Structural Sparsity of Complex Networks: Bounded Expansion in Random Models and Real-World Graphs

Erik D. Demaine[*]   Felix Reidl[†]   Peter Rossmanith[†]

Fernando Sánchez Villaamil[†]   Somnath Sikdar[†]   Blair D. Sullivan[‡]

Abstract

This research aims to identify strong structural features of real-world complex networks, sufficient to enable a host of graph algorithms that are much more efficient than what is possible for general graphs (and currently used for network analysis). Specifically, we study the property of *bounded expansion*[1]—roughly, that any subgraph has bounded average degree after contracting disjoint bounded-diameter subgraphs—which is the strongest formalization of the well-observed notion of "sparsity" that might possibly apply to real-world networks.

On the theoretical side, we analyze many previously proposed models for random networks and characterize which ones have bounded expansion. We show that, with high probability, (1) graphs sampled with either the Molloy–Reed configuration model (including a variation of the model which achieves high clustering) or the Chung–Lu model with a prescribed sparse degree sequence (including heavy-tailed degree distributions); (2) Erdős–Rényi random graphs starting from a bounded-degree graph; (3) stochastic block models with small probabilities; result in graphs of bounded expansion. We also prove that the Kleinberg model and the Barabási–Albert model, in fairly typical setups, contain large one-subdivisions of cliques and thus do not result in graphs of bounded expansion.

On the practical side, we give experimental evidence that many complex networks have bounded expansion, by measuring the closely related "low treedepth coloring number" on a corpus of real-world data.

On the algorithmic side, we show how tools provided by the bounded expansion framework can be used to efficiently solve the following common network analysis problems: for a fixed graph H, we obtain the fastest-known algorithm for counting the number of induced H-subgraphs and the number of H-homomorphisms; and we design linear algorithms for computing several centrality measures.

Keywords: complex networks · social networks · sparsity · bounded expansion · Molloy-Reed · scale-free · treedepth · parameterized algorithms

[*]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA. edemaine@mit.edu.

[†]Theoretical Computer Science, Dept. of Computer Science, RWTH Aachen University, Aachen, Germany, {reidl,rossmani,fernando.sanchez,sikdar}@cs.rwth-aachen.de.

[‡]Department of Computer Science, North Carolina State University, Raleigh, NC, blair_sullivan@ncsu.edu.

[1]Not to be confused with the notion of expansion related to expander graphs.

1 Introduction

Complex networks vs. structural graph algorithms. Social networks (such as Facebook or physical disease propagation networks), biological networks (such as gene interactions or brain networks), and informatics networks (such as autonomous systems) are all examples of *complex networks*, which have been the attention of much study in recent years, given the surge of available network data. Modeled as graphs, complex networks seem to share several structural properties. Perhaps most famous is the *small-world* property ("six degrees of separation"): typical distances between vertex pairs are small compared to the size of the network. Another important property is that their degree distribution tends to be heavy-tailed, i.e. distributions whose tails are not exponentially bounded. In many cases, this degree distribution is close to a *power law*: the fraction of vertices of degree k is proportional to $k^{-\gamma}$, for some constant γ typically between 2 and 3. Networks furthermore often exhibit *high clustering* and admit a natural division into a *community structure*. Despite the clustering property, complex networks are *sparse*: the ratio of edges to vertices (edge density) is usually small.

On the other hand, the field of structural graph algorithms has led to impressively efficient and precise algorithms (efficient PTASs, subexponential fixed-parameter algorithms, linear kernelizations, etc.) for increasingly general families of graphs; see, e.g., [23–26, 29, 33, 42]. Many such results proved initially for planar graphs have since been extended to bounded-genus graphs, graphs of bounded local treewidth, and graphs excluding a fixed minor; yet such results are known to be impossible for general graphs. Can we apply these powerful algorithms to analyze complex networks?

Unfortunately, the above-mentioned structural properties of complex networks seem too weak to enable better algorithms, while the discussed graph classes seem too restrictive to apply to complex networks. The goal of this paper is to bridge this gap, by identifying a more general graph class that simultaneously enables better algorithms and includes many complex networks.

Bounded expansion. In general, complex networks seem to exhibit an intermediate-scale structure composed of small dense parts—representing clusters/communities—that are sparsely interconnected. This hierarchical behavior has been established for many networks [52] and is consistent with the tree-like intermediate structure observed in [4].

How can this notion be captured formally? If we contract disjoint small-diameter subgraphs (representing potentially dense local substructures in the network), then the resulting graph (representing the global connectivity of these substructures) should be sparse. This is the notion of an *r-shallow minor*,

2

where r is the maximum diameter of the subgraphs that were contracted in the construction process. (For formal definitions, see Section 2.) We cannot expect the edge density of all r-shallow minors to be constant (then r would play no role), but we require it to grow as any function of r, and thus be independent of the size of the graph. A graph class for which this property holds has *bounded expansion*, a concept introduced by Nešetřil and Ossona de Mendez [65].

