Local Structure Theorems for Erdős–Rényi Graphs and their Algorithmic Application

Jan Dreier¹, Philipp Kuinke¹, Ba Le Xuan², and Peter Rossmanith¹

¹ Theoretical Computer Science, Dept. of Computer Science,
RWTH Aachen University, Aachen, Germany
{dreier,kuinke,rossmani}@cs.rwth-aachen.de
² The Sirindhorn International Thai-German Graduate School of Engineering,
King Mongkut’s University of Technology North Bangkok, Thailand
ba.l-sse2015@tggs-bangkok.org

Abstract. We analyze some local properties of sparse Erdős–Rényi graphs, where \(d(n)/n\) is the edge probability. In particular we study the behavior of very short paths. For \(d(n) = n^{o(1)}\) we show that \(G(n, d(n)/n)\) has asymptotically almost surely (a.a.s.) bounded local treewidth and therefore is a.a.s. nowhere dense. We also discover a new and simpler proof that \(G(n, d/n)\) has a.a.s. bounded expansion for constant \(d\). The local structure of sparse Erdős–Rényi graphs is very special: The \(r\)-neighborhood of a vertex is a tree with some additional edges, where the probability that there are \(m\) additional edges decreases with \(m\). This implies efficient algorithms for subgraph isomorphism, in particular for finding subgraphs with small diameter. Finally we note that experiments suggest that preferential attachment graphs might have similar properties after deleting a small number of vertices.

Keywords: graph theory, random graphs, sparse graphs, graph algorithms

1 Introduction

One of the earliest and most intensively studied random graph models is the Erdős–Rényi model [12]. Graphs from this class are usually depicted as a random variable \(G(n, p)\), which is a graph consisting of \(n\) vertices where each pair of vertices is connected independently uniformly at random with probability \(p\). The edge probability \(p\) may also depend on the size of the graph, e.g. \(p = d/n\). Many properties of Erdős–Rényi graphs are well studied, including but not limited to, threshold phenomena, the sizes of components, diameter and length of paths [1]. One particular impressive result is the 0-1 law: We fix a first-order formula \(\varphi\), if we take a random graph \(G = G(n, 1/2)\), then the probability of \(G \models \varphi\) is either 0 or 1 as \(n \to \infty\) [3].

Unfortunately, there are not so many results about efficient algorithms on random graphs. One can find an optimal coloring of \(G(n, p)\) in expected linear time for \(p < 1.01/n\) [3]. The 0-1 law on the other hand has (not yet) an efficient
accompanying algorithm that can decide, whether \( G \models \varphi \) for \( G = G(n, 1/2) \) and a fixed formula \( \varphi \).

One possibility to open up a whole graph class to efficient algorithms are algorithmic meta-theorems. Such meta-theorems were developed for more and more general graph classes: planar, bounded genus, bounded degree, \( H \)-minor free, \( H \)-topological minor free etc. In all these graph classes we can decide properties that are expressible in first-order logic in linear time for a fixed formula \( \varphi \). Unfortunately, random graph classes do not belong to any of these classes. For example \( G(n, 1.1/n) \) has a.a.s. linear treewidth and does contain constant-size cliques of arbitrary size [7]. Recently, however, graph classes of bounded expansion were introduced by Neˇ setˇ ril and Ossona de Mendez [8]. These classes also admit linear time FO-model checking and generalize the older meta-theorems [9]. The most general model checking algorithm runs in \( O(n^{1+\epsilon}) \) on nowhere-dense classes [10]. In \( G(n, d/n) \), the value \( d \) is the expected density of a random graph. For constant \( d \) it was shown that \( G(n, d/n) \) has a.a.s. bounded expansion [11]. Unfortunately, this does not automatically imply that one can test first-order properties on \( G(n, d/n) \) in linear (expected) time. It implies only that we can test such a property in linear time with a failure probability of \( o(1) \), while the expected runtime might be unbounded. This is for example the case if the runtime grows faster than the failure probability converges to zero. One example of an (expected-time) fpt-algorithm is one, that finds a \( k \)-clique in \( G(n, p(n)) \) in time \( f(k)n^{O(1)} \), for many choices of \( p \) [12].

In Section 3 we find an easier proof for the fact that \( G(n, d/n) \) has a.a.s. bounded expansion for constant \( d \) and give concrete probability bounds, something that was not given before. And we investigate local properties of Erd˝ os–R´ enyi graphs. The expected density of \( G(n, d(n)/n) \) is \( d(n) \) and therefore, for not constant \( d(n) \), unbounded. This implies that \( G(n, d(n)/n) \) does a.a.s. not have bounded expansion. Nevertheless, we show that subgraphs with small diameter are tree-like with only a few additional edges. From this we follow that \( G(n, n^{o(1)}/n) \) has a.a.s. locally bounded treewidth, which implies that they are a.a.s. nowhere dense. Locally bounded treewidth [13] and more generally, locally excluding a minor [14] are useful concepts for developing first-order model checking algorithms that run in time \( O(n^{1+\epsilon}) \).

We discussed that a random graph class which is a.a.s. nowhere dense or has a.a.s. bounded expansion does not directly admit efficient algorithms. Let \( d(n) = n^{o(1)} \). Nevertheless, it is known [15,16] that one can check first-order properties in \( G(n, d(n)/n) \) in time \( O(g(|\varphi|)n^{1+o(1)}) \), for \( d(n) = n^{o(1)} \) and some function \( g \). For constant \( d \) one can check first-order properties in time \( O(g(|\varphi|)n) \).

