
On the tree-depth of random graphs

G. Perarnau¹ and O. Serra¹

Departament de Matemàtica Aplicada IV
Universitat Politècnica de Catalunya
Jordi Girona 1-3, 08034 Barcelona.
`{guillem.perarnau,oserra}@ma4.upc.edu`

Abstract. The *tree-depth* of a graph G is a parameter that plays a crucial role in the theory of bounded expansion classes and has been introduced under numerous names.

We describe the asymptotic behaviour of this parameter in the case of dense and of sparse random graphs. The results also provide analogous descriptions for the tree-width of random graphs.

Key words: tree-depth, tree-width, random graphs, threshold function

1 Introduction

Let $G = (V, E)$ be an undirected simple graph with $|V| = n$ and $|E| = m$

The tree-depth $\text{td}(G)$ of a graph G is a measure introduced by Nesetril and Ossona de Mendez [10] in the context of bounded expansion classes (see Section 2 for a precise definition). The notion of the tree-depth is closely connected to the tree-width. The tree-width of a graph tells us how similar is G to a tree, while the tree-depth takes also into account the height of the tree.

Bounded expansion classes are defined in terms of shallow minors and its connection to tree-depth is highlighted by the following deep result. The k -th chromatic number of a graph $\chi_k(G)$ is defined as the minimum number of colors needed to color a graph in such a way that the subgraphs H induced by any i classes of colors, $i \leq k$, satisfy $\text{td}(H) \leq i$. Thus $\chi_1(G)$ is the ordinary chromatic number and $\chi_2(G)$ is the so-called star chromatic number. The main theorem in this context states that a class of graphs \mathcal{C} has bounded expansion if and only if $\limsup_{G \in \mathcal{C}} \chi_k(G) < \infty$ for any fixed $k > 0$. This is a clear motivation to study tree-depth.

This parameter has been introduced under numerous names in the literature. It is equivalent to rank function [12], vertex ranking number (or ordered coloring) [3] and upper chromatic number [10].

The following inequalities relate the tree-width and tree-depth of a graph:

$$\text{tw}(G) \leq \text{td}(G) \leq \text{tw}(G)(\log_2 n + 1) \quad (1)$$

Note that there are graphs that have bounded tree-width but unbounded tree-depth, for example trees.

To understand this new parameter, it is useful to know about its behaviour in certain classes of graphs. The main goal of this paper is to analyze how does it behave on random graphs.

We consider the ErdHos-Rényi model $G(n, p)$ for random graph. A *random graph* $G \in G(n, p)$ has n vertices and every pair of vertices is chosen independently to be an edge with probability p .

If \mathcal{P} is a property that a graph can have, we will say that this property holds *asymptotically almost sure (a.s.s.)* for random graphs $G \in G(n, p)$, if

$$\lim_{n \rightarrow \infty} \Pr(G \text{ has } \mathcal{P}) = 1$$

Throughout the paper, all the results and statements concerning random graphs must be understood in the asymptotically almost sure sense. We will occasionally make use of the $G(n, m)$ model of random graphs, where a graph with n vertices and m edges is chosen with the uniform distribution. As it is well-known the two models are closely connected and a.s.s. statements are usually transferred from one model to the other one.

Our first result states the value of tree-depth for dense random graphs.

Theorem 1. *Let $G \in G(n, p)$ be a random graph with $p = \frac{1}{o(n)}$, then G satisfy a.s.s.*

$$\text{td}(G) = n - o(n)$$

This theorem says that when G has super-linear number of edges its tree-depth attains a really large value. Actually our proof of Theorem 1 provides the same result for tree-width. To our knowledge, the tree-width of a dense random graph has not been studied until now.

But, what happens if the number of edges is linear? This case, the sparse case, is solved by the following theorem,

Theorem 2. *Let $G \in G(n, p)$ be a random graph with $p = \frac{c}{n}$, with $c > 0$,*

- (1) *if $c < 1$, then a.s.s. $\text{td}(G) = \Theta(\log \log n)$*
- (2) *if $c = 1$, then a.s.s. $\text{td}(G) = \Theta(\log n)$*
- (3) *if $c > 1$, then a.s.s. $\text{td}(G) = \Theta(n)$*

This last theorem is closely related with a conjecture of Klops announced in [7] on the linear behaviour of tree-width for random graphs with $c > 1$. This conjecture has been recently proved by Lee, Lee and Oum [8]. Here we give a proof of Theorem 2.(3) which, in view of inequality (1), also provides a simpler proof of Klops conjecture. Our proof uses, as the one in [8], the same

essential result of Benjamini, Kozma and Wormald [1] on the existence of an expander of linear size in a sparse random graph for $c > 1$.