Theoretical results. Since the definition of bounded expansion applies to graph classes instead of individual graphs, it is impossible to settle this question empirically. To ground our hypothesis in theory, we analyze several random graph models which were designed to mimic the properties of specific types of real world networks. Although not perfect, random graph models play a central role in network analysis, both to guide our understanding of complex networks and as a convenient source of synthetic data for algorithm testing and validation. In our case, random graph models allow us to determine whether (synthetic) complex networks have bounded expansion with high probability. We analyze several popular random graph models:

(i) the Configuration Model [63] and the Chung–Lu Model [19, 20] with specified asymptotic degree sequences, which includes graphs with heavy-tailed degree distributions;

(ii) a variant of the Configuration Model which achieves high clustering [8];

(iii) a significant generalization of Erdős–Rényi graphs (allowing the network to be built on top of an arbitrary base graph of bounded degree), which includes the stochastic block model with small probabilities;

(iv) the Kleinberg Model [43, 44];

(v) and the Barabási–Albert Model [6, 9].

We will show that the Configuration Model, the Chung–Lu Model and our extension of Erdős–Rényi graphs have bounded expansion, while the Kleinberg Model and the Barabási–Albert Model do not (actually they are not even *nowhere dense*, a strict generalization of bounded expansion).

Experimental results. We present an experimental study suggesting that important real-world networks have small *grad*, which is the density measure for single graphs that defines the expansion for graph classes. Interestingly, the algorithmic tools that become efficient when the grad of a graph is small can be directly applied without knowing its actual value.

We will make extensive use of *low-treedepth colorings*: for any integer p, a graph can be colored with $f(p)$ colors such that any set of at most $p - 1$

colors induced a graph of low treedepth, where f only depends on the grad of the graph. Generally the running time of algorithms based on low-treedepth colorings depends heavily on the number of colors $f(p)$. In Section 5 we present experimental results obtained by computing and evaluating low-treedepth colorings with a simple algorithm in a variety of real-world networks.

Algorithmic results. With both theoretical and empirical results in hand, we exploit the aforementioned tools for graphs with small grad in Section 6 to solve typical problems for complex networks: First we develop a faster algorithm than the one presented in [67] to count subgraph homomorphisms based on low treedepth colorings. Counting subgraphs is fundamental to *motif counting*, a widely used concept to analyze networks. Then we develop an algorithm which computes localized versions of several centrality measures in linear time on graphs of bounded expansion. Specifically, we present:

i. A linear-time algorithm to count the number of times a fixed subgraph appears as an (induced) subgraph/homomorphism in graphs of bounded grad. We do this by improving the previously best known algorithm to count the appearances of a structure of size h on graphs of treedepth t from $O(2^{ht}ht \cdot n)$ to $O(t^h 8^h h^2 \cdot n)$, thus removing the exponential dependency on t, while keeping the algorithm simple and avoiding big hidden constants.

ii. A linear-time algorithm to compute localized variants (i.e., computed in a constant-radius neighborhood around each vertex) of the *closeness centrality* and two other related measures. The constant in the running time depends on the radius and the grad of the graph.

For the second algorithm we provide experimental results which indicate that the local variants of these centrality measures can be used to estimate the top 10 percent of the most central nodes.

Previous results. There is substantial empirical work studying structural properties of complex networks, so we focus here on work closest to structural graph theory. In general, large real-world complex networks are not easily classified as either low- or high-dimensional. In particular, data-mining tools which implicitly assume low dimensionality (such as singular value decomposition) produce models and results incompatible with observed structure and dynamics, yet traditional high-dimensional tools (like sampling) often fail to achieve measure concentration due to the extreme sparsity of the networks. Adcock et al. [4] recently empirically established that, when compared with a suite of idealized graphs[2], realistic large social and informatics networks

_____

[2]representing low-dimensional structures, common hierarchical models, constant-degree expanders, etc.

exhibit meaningful "tree-like" structure at an intermediate scale. Their work related the k-core structure (whose extremal statistic is the degeneracy) to the networks' Gromov hyperbolicity and tree decompositions. Unfortunately, they showed that straightforward applications of these measures are often insufficient to reveal meaningful structure because of noisy/random edges in the network which contradict the strict structural requirements. (For example, several families of popular random graph models have been shown to have very large treewidth [34].)