In Section 4 we use the locally tree-like structure of Erd˝ os–R´ enyi graphs to construct an efficient algorithm for subgraph isomorphism. We show that one can find a subgraph \( H \) with \( h \) vertices and radius \( r \) in \( G(n, d(n)/n) \) in time \( O(2^h h^{3/2} d(n)^3 r n) \), while a naive algorithm may need time \( O(d(n)^h n) \). Therefore, our method may be faster for finding large pattern graphs with small radius.

It can be argued that Erd˝ os–R´ enyi graphs are not a good model for real-world networks and therefore efficient algorithms for Erd˝ os–R´ enyi graphs admit only
limited practical applications. Recently, there were more and more efforts to model real world networks with random graph models. One candidate to meet this goal were the Barabási–Albert graphs, which use a preferential attachment paradigm to produce graphs with a degree distribution that tries to mimic the heavy-tailed distribution observed in many real-world networks [17]. This model is particularly interesting from the point of mathematical analysis because of its simple formulation and interesting characteristics and the reason that Barabási–Albert graphs have been widely studied in the literature [18,19,20]. It was also shown that this model does not have a.a.s. bounded expansion [21].

In Section 5 we discuss experiments to see how similar the local structure of these graphs is to Erdős–Rényi graphs. Not surprisingly, it seems that they are quite different and contain dense subgraphs, and are likely to be somewhere dense.

If we, however, remove the relatively small dense early part of there graphs, the local structure of the remaining part looks quite similar to Erdős–Rényi graphs and indicators hint that the remaining part is indeed nowhere dense. As the dense part is quite small it gives us hope that hybrid algorithms exist that combine different methods for the dense part and the structurally simple part. To search for a subgraph $H$, for example, could be done by guessing which vertices of $H$ lie in the dense part and then using methods from Section 4 to find the remaining vertices in the simple part.

2 Preliminaries

In this work we will denote probabilities by $\mathbb{P}[\cdot]$ and expectation by $\mathbb{E}[\cdot]$. We use common graph theory notation [22]. For a graph $G$ let $V(G)$ be its vertex set and $E(G)$ its edge set. For $v \in V(G)$ we denote the $r$-neighborhood of $v$ by $N_r(v)$. The degree of a vertex $v$ in graph $G$ is denoted by $\deg(v)$. We write $G' \subseteq G$ if $G'$ is a subgraph of $G$. For $X \subseteq V(G)$ we denote by $G[X]$ the subgraph of $G$ that is induced by the vertices in $X$. The graph $G[V(G) \setminus X]$ obtained from $G$ by deleting the vertices in $X$ and their incident edges, is denoted by $G \setminus X$. The treewidth $\text{tw}(G)$ of a graph is a measure how tree-like a graph is. The treewidth of a tree is 1. We denote Erdős–Rényi graphs by a random variable $G(n,d/n)$. We distinguish between graphs with constant $d$ and graphs $G(n,d(n)/n)$ where we allow $d$ to grow (slowly) with $n$. We will use various ways to measure the sparsity of a graph or graph class.

Definition 1 (Shallow topological minor [8]). A graph $M$ is an $r$-shallow topological minor of $G$ if $M$ is isomorphic to a subgraph $G'$ of $G$ if we allow the edges of $M$ to be paths of length up to $r$ in $G$. We call $G'$ a model of $M$ in $G$. For simplicity, we assume by default that $V(M) \subseteq V(G')$ such that the isomorphism between $M$ and $G'$ is the identity when restricted to $V(M)$. The vertices $V(M)$ are called nails and the vertices $V(G') \setminus V(M)$ subdivision vertices. The set of all $r$-shallow topological minors of a graph $G$ is denoted by $G \triangleleft r$.

---

3 In the original definition paths up to length $2r$ were used.
4 Also known as principal vertices
In the classical definition an \( r \)-shallow topological minor consists of an up to \( 2r \) subdivision of the edges. This is only a slight change in favor of better readability of our results and does not change the implication of graphs having or not having such minors compared with the classical definition. With that we can define the clique size over all topological minors of \( G \) as

\[
\omega(G \hat{\nabla} r) = \max_{H \in G \hat{\nabla} r} \omega(H)
\]

**Definition 2 (Topological grad \([23]\)).** For a graph \( G \) and an integer \( r \geq 0 \), the topological grad at depth \( r \) is defined as

\[
\tilde{\nabla}_r(G) = \max_{H \in G \hat{\nabla} r} \frac{|E(H)|}{|V(H)|}
\]

For a graph class \( \mathcal{G} \), define \( \tilde{\nabla}_r(\mathcal{G}) = \sup_{G \in \mathcal{G}} \tilde{\nabla}_r(G) \).

**Definition 3 (Bounded expansion \([23]\)).** A graph class \( \mathcal{G} \) has bounded expansion if and only if there exists a function \( f \) such that for all \( r \geq 0 \), it holds that \( \tilde{\nabla}_r(\mathcal{G}) < f(r) \).

**Definition 4 (Locally bounded treewidth).** A graph class \( \mathcal{G} \) has locally bounded treewidth if and only if there exists a function \( f \) such that for all \( r \geq 0 \), every subgraph with radius \( r \) has treewidth at most \( f(r) \).

**Definition 5 (Nowhere dense \([23]\)).** A graph class \( \mathcal{G} \) is nowhere dense if there exists a function \( f \) such that for all \( G \in \mathcal{G} \) and all \( r \geq 0 \), it holds that \( \omega(G \hat{\nabla} r) < f(r) \)

If a graph class has locally bounded treewidth it is also nowhere dense \([23]\).