The paper is organized as follows. In Section 2, we define the notion of tree-depth and give some useful results concerning this parameter. Section 3 contains the proof of Theorem 1, which uses the relation connecting tree-width with balanced partitions. Finally Theorem 2 will be proved in Section 4. For $c < 1$ the result follows from the fact that the random graph is a collection of trees and unicyclic graphs of logarithmic size, which gives the upper bound, and there is one of these components with large diameter with respect to its size, providing the lower bound. For $c = 1$ we show that the giant component in the random graph has just a constant number of additional edges exceeding the size of a tree, which gives the upper bound, and rely on a result of Nachmias and Peres [9] on the concentration of the diameter of the giant component to obtain the lower bound. Finally, as we have already mentioned, for $c > 1$ the result follows readily from the existence of an expander of linear size in a sparse random graph for $c > 1$, a fact proved in Benjamini, Kozma and Wormald [1].

2 On tree-depth

Let T be a rooted tree. The *closure* of T is the graph that has the same set of vertices and an edge between every pair of vertices such that one is an ancestor of the other in the rooted tree. Consider a *rooted forest* as the disjoint union of rooted trees with height, the maximum of the heights of the trees. The *tree-depth* of a graph G is defined to be the minimum height of a rooted forest, whose closure contains G as a subgraph. Some examples are shown in Figure 1.

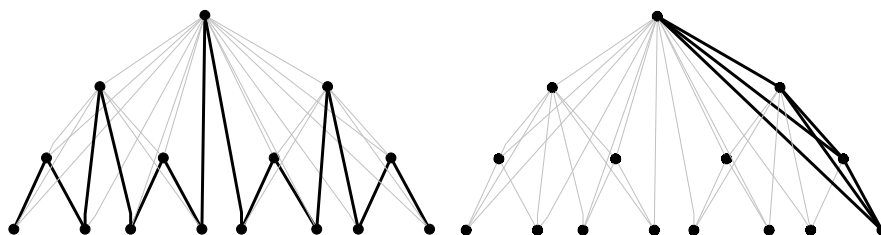


Fig. 1. The path of length 15 and the complete graph K_4 have tree-depth 4

Note that the following inequality holds,

$$\text{td}(G \setminus v) \geq \text{td}(G) - 1 \quad (2)$$

The former observation implies directly that $\text{td}(G) \leq n$. It follows from the definition that if G has connected components C_1, \dots, C_s , then,

$$\text{td}(G) = \max_{0 < i \leq s} \text{td}(C_i) \quad (3)$$

It is easy to check that, for a tree T ,

$$\text{td}(T) \leq \log_2 n + 1 \quad (4)$$

and, if P_n is a path with n vertices, then

$$\text{td}(P_n) = \log_2 n + 1 \quad (5)$$

It is well known that the class of graphs with bounded tree-width at most k is closed under taking minors. By (2), the same is true for the class of graphs with tree-depth at most k . Given a class of graphs \mathcal{C} , its tree-width is bounded if and only if all the graphs in the class exclude a certain grid as a minor. This role in tree-depth is played by paths. As every minor of a path is a path, it is natural to state the following proposition in terms of subgraphs:

Proposition 1 ([11]). *\mathcal{C} has bounded tree-depth if and only if every $G \in \mathcal{C}$ excludes a certain path as a subgraph*

3 Tree-depth for dense random graphs

This section is devoted to prove Theorem 1. We first recall the definition of the tree-width of a graph and of a balanced partition.

A k -tree can be defined in a recursive way,

- $(k + 1)$ -clique is a k -tree.
- a k -tree can be obtained by adding a new vertex to another k -tree and linking it with all the vertices of a certain k -clique.

A *partial k -tree* is a subgraph of a k -tree. The *tree-width* of a graph G is the minimum k such that G is a partial k -tree.

A partition (A, S, B) of the vertices is a *balanced k -partition* if the following three conditions are satisfied:

1. $|S| = k + 1$
2. $\frac{1}{3}(n - k - 1) \leq |A|, |B| \leq \frac{1}{2}(n - k - 1)$
3. S separates A and B .

The following result connecting balanced partitions and tree-width is due to Kloks [7].