Some simple preliminary observations about networks and bounded expansion were made in [32], such as the *linear growth model* not having bounded expansion since it was known that it contains growing bi-cliques, and they conjecture that Barabási-Albert is somewhere dense a.a.s. Here we prove that it is at least somewhere dense with non-vanishing probability. For the *random intersection graph* model, which is used to model real world networks where connections depend on shared attributes, it was shown in [28] that it has bounded expansion exactly when it is degenerate.

2 Preliminaries

For a natural number n we use the notation $[n] := \{1, \ldots, n\}$. For a graph G, we denote by $\Delta(G)$ its maximal degree and by $\omega(G)$ its clique number. We will make use of the following graph operations. For graphs G_1, G_2, the *complete join* $G_1 * G_2$ is the graph obtained by first taking the disjoint union of G_1, G_2 and then connecting every vertex of $V(G_1)$ to every vertex $V(G_2)$. For example, $G * \overline{K}_2$ is the graph obtained from G by adding two universal vertices. The *lexicographic product* $G_1 \bullet G_2$ is the graph with vertices $V(G_1) \times V(G_2)$ and edges

$$(u, x)(v, y) \in E(G_1 \bullet G_2) \iff uv \in E(G_1) \text{ or } (u = v \text{ and } xy \in E(G_2)).$$

2.1 Graphs of bounded expansion

We will use $(\leq r)$-*subdivisions* to formalize the notion of shallow topological minors. A $(\leq r)$-subdivision of a graph M is any graph which can be obtained from M by replacing edges with disjoint paths of length at most $r + 1$.

Definition 1 (Shallow topological minor [65]) *A graph M is an r-shallow topological minor if a $(\leq 2r)$-subdivision of M is isomorphic to a subgraph G' of 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 $G \, \widetilde{\triangledown} \, r$.

We will also need an associated density measure for shallow topological minors:

Definition 2 (Topological grad) *For a graph G and an integer $r \geq 0$, the topological greatest reduced average density (top. grad) at depth r is defined as*

$$\widetilde{\nabla}_r(G) = \max_{H \in G \, \widetilde{\triangledown} \, r} \frac{|E(H)|}{|V(H)|}.$$

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

We can now define what it means for a class to have bounded expansion.

Definition 3 (Bounded expansion) *A graph class $\mathcal{G}$ has bounded expansion if and only if there exists a function f such that for all $r \geq 0$, we have $\widetilde{\nabla}_r(\mathcal{G}) < f(r)$.*

When introduced in [65], bounded expansion was originally defined using a characterization based on the notion of *shallow minors*: H is an r-shallow minor of G if H can be obtained from G by contracting disjoint r-balls and then taking a subgraph. Taking the maximum over the density of all r-shallow minors then defines the *grad*. An r-ball in a graph G is a subgraph $G' \subseteq G$ with the property that there exists $v \in V(G')$ such that for all $u \in V(G')$, $d_{G'}(u, v) \leq r$. Both characterizations are equivalent, hence we will for brevity's sake simply talk about the grad.

We note that graphs excluding a topological minor—in particular planar graphs and bounded-degree graphs—have bounded expansion. This generalizes to graphs excluding a minor (and thus to those of bounded treewidth). Finally, we point out that bounded expansion implies bounded degeneracy, where a graph G is d-degenerate if any subgraph of G contains a node of degree smaller than d. The converse does not hold.

The following alternative characterization of bounded expansion uses a special coloring number with nice algorithmic properties.

Definition 4 (p-centered coloring [65]) *Given a graph G, let $c \colon V(G) \to [r]$ be a vertex coloring of G with r colors. We say that the coloring c is p-centered, for $p \geq 2$, if any connected subgraph of G either receives at least p colors or contains some color exactly once. Define $\chi_p(G)$ to be the minimum number of colors needed for a $(p+1)$-centered coloring.*

While this definition looks rather cryptic, it is easy to see that every graph has a p-centered coloring for any p: simply assign a distinct color to each vertex of the graph. Note that p-centered colorings are proper colorings for $p \geq 2$ and in particular, χ_1 is precisely the chromatic number. Typically, the number of colors q is much larger than p and one is interested in minimizing q.

The following structural property of p-centered colorings make them an attractive tool for algorithm design. Suppose that c is indeed a p-centered coloring of G using r colors. Then any $i < p$ color classes induce a graph of *treedepth* at most i. For those unfamiliar with this notion, the *treedepth* of a graph H is the lowest rooted forest whose closure contains G as a subgraph. A *treedepth decomposition* of H is simply a forest with vertex set $V(H)$ witnessing this fact. To put this width measure into perspective: a graph of treedepth at most t cannot contain a path of length 2^t *and* has pathwidth at most $t-1$. An example of a 5-centered coloring and treedepth-decompositions of a selected subset of colors can be found in Figure 2 on page 36.