### 3 Local Structure and Algorithmic Applications

We want to look at the local structure of Erdős–Rényi graphs and how to exploit it algorithmically. It is already known that Erdős–Rényi graphs have a.a.s. bounded expansion if the edge probability is \( d/n \) for constant \( d \) \([11]\). We will present a simpler proof via a direct method, that also gives concrete probability bounds. The original proof did not give such concrete bounds so we feel that this new proof has applications in the design of efficient algorithms. To make our calculations easier we assume that \( d \geq 2 \), since Erdős–Rényi graphs are only sparser for smaller \( d \), our techniques will also work in this case.

### 3.1 Bounded Expansion

The technique we use to bound the probability that certain \( r \)-shallow topological minors exists is to bound the probability that a path of length at most \( r \) exists between two arbitrary vertices.
Lemma 1. Let \( p_r \) be the probability that there is a path of length at most \( r \) between two arbitrary but fixed vertices in \( G(n, d/n) \). It holds that
\[
\frac{d}{n} \leq p_r \leq \frac{2d^r}{n}.
\]

Proof. Since all edges are independent, we do not need to identify the start and end vertices of the path. We prove by induction over \( r \), that the probability of the existence of a path of length exactly \( r \) is bounded by \( d^r n \). For \( r = 1 \) the statement holds: \( p_1 \leq d n \). We will use the union bound. The probability of a path of length \( r \) is at most that of all paths of length \( r - 1 \) and then the probability of a single edge:
\[
p_r \leq \sum_{k=0}^{n} p_{r-1} p_1 \leq \sum_{k=0}^{n} \frac{d^{r-1} d}{n^2} \leq \frac{d^r}{n}
\]
By using the union bound and assuming that \( d \geq 2 \), it holds that the joint probability is bounded by \( \frac{2d^r}{n} \).

Having this bound in place, we can show that \( G(n, d/n) \) has a.a.s. no \( r \)-shallow topological minors of large density from which it follows that they are contain in a graph class of bounded expansion a.a.s.

Theorem 1 (*). \( G(n, d/n) \) is a.a.s. contained in a graph class of bounded expansion. In particular, for \( d \geq 16 \) the probability that such a random graph contains some \( r \)-shallow topological minor of size \( k \) and at least \( 8k d^r \) edges is at most \( \max\{n^{-2k}, 2^{-n^{2/3}}\} \). For \( d < 16 \) the same result holds for at least \( 8k 16^r \) edges.

Proof. We will now investigate the probability that a random graph \( G = G(n, d/n) \) contains some model of an \( r \)-shallow topological minor \( H \) with nail set \( v_1, \ldots, v_k \). Such a model consists of the nails themselves and vertex-disjoint paths of maximal length \( r \) between them. Each such path models one edge of \( H \). Assume that \( V(H) = \{u_1, \ldots, u_k\} \) and that \( u_i \) is modeled by \( v_i \). Then an edge \( u_i u_j \in E(H) \) is modeled by a path from \( u_i \) to \( u_j \) in \( G \) and all these paths are vertex-disjoint. What is the probability that such a model exists? A first path exists with a probability of \( p_r \). The probability that the second path exists under the condition that it does not cross all candidates for the first path is slightly less than \( p_r \). Continuing this argument shows that the probability of finding such a model is at most \( p_r^{\left| E(H) \right|} \) and more specifically it is at most the probability of getting \( \left| E(H) \right| \) heads after \( \left| E(H) \right| \) independent coin tosses with a head probability of \( p_r \). Moreover, this implies that the probability of finding a model for some \( H \) with \( m \) edges is at most the probability of getting at least \( m \) heads after tossing \( \binom{k}{2} \) such coins.

Let \( X \) be the sum of \( \binom{k}{2} \) independent Bernoulli variables with \( \Pr[X = 1] = p_r \). Using the bounds of Lemma 1 we have
\[
\frac{4k^2}{n} \leq \frac{k^2 d}{n} \leq E[X] \leq \frac{k}{2} \frac{2d^r}{n} \leq \frac{k^2 d^r}{n}.
\]
Using Chernoff bounds with $\delta = \frac{8\kappa}{r} - 1$ we get

$$
\Pr \left[ X > (1 + \delta) \frac{k^2 r^r}{n} \right] \leq \Pr[X > (1 + \delta) \mathbb{E}[X]] \leq \left( \frac{e^\delta}{(1 + \delta)^{1 + \delta}} \right)^{\mathbb{E}[X]} \leq \left( \frac{ek}{8n} \right)^{32k} .
$$

This means for fixed $k$ nails, with probability of at most $\left( \frac{ek}{8n} \right)^{32k}$ the graph $G$ contains a model for an $r$-shallow topological minor with these nails and at least $(1 + \delta) \frac{k^2 r^r}{n} = 8kd^r$ edges. The density of such a topological minor is therefore $8d^r$. There are only $\binom{n}{k} \leq \left( \frac{nk}{k} \right)^k$ possibilities to choose the nails, so an $r$-shallow topological minor with $k$ nails and density at least $8d^r$ exist with a probability of at most $\left( \frac{nk}{k} \right)^k \left( \frac{ek}{8n} \right)^{32k}$, which is $n^{-2k}$ if $k \leq n^{2/3}$. For bigger $k$ it is bounded by $2^{-n^{2/3}}$. Therefore, every $r$-shallow topological minor in $G$ has a.a.s. a density of at most $8d^r$.

