement** Thanks to Oliver Riordan for useful discussions, and thanks to the referees for detailed reading and helpful comments.

**References**

[1] L. Addario-Berry, C. McDiarmid and B. Reed, Connectivity for bridge-addable monotone graph classes, *Combinatorics, Probability and Computing* 21 (2012) 803 – 815.

[2] N. Bernasconi, K. Panagiotou and A. Steger, The degree sequence of random graphs from subcritical classes, *Combinatorics, Probability and Computing* 18 (2009) 647 – 681.

[3] R. Carr, W. Goh and E. Schmutz, The maximum degree in a random tree and related problems, *Random Structures and Algorithms* 5 (1994) 13 – 24.

[4] G. Chapuy, E. Fusy, O. Giménez and M. Noy, The diameter of random planar graphs, *Combinatorics, Probability and Computing* 24 (2015) 145 – 178.

[5] M. Drmota, *Random Trees*, Springer Wien New York, 2009.

[6] M. Drmota, E. Fusy, M. Kang, V. Kraus and J. Rué, Asymptotic study of subcritical graph classes, *SIAM J. Discrete Math.* 25 (2011) 1615 – 1651.

[7] M. Drmota, O. Gimenez, M. Noy, K. Panagiotou and A. Steger, The maximum degree of random planar graphs, *Proc. London Math. Soc.* (3) 109 (2014) 892 – 920.

[8] M. Drmota and M. Noy, Extremal parameters in sub-critical graph classes, Proceedings of Analco 2013, SIAM (Markus Nebel and Wojciech Szpankowski eds.) (2013) 1–7.

[9] P. Flajolet and A. Odlyzko, The average height of binary trees and other simple trees, *J. Comput. System. Sci* 25 (1982) 171 – 213.

[10] N. Fountoulakis and K. Panagiotou, 3-connected cores in random planar graphs, *Combinatorics, Probability and Computing* 20 (2011) 381 - 412.

[11] O. Giménez, D. Mitsche and M. Noy, Maximum degree in minor-closed classes of graphs, *Europ. J. Comb.* 55 (2016) 41 – 61.

[12] O. Giménez and M. Noy, Asymptotic enumeration and limit laws of planar graphs, *J. Amer. Math. Soc.* 22 (2009) 309–329.

[13] O. Giménez and M. Noy, Counting planar graphs and related families of graphs, in *Surveys in Combinatorics 2009*, 169 – 329, Cambridge University Press, Cambridge, 2009.

[14] O. Giménez, M. Noy and J. Rué, Graph classes with given 3-connected components, *Random Structures and Algorithms* 42 (2013) 438 – 479.

[15] H. Kajimoto, An extension of the Prüfer code and assembly of connected graphs from their blocks, *Graphs and Combinatorics* 19 (2003) 231–239.

[16] M. Kang and K. Panagiotou, On the connectivity of random graphs from addable classes, *J. Comb. Theory Ser. B* 103 (2013) 306 – 312.

[17] C. McDiarmid, Random graphs from a minor-closed class. *Combinatorics, Probability and Computing* 18 (2009) 583–599.
[18] C. McDiarmid and B. Reed, On the maximum degree of a random planar graph. *Combinatorics, Probability and Computing* **17** (2008) 591–601.
[19] C. McDiarmid, A. Steger and D. Welsh, Random planar graphs. *J. Combin. Theory Ser. B* **93** (2005) 187–205.
[20] J.W. Moon, On the maximum degree in a random tree, *The Michigan Mathematical Journal* **15** (1968) 429 – 432.
[21] M. Noy, Random planar graphs and beyond, Proceedings of the ICM, Volume IV, 407-430, Seoul 2014.
[22] K. Panagiotou and A. Steger, Maximal biconnected subgraphs of random planar graphs, *ACM Transactions on Algorithms* **6** (2010) art. no. 31.
[23] K. Panagiotou, B. Stufler and K. Weller, Scaling limits of random graphs from subcritical classes, arXiv:1411.1865v2 [math.PR], *Ann. of Prob.*, to appear.
[24] G. Perarnau and G. Schaeffer, Connectivity in bridge-addable graph classes: the McDiarmid-Steger-Welsh conjecture, arXiv: 1504.06344v1, April 2015.
[25] A. Rényi and G. Szekeres, On the height of trees, *J. Austral. Math. Soc* **7** (1967) 497 – 507.
[26] J.H. van Lint and R.M. Wilson, *A Course in Combinatorics*, CUP, 2nd ed. 2001.

1 Department of Statistics 1 South Parks Road Oxford, OX1 3TG, United Kingdom; 2 Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG, UK