Nešetřil and Ossona de Mendez show that graph classes of bounded expansion are precisely those for which there exists a function f such that every member G of the graph class satisfies $\chi_p(G) \leq f(p)$ (see Theorem 7.1 in [65]). In [66], the authors also showed how to obtain a p-centered coloring with at most $P(f(p))$ colors for each fixed p in linear time, where P is some polynomial of degree roughly 2^{2^p}. We will make use of this algorithm in Sections 5 and 6 and see that the actual number of colors is manageable.

When working with random graphs, we will make heavy use of the following alternative characterization of graphs of bounded expansion:

Proposition 1 ([67,68]) *A class $\mathcal{C}$ of graphs has bounded expansion if and only if there exist real-valued functions $f_1, f_2, f_3, f_4 \colon \mathbf{R}^+ \to \mathbf{R}$ such that the following two conditions hold:*

i. *For all $\varepsilon > 0$ and for all $G \in \mathcal{C}$ with $|G| > f_1(\varepsilon)$, it holds that*

$$\frac{1}{|V(G)|} \cdot |\{v \in V(G) \colon \deg(v) \geq f_2(\varepsilon)\}| \leq \varepsilon.$$

ii. *For all $r \in \mathbf{N}$ and for all $H \subseteq G \in \mathcal{C}$ with $\widetilde{\nabla}_r(H) > f_3(r)$, it follows that*

$$|V(H)| \geq f_4(r) \cdot |V(G)|.$$

Intuitively, Proposition 1 characterizes classes of graphs with bounded expansion as those where:

i. all sufficiently large members of the class have a small fraction of vertices of large degree;

7

ii. all subgraphs of $G \in \mathcal{C}$ whose shallow topological minors are sufficiently dense must necessarily span a large fraction of the vertices of G.

Nowhere dense is a generalization of bounded expansion in which we measure the *clique number* instead of the edge density of shallow minors. See [67] for a definition and many equivalent notions. In particular, a class is *somewhere dense* if for some r there exist arbitrarily large cliques as r-shallow (topological) minors of the members of that class.

2.2   Random graph models

A *random graph model* is a sequence $\mathscr{G} := (\mathscr{G}_n)_{n \in \mathbf{N}}$ of random variables $\mathscr{G}_n$ defined over the set of all n-vertex graphs. For convenience, we assume that the vertex set of $\mathscr{G}_n$ is $[n]$. Since models usually come with tunable parameters—like the edge probability in the Erdős–Rényi model—we actually have such a sequence for every possible choice of parameters (which can depend on n).

Extending the notation of the previous section, for a random graph $\mathscr{G}_n$ and an integer r the notation $\mathscr{G}_n \widetilde{\triangledown} r$ denotes a random variable over sets of graphs with at most n vertices whose probability distribution is given by

$$\mathbf{Pr}\left[\mathscr{G}_n \widetilde{\triangledown} r = A\right] = \sum_{G : A = G \widetilde{\triangledown} r} \mathbf{Pr}\left[\mathscr{G}_n = G\right]$$

where A is a set of graphs. As usual, we study the properties of random graphs in the limit.

Definition 5 *A graph model $\mathscr{G}_n$ has* bounded expansion asymptotically almost surely (a.a.s.) *if there exists a function f such that for every $r \geq 0$*

$$\lim_{n \to \infty} \mathbf{Pr}\left[\widetilde{\nabla}_r(\mathscr{G}_n) < f(r)\right] = 1.$$

It has bounded expansion with high probability (w.h.p.) *if for every $c \geq 1$ there exists a function f such that*

$$\mathbf{Pr}\left[\widetilde{\nabla}_r(\mathscr{G}_n) < f(r)\right] \geq 1 - O(n^{-c}).$$

The following notions are needed to prove negative results about random models.

Definition 6 *A graph model $\mathscr{G}_n$ is* a.a.s. somewhere dense *if there exists $r \in \mathbf{N}$ such that for all functions f it holds that*

$$\lim_{n \to \infty} \mathbf{Pr}\left[\omega(\mathscr{G}_n \widetilde{\triangledown} r) > f(r)\right] = 1.$$

It is not a.a.s. nowhere dense *if there exists* $r \in \mathbf{N}$ *such that for all functions f it holds that*

$$\lim_{n\to\infty} \mathbf{Pr}\left[\omega(\mathscr{G}_n \,\widetilde{\nabla}\, r) > f(r)\right] > 0.$$

Where $\omega(\mathscr{G}_n \,\widetilde{\nabla}\, r) = \max_{H \in \mathscr{G}_n \,\widetilde{\nabla}\, r} \omega(H)$.

In particular, a graph model that is a.a.s. somewhere dense has unbounded expansion.