Theorem 3 ([7]). *Let G be a graph with n vertices and $\text{tw}(G) \leq k$ such that $n \geq k - 4$. Then G has a balanced k -partition.*

Therefore, the non existence of balanced partitions allows as to lower bound the tree-width, and, by inequality (1), this will also provide lower bounds for the tree-depth.

Proof of Theorem 1. Using the relation connecting tree-width with balanced partitions (Theorem 3), the idea is to prove that G does not contain a balanced separator of size less than $n - o(n)$.

Fix $\beta < 1$. Suppose that we are searching for a balanced separator S of size $k \leq \beta n$. This set separates the graph into two subsets A and B . As S is balanced, we can assume that $|A| \geq |B| \geq \frac{1-\beta}{3}n$.

The probability that a given set $S \subset V$ separates the graph is,

$$\Pr(S \text{ balanced sep. } G) = (1 - p(n))^{|A||B|} \leq (1 - p(n))^{\frac{(1-\beta)(2-\beta)}{9}n^2}. \quad (6)$$

For the last inequality we consider the worst case where $|A| = ((2 - \beta)/3)n$ and $|B| = ((1 - \beta)/3)n$.

We can bound the number of possible balanced partitions with separators S of size at most βn by the number of partitions in three sets, which is obviously less than 3^n .

$$\begin{aligned} \Pr(\exists S \text{ balanced sep. } G) &= \Pr\left(\bigcup_S \{S \text{ is a balanced sep. } G\}\right) \\ &\leq \sum_S \Pr(S \text{ is a balanced sep. } G) \\ &\leq \sum (1 - p(n))^{\frac{(1-\beta)(2-\beta)}{9}n^2} \\ &\leq 3^n (1 - p(n))^{\frac{(1-\beta)(2-\beta)}{9}n^2} \\ &\leq 3^{n + \log(1-p(n))\frac{(1-\beta)(2-\beta)}{9}n^2} \end{aligned}$$

So, it remains to prove that the exponent tends to $-\infty$. As $p(n) < 1$, we can use the inequality of $\log(1 - x) \leq -x$.

$$\begin{aligned} n + \log(1 - p(n))\Theta(n^2) &\leq n - p(n)\Theta(n^2) \\ &= n - \frac{\Theta(n^2)}{o(n)} \rightarrow -\infty \quad (n \rightarrow \infty). \end{aligned}$$

Hence $\Pr(\exists S \text{ balanced sep. } G) \rightarrow 0$.

Since there is no set of this size separating G , we have $\text{td}(G) \geq \text{tw}(G) > \beta n$. Since the above inequality is valid for all $\beta < 1$, we have $\text{td}(G) \geq n - o(n)$. Observe that, $\text{td}(G) \leq n$. \square

4 Tree-depth for sparse random graphs

In this section Theorem 2 will be proved.

4.1 $c < 1$

Let $G \in G(n, p = c/n)$ with $0 < c < 1$. Our objective is to show that $\text{td}(G) = \Theta(\log \log n)$.

First we will prove the upper bound.

A *unicyclic* graph is a connected graph that has the same number of vertices than edges. Note that such a graph consists of a cycle, with some attached trees to its vertices. A *pseudotree* is a graph that is either a tree or a unicyclic graph. A *pseudoforest* is a graph composed by different connected components that are pseudotrees.

Lemma 1. *If G is a pseudoforest, then $\text{td}(G) \leq \log n_c + 2$, where $n_c = \max\{|C| : C \text{ connected component}\}$.*

Proof. Take the largest connected component of G . As tree-depth is an increasing parameter with the subgraph partial ordering, we can assume that G is a unicyclic graph. Let C the cycle and $x \in V(C)$, then $T = G \setminus \{x\}$ is a tree.

By using (2) and (4) it follows that

$$\text{td}(G) \leq \log n_c + 2.$$

□

A famous result of Erdős and Rényi [4] states that, if $0 < c < 1$, then G is a pseudoforest. Bólobas [2, Corollary 5.11] showed that, if $0 < c < 1$, then the size of the largest tree component in the random graph has order $\Theta(\log n)$. Moreover, for any function $\vartheta(n) \rightarrow \infty$, there are at most $\vartheta(n)$ vertices belonging to unicyclic graphs [2, Corollary 5.8]. Taking $\vartheta(n) < \log n$ we can ensure that the largest connected component of G is of size $\Theta(\log n)$. Therefore, by Lemma 1 and $n_c = \Theta(\log n)$,

$$\text{td}(G) = O(\log \log n).$$

The lower bound is more involved.