### 3.2 Locally Simple Structure

It is known that even for constant $d$, the treewidth of $G(n, d/n)$ grows with $\Omega(n)$ [7]. Furthermore, $G(n, d(n)/n)$ does a.a.s not have bounded expansion if $d(n)$ is unbounded. We now show that $G(n, n^{o(1)}/n)$, nevertheless has locally bounded treewidth and thus is a.a.s. nowhere dense. We start by counting the expected number of occurrences of a certain subgraph in $G(n, d(n)/n)$.

**Lemma 2.** The expected number of subgraphs with $k$ vertices and at least $k + m$ edges in $G(n, d(n)/n)$ is at most

\[
\left( \frac{d(n)^{k+m}}{n^m} \right) \text{ for all } n.
\]

**Proof.** There are $\binom{n}{k}$ subsets $S$ of size $k$ in $G$. If we fix $k$ vertices, we want $k + m$ edges to be present, which has probability $\left( \frac{d(n)}{n} \right)^{k+m}$ by independence of edges. Let us denote by $X_S$ the event that vertex set $S$ has $k + m$ edges. The expected number of desired subgraphs is then $\mathbb{E}[\sum S X_S]$. We note that the events $X_S$ are not independent of each other, but with linearity of expectation we have:

\[
\mathbb{E} \left[ \sum S X_S \right] = \sum S \mathbb{E}[X_S] \leq \binom{n}{k} \left( \frac{d(n)}{n} \right)^{k+m} \leq \frac{n^k d(n)^{k+m}}{n^{k+m}} \leq \frac{d(n)^{k+m}}{n^m}
\]

Which proves the desired upper bound. \(\square\)

From Lemma 2 we can conclude a well known property of Erdős–Rényi graphs, that the expected number of cycles of length $r$ is around $d(n)^r$ (which is a constant if $d$ is constant), by setting $k = r$ and $m = 0$. We now use this Lemma to make statements about the density of neighborhoods.

**Lemma 3.** The probability that there exists an $r$-neighborhood of $G(n, d(n)/n)$ which contains $m$ more edges than vertices is bounded by $d(n)^{2r} \left( \frac{d(n)^{2r+1}}{n} \right)^m$.

**Proof.** Consider any $r$-neighborhood with $\ell$ vertices. Assume the neighborhood contains at least $m$ more edges than vertices. Let $T$ be a breadth-first search
spanning tree of this neighborhood. Since $T$ contains $\ell$ vertices and $\ell - 1$ edges, there are $m + 1$ edges which are not contained in $T$. Each extra edge is consists of two vertices. Let the set of these vertices be $U$. Let $H$ be the graph induced by the union of the $m + 1$ extra edges and for each $u \in U$ the unique path in $T$ from $u$ to the root of $T$. Since $|U| \leq 2(m + 1)$ and each path to the root in the breadth-first-search tree $T$ has length at most $r$, the number of vertices of $H$ is bounded by $2r(m + 1)$.

In summary, if there exists an $r$-neighborhood with at least $m$ more edges than vertices then there exists a subgraph with $k \leq 2r(m + 1)$ vertices and $m$ more edges than vertices. But according to Lemma 2, the expected number of such subgraphs is bounded by

$$d(n)^{2r(m+1)+m} = d(n)^{2r} \left( \frac{d(n)^{2r+1}}{n} \right)^m.$$ 

By employing a first moment argument, the same value also bounds the probability that such a subgraph exists. 

**Theorem 2.** $G(n, d(n)/n)$ has a.a.s. locally bounded treewidth, for $d(n) = n^{o(1)}$.

**Proof.** Since $d(n) = n^{o(1)}$, there exists a monotone decreasing function $g(n)$ with $d \leq n^{g(n)}$ and $\lim_{n \to \infty} g(n) = 0$. Let $f(r)$ be the inverse function of $1/(8g(n))$. Since $g(n)$ is monotone decreasing, $f(r)$ exists and is monotone increasing. We show that for all $r \geq 0$ every subgraph with radius $r$ has a.a.s. treewidth at most $f(r)$. We distinguish between two cases. The first case is $r < 1/(8g(n))$. According to Lemma 3, an $r$-neighborhood of $G$ has more edges than vertices with probability at most

$$\frac{d(n)^{2r} n^{m}}{n^{m}} \leq d(n)^{2r} \left( \frac{d(n)^{2r+1}}{n} \right)^m \leq n^{-1/2 + g(n)} = o(1)$$

We can conclude that every $r$-neighborhood has a.a.s. treewidth at most $2 \leq f(r)$.

The second case is $r \geq 1/(8g(n))$. This means $f(r) \geq n$. The treewidth of a graph with $n$ vertices is bounded by $n \leq f(r)$. 

**4 Algorithm for Subgraph Isomorphism**

In this section we solve Subgraph Isomorphism, which given a graph $G$ and a graph $H$ asks, whether $G$ contains $H$ as a subgraph. This is equivalent to FO-model checking restricted to only $\exists$ quantifiers. We need the following lemma.