3 Graph Models of Bounded Expansion

Often, analytic methods are selected based on their behavior when applied to random graphs that are believed to mimic characteristics of the particular networks under consideration. We will not digress into the arguments for and against this methodology, but simply note that it is a *de facto* part of standard practice at this point in time. Accordingly, we must be able to establish whether or not graphs generated by such models have bounded grad. For more information on random graph models, we refer the readers to the surveys in [63, 64, 72, 74]. In this section, we determine whether several such models have bounded expansion (as by Definition 5.)

We show that the popular configuration model [63, 64], including the version with households exhibiting high clustering [8], and the Chung–Lu model [19, 20], has bounded expansion w.h.p. for interesting degree distributions (cf. Table 1 for examples.) Prior work [68] has shown that the Erdős–Rényi model has bounded expansion a.a.s. Unfortunately, empirical analysis of real-world networks (including friendships/social networks, telephone/communication networks, and biological/neural networks) has shown that typical degree distributions are measurably different from the Poisson distribution exhibited by Erdős–Rényi graphs(see [74] and the references therein.) We extend the results of Nešetřil et al. to a model which allows the inclusion of an arbitrary bounded degree graph in addition to probabilistically generated edges (with non-identical probabilities, subject to a uniform bound) and show this model to have bounded expansion with high probability. This in particular strengthens the aforementioned result on the Erdős–Rényi model in terms of speed of convergence. Including a base graph and allowing non-uniform edge probabilities drastically increases the structural variability in the model's output. We will further argue that using graphs with unbounded degree as the base graph will necessarily generate structurally dense graphs; in that sense our result is tight. Finally, we comment that this model includes as a special case certain types of stochastic block models.

3.1 The positive toolkit

In the following we identify a useful—albeit technical—property of random graph models which helps us to apply Proposition 1. In essence, we show that a strong enough bound for the probability that some random vertices form a dense shallow topological minor suffices to show that no such minor appears with high probability. Proving that such a model has bounded expansion then boils down to showing that it does not have too many high-degree vertices.

Definition 7 *Let $\alpha, \beta : \mathbf{N} \to \mathbf{R}$ be functions. A random graph model $\mathscr{G}$ is (α, β)-sparse if there exists constants n_0, c_0 and ξ_0 such that for every $n > n_0$, every $k < n/c_0$ and every $r \in \mathbf{N}$, the probability that k distinct vertices chosen uniformly at random from $\mathscr{G}_n$ are connected by at least $\xi k > \xi_0 k$ internally vertex-disjoint paths each of length at most $2r+1$ is bounded by*

$$O\left(\binom{\binom{k}{2}}{\xi k} \left(\frac{\alpha(r)}{n - \beta(r)\xi k} \right)^{\xi k} \right).$$

The following lemma shows that such graph models do not contain dense shallow topological minors with high probability.

Lemma 1 *Let $\mathscr{G}$ be an (α, β)-sparse random graph model with constants n_0, c_0 and ξ_0. Then for every $\xi \geq \max\{\xi_0, (1+\sqrt{5})/2\}$ and every $r \in \mathbf{N}$ there exist constants n_1, c_1 such that for any r-shallow topological minor H of $\mathscr{G}_n, n > n_1$ with at most $c_1 \cdot n$ vertices we have that*

$$\mathbf{Pr}\left[\frac{|E(H)|}{|V(H)|} > \xi \right] \leq 1/n^\xi.$$

Proof: Using the terminology from Definition 1, let $k := |V(H)|$ so that the model of H (which is a subgraph of $\mathscr{G}_n$) has k nails and at most $k \cdot \xi \cdot (2r+1)$ subdivision points.

The probability that we wish to estimate is simply the probability that a set of $k \leq n/c$ vertices, where c is some constant to be fixed later, chosen uniformly at random have at least ξk paths of length at most $2r+1$ between them in $\mathscr{G}_n$. By the properties of $\mathscr{G}$, this probability is bounded by:

$$\sum_{k=1}^{n/c} \binom{n}{k} \binom{\binom{k}{2}}{\xi k} \left(\frac{\alpha(r)}{n - \beta(r)\xi k} \right)^{\xi k}, \tag{1}$$

The expression in the sum of (1) is bounded from above by:

$$\left(\frac{ne}{k} \right)^k \left(\frac{ke}{\xi} \right)^{\xi k} \left(\frac{\alpha(r)}{n - \beta(r)\xi k} \right)^{\xi k} = \left(e \cdot \left(\frac{e\alpha(r)}{\xi(1 - \beta(r)\xi k/n)} \right)^\xi \left(\frac{k}{n} \right)^{\xi-1} \right)^k. \tag{2}$$

Choosing $c \geq 2\beta(r)\xi$ we see that $\beta(r)\xi k/n \leq \beta(r)\xi f_4(r) \leq 1/2$ and therefore (2) is bounded by

$$\left(e \cdot \left(\frac{2e\alpha(r)}{\xi}\right)^{\xi} \cdot \left(\frac{k}{n}\right)^{\xi-1}\right)^k \leq \left(\gamma(r,\xi) \cdot \left(\frac{k}{n}\right)^{\xi-1}\right)^k \tag{3}$$

where $\gamma(r,\xi)$ denotes the expression $e(2e\alpha(r)/\xi)^{\xi}$.