As we have seen, G is composed of pseudotrees of logarithmic size. Let T be a tree and denote by d its diameter. By (5) and the fact that the tree-depth is monotonically increasing with respect to the subgraph partial ordering,

$$\text{td}(T) \geq \log d.$$

The question now is: which is the diameter of these random trees? Rényi and Szekeres [13] proved that, if H_k is the height of a random labeled rooted tree on k vertices, then

$$\mathbb{E}(H_k) \sim \sqrt{2\pi k}$$

and

$$\mathbf{Var}(H_k) \sim \frac{\pi(\pi-3)}{3}k$$

Since $H_k \leq D_k \leq 2H_k$ and every rooted tree on k vertices has exactly k ways to be rooted, then the diameter D_k satisfies $\mathbb{E}(D_k) = \Theta(\sqrt{k})$ and $\mathbf{Var}(D_k) = \Theta(k)$. Hence our proof for this case will be completed if we show that the random graph contains a tree T with size $\Theta(\log n)$ and diameter $\Theta(\log |T|) = \Theta(\log \log n)$.

Recall that every labeled tree on k vertices has the same probability to appear, and the diameter of each individual tree can be arbitrarily small as n tends to infinity. Since the expectation of D_k has the same order of magnitude as the standard deviation, we can not ensure that its value is highly concentrated. Thus we must show that there are sufficiently many large trees in the random graph.

ErdHos and Rényi [4] showed that X_k , the number of trees of order k in $G \in G(n, m(n))$ with $m(n)/n^{\frac{k-2}{k-1}} \rightarrow \infty$, has a normal distribution with expectation and variance M_n , where

$$M_n = n \frac{k^{k-2}}{k!} \left(\frac{2m}{n} \right)^{k-1} \exp \left(-\frac{2km}{n} \right)$$

Moving back to the random graph model $G(n, p)$ with $p = c/n$, and noting that $\mathbb{E}(m) = \frac{cn}{2}$ we get the analogous result.

$$M_n = n \frac{k^{k-2}}{k!} c^{k-1} \exp(-kc)$$

If $k = \log n$ then,

$$M_n = \frac{n^{\log \log n - \alpha}}{c(\log^2 n)(\log n)!}$$

where $\alpha = c - 1 - \log c$.

It is easy to see that for any c , $\log n \leq M_n$. By using Chebyshev's inequality with $\mu = \sigma^2 = M_n$, we can see that, for some $\omega(n) = o(\sqrt{M_n})$ s.t. $\omega(n) \rightarrow \infty$,

$$\Pr(|X_{\log n} - M_n| \geq \omega(n)\sqrt{M_n}) \leq \frac{1}{\omega(n)^2} \rightarrow 0 \quad (n \rightarrow \infty)$$

This ensures that at least $K = M_n - o(M_n) = \Omega(\log n)$ tree components have size $\log n$.

Let $\overline{D}_{\log n}$ be the mean of the diameter among all the components of size $\log n$. Clearly $\mathbb{E}(\overline{D}_{\log n}) = \Theta(\sqrt{\log n})$, but now we have that $\mathbf{Var}(\overline{D}_{\log n}) =$

$O(\log n/K)$, where $K = \Omega(\log n)$, and so $\sigma(\overline{D}_{\log n}) \leq 1$. Hence, by using again Chebyshev inequality on $\overline{D}_{\log n}$, we can ensure that there exists some tree T with diameter $\Omega(\sqrt{\log n})$.

Finally,

$$\text{td}(G) = \Omega(\log \sqrt{\log n}) = \Omega\left(\frac{1}{2} \log \log n\right) = \Omega(\log \log n).$$

4.2 $c = 1$

Now we look at the critical point where $c = 1$. Another famous result of Erdős and Rényi [4] states that, in this case, the random graph has the so-called giant component (GC) of order $\Theta(n^{2/3})$. The graph has at most $n/2 + O(\sqrt{n})$ edges.

The idea is to prove, that the GC is almost a tree. In [6] a useful concept is defined. For $l \geq -1$ an l -component is a connected component with k vertices and $k + l$ edges. For example, (-1) -components are trees and 0-components are unicyclic graphs. A *complex component* is an l -component with $l > 0$. Let $\{\tilde{G}_t\}_0^N$ denote a graph process where edges are added at random on n points and $N = \binom{n}{2}$.