**Lemma 4.** The expected size of an $r$-neighborhood in $G(n, d(n)/n)$ is bounded by $2d(n)^r$.

**Proof.** Let $v$ be an arbitrary but fixed vertex and let $X_i$ be a random variable with $X_i = 1$ if and only if there is a path of length at most $r$ from $v$ to $i$. From Lemma 1 we know that $\mathbb{E}[X_i] \leq \frac{2d(n)^r}{n}$. Let $N_r(v)$ be the size of an $r$-neighborhood of
vertex \( v \). We will use linearity of expectation that does not need the \( X_i \) to be independent.

\[
E[N_r(v)] \leq E\left[\sum_{i=1}^{n} X_i\right] \leq \sum_{i=1}^{n} E[X_i] \leq \sum_{i=1}^{n} \frac{2d(n)^r}{n} \leq 2d(n)^r
\]

Thus we have proven the claim. \( \square \)

Let \( H \) be a connected graph with \( h \) vertices and radius \( r \). In this section we discuss how fast it can be decided whether \( G(n, d(n)/n) \) contains \( H \) as a subgraph. We first discuss the runtime of simple branching algorithms on Erdős–Rényi graphs and how exploiting local structure may lead to better runtimes. We discovered that if the radius \( r \) of the pattern graph is small, an approach based on local structure is significantly faster.

For low-degree graphs there exists a simple branching algorithm to decide whether a graph \( G \) contains \( H \) as a subgraph in time \( O(\Delta^h n) \), where \( \Delta \) is the maximal degree in \( G \). Let us first assume that \( d(n) = d \) is constant. There is nevertheless a non-vanishing probability that the maximal degree of \( G(n, d/n) \) is as large as \( \sqrt{\log n} \). Therefore, the maximal degree cannot be bounded by any function of \( d \). This implies that a naive, maximal degree based algorithm may have at least a quasi-linear dependence on \( n \), while we present an algorithm which has only a linear dependence on \( n \).

Let us also assume that that \( d(n) \) is of order \( \log n \) and even that the maximum degree is bounded by \( O(d(n)) \). A naive branching algorithm may therefore decide whether \( G(n, d(n)/n) \) contains \( H \) in expected time \( O(d(n))^hn) \). We improve this result, not making any assumption about the maximal degree, by replacing the factor \( O(d(n))^h \) in the runtime with \( 2^h h^{3/2} d(n)^{3r} \), where \( r \) is the radius of \( H \). For graphs with small radius, the runtime is no more longer dominated by a factor \( O(d(n))^h \), but by \( 2^h h^{3/2} \), which may be significantly smaller when \( d(n) \) is, for example, of order \( \log n \).

So far we only discussed connected subgraphs. Using color-coding techniques, the results in this section can easily be extended to disconnected subgraphs, where the radius of each component is bounded by \( r \). Color-coding may, however, lead to an additional factor of \( c^h \) in the runtime: Assume \( H \) has \( c \) components where the size of \( H \) is \( h \). We want to color each vertex of \( G \) uniformly at random. Assume \( G \) contains \( H \), then the probability that every component of \( H \) can be embedded using vertices of a single color is at least \( 1/c^h \). So if \( H \) can be embedded in \( G \) we will answer yes after an expected number of \( c^h \) runs.

For the following result notice that if \( d(n) \) is polylogarithmic in \( n \) the runtime is quasi-linear in \( n \). For \( d(n) = n^{o(1)} \) the dependence on \( n \) is \( n^{1+o(1)} \). The algorithm is given in the proof for Theorem 5.

**Theorem 3.** Let \( H \) be a connected graph with \( h \) vertices and radius \( r \). If \( d(n)^{2r+1} < n \) and \( h \leq \log(n/d(n)^{2r+1}) \) one can decide whether \( G(n, d(n)/n) \) contains \( H \) as a subgraph in expected time \( O(2^h h^{3/2} d(n)^{3r} n) \).

**Proof.** In the following we describe an algorithm to solve this problem. First, we enumerate all vertices \( v \in V \) and construct a spanning tree \( T \) for \( N_r(v) \). Let \( M \)
be the set of edges in $N_r(v)$ which are not contained in $T$. Let $m = |M|$ and enumerate all sets $S \subseteq E(H)$ with $|S| \leq m$ and all embeddings $f': H[S] \rightarrow M$. We can answer with no, if there is no $S$, such that $H - S$ is a forest. If for some $S$, $H - S$ is a forest and there exists an embedding $f: (H - S) \rightarrow T$ which is consistent with $f'$ we can answer yes.

If there exists no embedding of $H$, the described algorithm will answer with no. We will now show that the algorithm will answer with yes, if such an embedding exists. If there exists an embedding of $H$ there also exists a $v \in V$ such that $H$ can be embedded into $N_r(v)$. Let again $T$ be a spanning tree of $N_r(v)$ and $M$ be the set of edges in $N_r(v)$ not contained in $T$. There exists $S \subseteq E(H)$ such that $S$ is the set of exactly those edges in $H$ which are mapped to $M$. It follows that $|S| \leq |M| = m$. It holds, that no edge from $H - S$ is mapped to $M$, therefore, $H - S$ can be embedded into $T$. The graph $H - S$ is a forest. Let $f': H[S] \rightarrow M$ be a restricted embedding. The embedding of $H$ to the neighborhood can be reconstructed by extending $f'$. Therefore the forest $H - S$ can be embedded into the tree $T$ in a way consistent with $f'$.