Notice that an r-shallow topological minor with k nails cannot have a density larger than $(k-1)/2$. Therefore, we sharpen the bound given in Equation 1 by summing from $2\xi+1$ instead of 1. We claim that

$$\sum_{k=2\xi+1}^{n/c} \left(\gamma(r,\xi) \cdot \left(\frac{k}{n}\right)^{\xi-1}\right)^k \leq \frac{1}{n^{\xi}} \tag{4}$$

The first term in this sum is

$$\left(\gamma(r,\xi) \cdot \left(\frac{2\xi+1}{n}\right)^{\xi-1}\right)^{2\xi+1} = \frac{\gamma(r,\xi)^{2\xi+1}(2\xi+1)^{2\xi^2-\xi-1}}{n^{2\xi^2-\xi-1}} \leq \frac{1}{2n^{\xi}},$$

where the last inequality holds for $n > 2\gamma(r,\xi)^{2\xi+1}(2\xi+1)^{2\xi^2-\xi-1}$ and $\xi \geq (1+\sqrt{5})/2$. The remaining terms decrease at least geometrically by a factor of

$$\frac{\left(\gamma(r,\xi)((k-1)/n)^{\xi-1}\right)^{k-1}}{(\gamma(r,\xi)(k/n)^{\xi-1})^k} = \frac{n^{\xi-1}}{\gamma(r,\xi)k^{\xi-1}}\left(\frac{k-1}{k}\right)^{k-1}$$

$$\geq \frac{n^{\xi-1}}{\gamma(r,\xi)k^{\xi-1}} \cdot \frac{1}{e} \geq \frac{n^{\xi-1}}{\gamma(r,\xi)(n/c)^{\xi-1}} \cdot \frac{1}{e}$$

$$= \frac{c^{\xi-1}}{e\gamma(r,\xi)}.$$

For $c^{\xi-1} \geq 2e\gamma(r,\xi)$, this factor is at least 2. Hence the overall sum is at most twice the first term and therefore in total at most $1/n^{\xi}$, as claimed. This proves that

$$\mathbf{Pr}\left[\frac{|E(H)|}{|V(H)|} > \xi\right] \leq 1/n^{\xi},$$

for any $\xi \geq (1+\sqrt{5})/2$, setting $c_1 = \max\{c_0, (2e\gamma(r,\xi))^{1/(\xi-1)}, 2\beta(r)\xi\}$ and $n_1 = 2\gamma(r,\xi)^{2\xi+1}(2\xi+1)^{2\xi^2-\xi-1}$. $\qquad\square$

In order to make the rather technical definition of (α,β)-sparse random graph models more applicable, we prove the following sufficient condition: a model is (α,β)-sparse if a fixed set of q edges is present with probability proportional to n^{-q}. This property is useful for graph models in which edge probabilities are not independent.

Lemma 2 *Let $\mathcal{G}$ be a random graph model with the following property: There exists a constant $c > 1$ such that for every edge set $F \subseteq [n] \times [n]$, $|F| \le n/c$, chosen uniformly at random it holds that:*

$$\mathbf{Pr}\left[F \subseteq E(\mathcal{G}_n)\right] = O((\theta/n)^{|F|}),$$

where θ is a constant that depends on $\mathcal{G}$. Then $\mathcal{G}$ is $((2r+1)e\theta^{2r+1}, 2r)$-sparse for all $r \in \mathbf{N}$.

Proof: Suppose that k randomly chosen vertices had q vertex-disjoint paths between them of lengths $l_1, \ldots, l_q$ such that $1 \le l_i \le 2r + 1$ for all $1 \le i \le q$. For a path of length l_i to exist, there exist $l_i - 1$ internal vertices that are nodes in a path between two of the k vertices chosen uniformly at random. By the properties of $\mathcal{G}$, the probability that there exist q vertex-disjoint paths between k randomly chosen vertices is given by:

$$\binom{\binom{k}{2}}{q} \cdot \sum_{i:\, 1 \le l_i \le 2r+1} O\left(\left(\frac{\theta}{n}\right)^{\sum_{i=1}^{q} l_i}\right) \cdot T(n; l_1, \ldots, l_q), \qquad (5)$$

where we use $T(n; l_1, \ldots, l_q)$ to denote:

$$\binom{n}{l_1 - 1}(l_1 - 1)!\binom{n - l_1 + 1}{l_2 - 1}(l_2 - 1)! \cdots \binom{n - \sum_{i=1}^{q-1}(l_i - 1)}{l_q - 1}(l_q - 1)!.$$