Proposition 2. *For a random graph process $\{\tilde{G}_t\}$ with $t = \frac{n}{2} + O(n^{2/3})$, GC is an l -component with $l = O(1)$.*

Proof. When $t = \frac{n}{2} - s$ with $s \gg n^{2/3}$, asymptotically there are no complex components (Theorem 5.5 [6]). Let C be a component of size $\Theta(n^{2/3})$. In the t -th step, when we add a random edge e_t , the excess of C , can only increase in one of the following situations:

- (A) e_t links two vertices that belong to C .
- (B) e_t links a vertex in C with another vertex in a unicyclic component.

(A) Let Y_t a random variable that is 1 when e_t is in internal in the GC and 0 otherwise. Then asymptotically

$$\Pr(Y_t = 1) = \frac{\binom{n^{2/3}}{2}}{\binom{n}{2}} = O(n^{-2/3})$$

Defining $S_t = \sum_{i=1}^t Y_i$, and as $t = O(n^{2/3})$, $\mathbb{E}(S_t) = O(1)$, implying that $S_t = O(1)$.

(B) First of all, note that if at most $O(n^{2/3})$ vertices belong to unicyclic components, the same argument can be applied. Let U the number of these vertices. If T are the vertices that belong to trees, Theorem 5.7 in [2], implies due the continuity of $T = T(c)$, that

$$\mathbb{E}(T) = n - o(n)$$

But this only imply that U has not linear size. We find the solution in the same chapter. Theorem 5.23 [2] ensures that the U distribution satisfies,

$$\mathbb{E}(U) \sim 0.406n^{2/3} \quad \sigma^2(U) = O(n^{4/3})$$

Which implies by Chebyshev that U has size at most $O(n^{2/3})$, and the proposition follows.

Since, there was no complex component and the increment of the excess is constant in the GC, we can ensure that $l = O(1)$.

□

The GC of G contain k vertices and $k + l$ edges, and if we delete $l = O(1)$ vertices, we will get a tree on $(k - l) = O(n^{2/3})$ vertices. Since the remaining components have negligible order, the tree-depth satisfies

$$\text{td}(G) \leq l + O\left(\log\left(n^{2/3}\right)\right) = O(1) + O\left(\frac{2}{3}\log n\right) = O(\log n)$$

To prove the lower bound, we use the following result which follows from a more general statement due to Nachmias and Peres [9].

Theorem 4 ([9]). *Let C be the largest component of a random graph in $G(n, p)$ with $p = 1/n$. Then, for any $\varepsilon > 0$, there exists $A = A(\varepsilon)$ such that*

$$\Pr(\text{diam}(C) \notin (A^{-1}n^{1/3}, An^{1/3})) < \varepsilon.$$

It follows from the monotonicity of tree-depth (2) and from (5) that a graph with diameter d satisfies $\text{td}(G) = \Omega(\log d)$. Hence, it follows from Theorem 4 that

$$\text{td}(G) = \Omega(\log n^{1/3}) = \Omega(\log n).$$

This concludes the proof of the case $c = 1$.

4.3 $c > 1$

The *Cheeger constant* of a graph G can be defined as:

$$\Phi(G) = \min_{0 < |X| \leq n/2} \frac{E(X, V \setminus X)}{|X|}$$

This coefficient measures the expansion of the graph. A graph G is said to be an α -expander if $|N(X) \setminus X| \geq \alpha|X|$ for every set X of vertices with $|X| \leq |V(G)|/2$, where $N(X)$ denotes the vertex neighborhood of X . Note that, for an α -expander graph, $\Phi(G) \geq \alpha$.

Proposition 3. *Let $\alpha > 0$. Let G be a graph that contains H an α -expander, of size k . Then $\text{tw}(G) = \Omega(k)$.*

Proof. Denote by $\text{tw}(G) = t_G$ and $\text{tw}(H) = t_H$. As tree-width is closed under subgraph relation and $H \subset G$, we know that $t_H \leq t_G$.