Let us now analyze the runtime. We first calculate the expected runtime $E_v$ for deciding whether $H$ can be embedded into $N_r(v)$. Let $T$, $M$, $m$ as usual. We need to iterate over all sets $S$ and all partial embeddings $f'$. There are

$$\sum_{s=1}^{h} \binom{h}{s} \binom{m}{s} s! \leq 2^h m^h$$

distinct ways in which one can choose a subset $S \subseteq E(H)$ and map it to $M$. One can decide in time $O((l^{3/2}/\log(l))n)$ whether a graph of size $l$ can be embedded in a tree of size $n$ [24]. Let $N = |N_r(v)|$. The runtime for a fixed $m$ therefore is

$$2^h m^h O\left(\frac{h^{3/2} N}{\log(h)}\right).$$

Large values for $m$ lead to high run-times, but are unlikely. According to Lemma 3 the expected runtime $E_v$ can be bounded as follows.

$$E_v \leq \sum_{m=1}^{\infty} 2^h m^h O\left(\frac{h^{3/2} N}{\log(h)}\right) \sum_{r=0}^{\infty} d(n)^{2r} \left(\frac{d(n)^{2r+1}}{n}\right)^m$$

$$= O(2^h h^{3/2} d(n)^{2r} N) \sum_{m=1}^{\infty} m^h \left(\frac{d(n)^{2r+1}}{n}\right)^m$$

According to Lemma 4 the expected size of an $r$-neighborhood is bounded by $2d(n)^r$. By linearity of expectation, this leads to the following expected runtime $E$.

$$E = \sum_{v \in V} O(2^h h^{3/2} d(n)^{2r} \mathbb{E}[|N_r(v)|]) \sum_{m=1}^{\infty} m^h \left(\frac{d(n)^{2r+1}}{n}\right)^m$$

$$= O(2^h h^{3/2} d(n)^{3r} n) \sum_{m=1}^{\infty} m^h \left(\frac{d(n)^{2r+1}}{n}\right)^m$$
Because \( h \leq \log(n/d(n)^{2r+1}) \) we get \( m^h \leq (n/d(n)^{2r+1})^{\log m} \). In this way we can estimate the infinite sum in the previous equation by a geometric sum as follows:

\[
\sum_{m=1}^{\infty} m^h \left( \frac{d(n)^{2r+1}}{n} \right)^m = O(1) + \sum_{m=2}^{\infty} \left( \frac{d(n)^{2r+1}}{n} \right)^{m/2} = O(1).
\]

This means we can ignore the contribution of the sum and get a runtime of \( O(2^h h^{3/2} d(n)^{3r} n) \).

\[ \square \]

5 Experimental Evaluation of Barabási–Albert-Graphs

In the previous section, we showed that Erdős–Rényi graphs have bounded expansion for edge probability \( p = d/n \) (with constant \( d \)) and are nowhere dense with \( p = n^{o(1)}/n \). In this section, we discuss the sparsity of the Barabási–Albert model. It is known that this model has not a.a.s. bounded expansion, because it contains an unbounded clique with non-vanishing probability [21]. It is not known, however, if it is (or is not) a.a.s. somewhere-dense. Our experiments seem to imply that on average Barabási–Albert graphs seem to be dense but that this density is limited to early vertices: In the Barabási–Albert model, vertices with high degree tend to be preferred for new connections. This means that edge probabilities are not independent. Moreover, the expected degree \( d(i) = \sqrt{n/i} \) for a vertex \( i \) is less uniform than it is for Erdős–Rényi graphs, where \( d(i) = pn \).

To evaluate the expansion properties of the Barabási–Albert-model, we compute transitive fraternal augmentations and \( p \)-centered colorings. These have been introduced by Nešetřil and Ossona de Mendez, and are highly related to bounded expansion and a tool for developing new and faster algorithms. A graph class has bounded expansion if and only if the maximum in-degree of transitive fraternal augmentations is bounded, or the graph admits a \( p \)-centered coloring with bounded number of colors.

**Definition 6 (Transitive fraternal augmentation [8]).** Let \( \vec{G} \) be a directed graph. A 1-transitive fraternal augmentation of \( \vec{G} \) is a directed graph \( \vec{H} \) with the same vertex set, including all the arcs of \( \vec{G} \) and such that, for any vertices \( x, y, z \),

- if \( (x, z) \) and \( (z, y) \) are arcs of \( \vec{G} \) then \( (x, y) \) is an arc of \( \vec{H} \) (transitivity),
- if \( (x, z) \) and \( (y, z) \) are arcs of \( \vec{G} \) then \( (x, y) \) or \( (y, x) \) is an arc of \( \vec{H} \) (fraternity).

A transitive fraternal augmentation of a directed graph \( \vec{G} \) is a sequence \( \vec{G}_1 \subseteq \cdots \subseteq \vec{G}_i \subseteq \cdots \subseteq \vec{G}_n \), such that \( \vec{G}_i+1 \) is a 1-transitive fraternal augmentation of \( \vec{G}_i \).

**Definition 7 (\( p \)-centered coloring [25]).** For an integer \( p \), a \( p \)-centered coloring of \( G \) is a coloring of the vertices such that any connected subgraph \( H \) induced on the vertices of an arbitrary set of \( i \) colors \( (i \leq p) \), \( H \) must have at least one color that appears exactly once.
Showing that the maximum in-degree of a transitive fraternal augmentation or the number of colors needed for a \( p \)-centered coloring does not grow with the size of the graph is a way to prove that a graph has bounded expansion [8]. When designing algorithms, \( p \)-centered colorings can be used to solve hard problems efficiently. By using \( p \)-centered colorings, we can decompose a graph into small, well-structured subgraphs such that \( \text{NP} \)-hard problems can be solved easily on each subgraph before combining these small solutions to get a solution for the entire graph. It is important that the number of colors needed for a \( p \)-centered coloring for a fixed \( p \) is small, as the runtime usually is a function of the number of colors needed. If a graph class does not have bounded expansion; that is, the number of colors grows with \( n \), but very slowly, such as \( \log \log n \), using these algorithms might still be practical.