Now $T(n; l_1, \ldots, l_q)$ simplifies to

$$\frac{n!}{(n + q - \sum_{i=1}^{q} l_i)!}.$$

Using the bound $(n/e)^n < n! < n(n/e)^n$, one can upper-bound the above as follows:

$$\frac{n!}{(n + q - \sum_{i=1}^{q} l_i)!} \le \frac{n^{n+1} \cdot e^q}{(n + q - \sum_{i=1}^{q} l_i)^{\left(n + q - \sum_{i=1}^{q} l_i\right)}}$$

$$\le \frac{n^{n+1} \cdot e^q \cdot (n + q)^{(2r+1)q}}{(n + q - q(2r+1))^{n+q}}$$

$$\le \frac{n^{n+1} \cdot e^q \cdot (n + q)^{(2r+1)q}}{(n - 2rq)^{n+q}}. \qquad (6)$$

Using (6), we can upper-bound expression (5) by:

$$\binom{\binom{k}{2}}{q} \cdot (2r + 1)^q \cdot \frac{n^{n+1} \cdot e^q \cdot (n + q)^{(2r+1)q}}{(n - 2rq)^{n+q}} \cdot \left(\frac{\theta}{n}\right)^{q(2r+1)}$$

which in turn is at most

$$\binom{\binom{k}{2}}{q} \cdot \left(\frac{(2r+1) \cdot e \cdot \theta^{2r+1}}{n - 2rq}\right)^q \cdot \left(1 + \frac{q}{n}\right)^{q(2r+1)} \cdot \frac{n}{\left(1 - \frac{2rq}{n}\right)^n},$$

which is

$$O\left(\binom{\binom{k}{2}}{q} \cdot \left(\frac{(2r+1) \cdot e\theta^{2r+1}}{n - 2rq}\right)^q\right). \tag{7}$$

$\square$

Finally, we note the following well-known property of grads: the grad of a graph only changes slightly if we add a universal vertex to it. We provide a proof for completeness.

Lemma 3 *For every graph G it holds that*

$$\frac{\widetilde{\nabla}_r(G)}{\widetilde{\nabla}_r(G) + 1}(\widetilde{\nabla}_r(G) + 1) < \widetilde{\nabla}_r(G * K_1) < \widetilde{\nabla}_r(G) + 1$$

Proof: It is easy to see that if H is an r-shallow topological minor of G, then $H * K_1$ is a r-shallow topological minor of $G * K_1$. The density of $H * K_1$ is then given by

$$\frac{|E(H * K_1)|}{|V(H * K_1)|} = \frac{|E(H)| + |V(H)|}{|V(H)| + 1}$$

And therefore we obtain upper and lower bounds via

$$\frac{1}{1 + (1/|V(H)|)}\left(\frac{|E(H)|}{|V(H)|} + 1\right) \leq \frac{|E(H * K_1)|}{|V(H * K_1)|} < \frac{|E(H)|}{|V(H)|} + 1$$

Observing that $|V(H)| \geq \widetilde{\nabla}_r(G)$ proves the claim. $\square$

Therefore even a constant number of universal vertices will not influence the grad too much. In terms of graph classes this means that if $\mathcal{G}$ has bounded expansion, then the class $\{G * K_c\}_{G \in \mathcal{G}}$ for any constant c also has bounded expansion.

3.2 The Configuration Model

Since empirical measurements of real-world graphs typically reveal non-Poisson degree distributions (in particular, they are often heavy tailed), one might wish to consider a random graph model which generates graphs with a given degree sequence uniformly at random. The goal of this section is to show that

asymptotically, random graphs with a degree sequence that is well-behaved in a certain sense, have bounded expansion.

The configuration model was introduced by Molloy and Reed [63] who also credit Bender and Canfield [10]. It was further refined by Bollobás [11] and Wormald [91], and has been discussed extensively in the literature [64, 71, 72, 74].

In order to consider the question of bounded expansion, we must first define what it means for an n-vertex random graph to realize a fixed degree sequence. The formalization we give in Definitions 8–9 appears in [63].

Definition 8 (Asymptotic degree sequence) *An asymptotic degree sequence is an infinite sequence of integer-valued functions $\mathcal{D} = d_1, d_2, \ldots$ such that*

i) $d_i(n) = 0$ for all $i \geq n$;

ii) $\sum_{i=0}^{n-1} d_i(n) = n$.