By Theorem 3, we know that there is a balanced partition $V(H) = (A, S, B)$, where S is a vertex separator of cardinality $t_H + 1$ and we can assume that $k/2 \geq |A| \geq (k - t_H - 1)/3$. Hence,

$$t_G \geq t_H \geq |S| - 1 \geq \alpha|A| - 1 \geq \alpha \frac{k - t_H - 1}{3} - 1 \geq \alpha \frac{k - t_G - 1}{3} - 1,$$

where we have used the fact that H is an α -expander. Thus, $t_G \geq \frac{\alpha(k-1)-3}{\alpha+3}$, and $\text{tw}(G) = \Omega(k)$. \square

The recent proof of Benjamini, Kozma and Wormald of the value of mixing time of the random walk on the giant component of a random graph $p = c/n$, $c > 1$, relies on the existence of an α -expander of linear size in the giant component. The following result follows from [1, Theorem 4.2] where the fact that the expander has linear size is stated within the proof of that Theorem.

Theorem 5 ([1]). *Let G be a random graph in $G(n, p)$ with $p = c/n$, $c > 1$. There are $\alpha, \delta > 0$ and a subgraph H of G such that H is an α -expander and $|V(H)| = \delta n$.*

The result follows from Proposition 3 and Theorem 5.

Remark 1 (Random regular graphs have linear tree-width). We note that Proposition 3 also shows that random d -regular graphs (d -RRG) have linear tree-depth and tree-width.

For d -regular graphs there is a two-sided equivalence between expanders and a nonzero value of the Cheeger constant. The latter can be lower bounded by $\lambda_2(G)/2$, where $\lambda_2(G)$ is the second largest eigenvalue of the adjacency matrix of the graph. Friedman, Kahn and Szemerédi [5] prove that this eigenvalue in d -regular random graphs is $O(\sqrt{d})$. Therefore it follows from Proposition 3 that $td(G) \geq tw(G) = \theta(n)$.

Acknowledgement. Work partially supported by the Catalan Research Council under grant 2009SGR01387 and the Spanish Council under project MTM2008-06620-C03-01. The first author acknowledges the grant FPI-UPC

References

- [1] Benjamini, I., Kozma, G. and Wormald, N.. The mixing time of the giant component of a random graph. *Preprint, arxiv:math/0610459v1*, 2006.

- [2] Bollobás, Béla. Random graphs. *Academic Press Inc. [Harcourt Brace Jovanovich Publishers]*, 1985.
- [3] Deogun, J. S. and Kloks, T. and Kratsch, D. and Müller, H.. On vertex ranking for permutation and other graphs, *STACS 94 (Caen, 1994), Lecture Notes in Comput. Sci.*, vol. 775, 747–758, Springer, 1994.
- [4] Erdős, P. and Rényi, A.. On the evolution of random graphs. *Bull. Inst. Internat. Statist.*, vol. 38, 343–347, 1961.
- [5] Friedman, J. and Kahn, J. and Szemerédi, E.. On the second eigenvalue of random regular graphs, *STOC '89: Proceedings of the twenty-first annual ACM symposium on Theory of computing*, ACM, 1989.
- [6] Janson, Svante and Luczak, Tomasz and Rucinski, Andrzej. Random graphs. *Wiley-Interscience Series in Discrete Mathematics and Optimization*, Wiley-Interscience, New York, 2000.
- [7] Kloks, Ton. Treewidth: Computations and approximations. *Lecture Notes in Computer Science*, vol. 842, Springer-Verlag, 1994.
- [8] Lee, C, Lee, J and Oum, S.. Rank-width of Random Graphs. *Preprint, arXiv:1001.0461*, 2010.
- [9] Nachmias, Asaf and Peres, Yuval. Critical random graphs: diameter and mixing time. *Ann. Probab., The Annals of Probability*, vol. 36, 1267–1286, 2008.
- [10] Nešetřil, Jaroslav and Ossona de Mendez, Patrice. Tree-depth, subgraph coloring and homomorphism bounds. *European J. Combin., European Journal of Combinatorics*, vol. 27, 1022–1041, 2006.
- [11] Nešetřil, Jaroslav and Ossona de Mendez, Patrice. Grad and classes with bounded expansion. I. Decompositions. *European J. Combin., European Journal of Combinatorics*, vol. 29, 760–776, 2008.
- [12] Nešetřil, Jaroslav and Shelah, Saharon. On the order of countable graphs. *European J. Combin., European Journal of Combinatorics*, vol. 24, 649–663, 2003.
- [13] Rényi, A. and Szekeres, G.. On the height of trees, *J. Austral. Math. Soc., Australian Mathematical Society. Journal. Series A. Pure Mathematics and Statistics*, vol. 7, 497–507, 1967.