One example problem which can be solved directly using \( p \)-centered colorings is **Subgraph Isomorphism**, where one asks if a graph \( H \) is contained in a graph \( G \) as a subgraph. In general graphs, this problem is \( \mathsf{W}[1] \)-hard when parameterizing by the size of \( H \) [20]. However, there exist an algorithm, whose runtime is a function of the number of colors needed for a \( p \)-centered coloring, where \( p \) depends on the size of \( H \) [23]. So, regardless of the fact whether Barabási–Albert graphs are theoretically sparse or not, calculating the number of colors of a \( p \)-centered coloring for different graph sizes has direct impact on the feasibility of a whole class of algorithms on these graphs.

### 5.1 Experiment Overview

We analyze the expansion properties of Barabási–Albert graphs by computing transitive fraternal augmentations and \( p \)-centered colorings. In the following, we describe the heuristics used to compute these. In order to compute the transitive fraternal augmentations of a graph \( G \), the graph is oriented to a directed graph \( \overrightarrow{G_1} \) by using low-degree orientation, in which every edge \((u,v)\) in \( G \) is transformed to an arc \((u,v)\) in \( \overrightarrow{G_1} \) if the degree of \( u \) is greater than the degree of \( v \). Then, transitive fraternal augmentations are applied to \( \overrightarrow{G_1} \), which yield a sequence \( \overrightarrow{G_1}, \overrightarrow{G_2}, \ldots, \overrightarrow{G_i} \). The augmentation heuristic we used was proposed in earlier work [27]: To build graph \( \overrightarrow{G_i} \) from \( \overrightarrow{G_{i-1}} \) we need to perform transitive fraternal augmentations. First we create the set \( F \) of fraternal edges of \( \overrightarrow{G_{i-1}} \). Let \( G_F \) be the graph induced by \( F \). Now we can orient the edges of \( G_F \) by the same low-degree orientation performed earlier to get the directed fraternal edges \( \overrightarrow{F} \) that are added to \( \overrightarrow{G_{i-1}} \) and result in \( \overrightarrow{G_i} \). Now we can color the undirected graph \( G_i \) of \( \overrightarrow{G_i} \) by iterating through the vertices in a descending-degree order and assign each vertex the lowest color that does not appear in its neighborhood. We then check whether that coloring is a \( p \)-centered coloring of the input graph \( G \). If this is not the case we repeat this procedure for \( G_{i+1} \).
Algorithm 1 Computing $p$-centered colorings

1: procedure Compute-$p$-centered-Colorings($G, p$)
2: Create $G_1$ from $G$ using low-degree orientation
3: $i \leftarrow 1$
4: loop
5: $c \leftarrow$ Greedy coloring of the undirected graph $G_i$
6: if $c$ is a $p$-centered coloring of $G$ then
7: return $c$
8: end if
9: Compute 1-transitive fraternal augmentation of $G_i$ to get $G_{i+1}$
10: $i \leftarrow i + 1$
11: end loop
12: end procedure

5.2 Barabási–Albert Graphs are Empirically Dense

First, we analyze the maximum in-degree of transitive fraternal augmentations. We ran the previously described algorithm on random Barabási–Albert graphs with $d = 2$ for different sizes ($500 \leq n \leq 3000$) and calculated the maximum in-degree of up to five transitive fraternal augmentation steps. The results are shown in Figure 1a. Each data point is an average over ten runs with the same $n$. For all graphs both the maximum in-degree grows with $n$, which would not be the case for graphs with bounded expansion.

To evaluate how well the expansion properties of Barabási–Albert graphs can be practically exploited, we analyzed the number of colors needed to construct $p$-centered colorings. We constructed 3- and 4-centered colorings with the same graph parameters and sizes than before. The results are shown in Figure 1b. For the analyzed range, the number of colors needed grows steadily. Furthermore, the number of colors needed to construct 4-centered colorings is substantially higher than the number of colors needed for 3-centered colorings. Computing higher order colorings or colorings for larger graphs was infeasible with the used algorithm. It seems practically impossible to use $p$-centered colorings algorithmically for Barabási–Albert graphs. We have to note that the used algorithm is only a heuristic and the real values might be much better than what we have computed. But since these heuristics work well for graphs that have low treedepth colorings, it is unlikely that the graphs have bounded coloring number for $p$-centered colorings.

5.3 Density Seems Limited to Early Vertices

Previously, we showed that the colors needed to construct $p$-centered colorings of small graphs can be very high. In this section we discover that the early vertices of the random process heavily affect these results. We remove the first 10% of the vertices added in the random process and analyze the maximum in-degree of transitive fraternal augmentations and number of colors needed
(a) Transitive fraternal augmentations

(b) $p$-centered colorings

Fig. 1: Results for Barabási–Albert graphs with $d = 2$ for increasing $n$.

(a) Transitive fraternal augmentations

(b) $p$-centered colorings

Fig. 2: Results for Barabási–Albert graphs with $d = 2$ for increasing $n$ after deleting the first 10% of vertices.
to construct $p$-centered colorings. By removing those 10%, we can construct $p$-centered colorings for much larger graphs ($5000 \leq n \leq 30000$), see Fig. 2. The required number of colors for $p$-centered colorings and maximum in-degree of transitive fraternal augmentations remain stable and do not seem to depend on the number of vertices. This suggests that these 10% of the early vertices contain almost all of the density of Barabási–Albert graphs. This is of course a linear factor and it remains to see if one can use much smaller functions of $n$, like for example $\log n$. The sizes of the graphs at hand, however, were not large enough to investigate sub-linear functions of $n$ with a meaningful result.

6 Conclusion

In this work we gave an alternative proof that $G(n, d/n)$ has a.a.s. bounded expansion and have shown that $G(n, d(n)/n)$ with $d(n) = n^{o(1)}$ has a.a.s. locally bounded treewidth. Our results are based on the fact that local neighborhoods of Erdős–Rényi graphs are tree-like with high probability. It is known [15] that for a graph $G = G(n, d(n)/n)$ with $d(n) = n^{o(1)}$ and a first-order formula $\varphi$ one can decide whether $G \models \varphi$ in expected time $f(|\varphi|)n^{1+o(1)}$ for some functions $f$ and $g$. This result can also be proven using our techniques. It remains to show whether it is possible to answer this question in linear expected fpt time (where $d(n)n$ is the expected number of edges), i.e. $O(f(|\varphi|)d(n)n)$. In this paper, we also presented a more efficient algorithm for the subgraph isomorphism problem on Erdős–Rényi graphs if the pattern graph has small radius. It would be interesting to consider other measures for the pattern graph as well, such as treewidth or treedepth. Furthermore, we gathered empirical evidence which suggests that Barabási–Albert graphs are somewhere dense. It would be interesting to prove this conjecture.

References

1. Bollobás, B.: Random Graphs. 2nd edn. Cambridge University Press (2001)
2. Erdős, P., Rényi, A.: On random graphs. Publicationes Mathematicae 6 (1959) 290–297
3. Fagin, R.: Probabilities on finite models. J. Symbolic Logic 41(1) (03 1976) 50–58
4. Coja-Oghlan, A., Taraz, A.: Colouring random graphs in expected polynomial time. In: STACS 2003. (2003) 487–498
5. Dawar, A., Grohe, M., Kreutzer, S.: Locally excluding a minor. In: 22nd Annual IEEE Symposium on Logic in Computer Science (LICS 2007). (July 2007) 270–279
6. Flum, J., Frick, M., Grohe, M.: Query evaluation via tree-decompositions. J. ACM 49(6) (November 2002) 716–752
7. Gao, Y.: Treewidth of Erdős–Rényi random graphs, random intersection graphs, and scale-free random graphs. Discrete Appl. Math. 160(4-5) (March 2012) 566–578
8. Nešetřil, J., Ossona de Mendez, P.: Grad and classes with bounded expansion I. Decompositions. European Journal of Combinatorics 29(3) (2008) 760–776
9. Dvořák, Z., Král, D., Thomas, R.: Testing first-order properties for subclasses of sparse graphs. J. ACM 60(5) (2013) 36:1–36:24
10. Grohe, M., Kreutzer, S., Siebertz, S.: Deciding first-order properties of nowhere dense graphs. In: Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing. STOC ’14, New York, NY, USA, ACM (2014) 89–98
11. Nešetřil, J., de Mendez, P.O., Wood, D.R.: Characterisations and examples of graph classes with bounded expansion. European Journal of Combinatorics 33(3) (2012) 350 – 373 Topological and Geometric Graph Theory.
12. Fountoulakis, N., Friedrich, T., Hermelin, D.: On the average-case complexity of parameterized clique. Theoretical Computer Science 576 (2015) 18 – 29
13. Frick, M., Grohe, M.: Deciding first-order properties of locally tree-decomposable structures. J. ACM 48(6) (November 2001) 1184–1206
14. Dawar, A., Grohe, M., Kreutzer, S.: Locally excluding a minor. In: 22nd IEEE Symposium on Logic in Computer Science (LICS 2007), 10-12 July 2007, Wroclaw, Poland, Proceedings, IEEE Computer Society (2007) 270–279
15. Grohe, M.: Logic, graphs, and algorithms (2007)
16. Grohe, M. In: Generalized Model-Checking Problems for First-Order Logic. (2001) 12–26
17. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. 286(5439) (1999) 509–512
18. Cohen, R., Havlin, S.: Scale-free networks are ultrasmall. Phys. Rev. Lett. 90 (Feb 2003) 058701
19. Kamrul, M.H., Hassan, M.Z., Pavel, N.I.: Dynamic scaling, data-collapse and self-similarity in Barabási–Albert networks. Journal of Physics A: Mathematical and Theoretical 44(17) (2011) 175101
20. Klemm, K., Eguluz, V.M.: Growing scale-free networks with small-world behavior. Phys. Rev. E 65 (May 2002) 057102
21. Demaine, E.D., Reidl, F., Rossmanith, P., Sánchez Villaamil, F., Sikdar, S., Sullivan, B.D.: Structural sparsity of complex networks: Random graph models and linear algorithms. CoRR abs/1406.2587 (2014)
22. Diestel, R.: Graph Theory. (2010)
23. Nešetřil, J., Ossona de Mendez, P.: Sparsity: Graphs, Structures, and Algorithms. (2014)
24. Shamir, R., Tsur, D.: Faster subtree isomorphism. Journal of Algorithms 33(2) (1999) 267 – 280
25. Nešetřil, J., Ossona de Mendez, P.: Grad and classes with bounded expansion II. Algorithmic aspects. European Journal of Combinatorics 29(3) (2008) 777–791
26. Downey, R.G., Fellows, M.R.: Parameterized Complexity
27. Reidl, F.: Structural Sparseness and Complex Networks. PhD thesis, RWTH Aachen University (2015)