In other words, the function $d_i(n)$ denotes the number of vertices of degree exactly i in an n-vertex graph. Condition (i) then simply states that no vertex in an n-vertex graph has degree n or more. Condition (ii) states that the sum over the number of vertices for each degree i equals the total number of vertices.

Given an asymptotic degree sequence $\mathcal{D}$ and an integer n we write $\mathcal{D}_n := d_1(n), d_2(n), \ldots, d_{n-1}(n)$ to denote the *concrete* degree-sequence for an n-vertex graph prescribed by $\mathcal{D}$. Accordingly, let $\Omega_{\mathcal{D}_n}$ denote the set of all graphs with vertex set $\{1, \ldots, n\}$ and degree sequence $\mathcal{D}_n$. Since we will be using asymptotic properties of random graphs with degree sequence $\mathcal{D}$, we want the sequences $\mathcal{D}_n$ to be similar in the sense that the fraction of vertices of degree i is roughly the same for every n.

Definition 9 *An asymptotic degree sequence $\mathcal{D}$ is*

i) feasible if $\Omega_{\mathcal{D}_n} \neq \varnothing$ for all $n \geq 1$.

ii) smooth if there exist constants λ_i such that $\lim_{n \to \infty} d_i(n)/n = \lambda_i$.

iii) sparse if there exists a constant $\mu \in \mathbf{R}^+$ such that, $\sum_{i=0}^{n-1} i \cdot d_i(n)/n = \mu + o(1)$, and $\sum_{i=0}^{\infty} i\lambda_i = \mu$.

We say that $\mathcal{D}$ converges to a (real-valued) function f if there exists a constant c such that for all $i \geq c$ it holds that $\lambda_i = f(i)$.

One major hurdle in studying random graphs on fixed degree sequences is the difficulty to generate such graphs directly. The configuration model addresses this problem by generating a *multigraph* with the correct degree sequence. As we will see later, for sparse degree sequences the probability of obtaining a simple graph is sufficiently high.

The model intuitively works as follows: One first generates n vertices with "half-edges" as prescribed by the degree sequence. Then these half-edges are randomly and independently connected in pairs to form full edges. The process might introduce some loops (when two half-edges belonging to the same vertex are paired up) and parallel edges (when two pairs of half-edges belonging to the same two vertices are paired up.) Such edges must be removed to obtain a simple graph. Consequently, the degree sequence will only be fully realized in the cases where the configuration outputs a simple graph directly.

Formally, given a degree sequence $\mathcal{D}$, a *random configuration with n vertices on $\mathcal{D}$* is generated as follows:

1. Create a a sequence $(\delta_i)_{1 \leq i \leq n}$ of integers such that $\sum_{i=1}^{n} [\delta_i = k] = d_k(n)$ for all $1 \leq k \leq n$ (where $[\cdot]$ are Iverson brackets), i.e. for each degree k the number k will appear exactly $d_k(n)$ times.

2. Create a set L containing δ_v distinct copies of each vertex $v \in \{1, \ldots, n\}$ denoted by $v_1, \ldots, v_{\delta_v}$.

3. Choose a random matching of the elements of L.

We adopt the following notation: for each vertex $v \in \{1, \ldots, n\}$ we write $\deg(v) = \delta_v$. Now it is easy to see that a graph with the degree sequence $\delta_1, \ldots, \delta_n$ is represented by a configuration whose matching obeys following properties:

i) no two vertices v_i, v_j for $1 \leq i, j \leq \deg(v)$ are matched,

ii) for all vertices $u, v \in \{1, \ldots, n\}$ at most one pair u_i, v_j for $1 \leq i \leq \deg(u)$ and $1 \leq j \leq \deg(v)$ is matched.

Conversely, any matching in the configuration model that satisfies the above two properties corresponds to a simple graph with the degree sequence $\delta_1, \ldots, \delta_n$ and hence has the degree distribution $\mathcal{D}$.

In the following, let $\mathscr{C}(\mathcal{D})$ be the random graph model that generates graphs in this fashion (including removing loops and parallel edges) according to the asymptotic degree sequence $\mathcal{D}$.

We will first show that graphs taken from $\mathscr{C}(\mathcal{D})$, where $\mathcal{D}$ is a well-behaved degree sequence, have bounded expansion w.h.p. We can then apply known

results to extend this to be a.a.s. true for the case were we sample uniformly at random from $\Omega_{\mathcal{D}_n}$.

We can now state our main technical result:

Theorem 1 *Let $\mathcal{D}$ be a feasible, smooth and sparse asymptotic degree sequence. Then $\mathscr{C}(\mathcal{D})$ has bounded expansion with high probability.*

Proof: We verify the two properties stated in Proposition 1. Property (i) directly follows from the sparseness and smoothness-conditions: suitable functions f_1, f_2 are implied by the condition that $\sum_{i=0}^{n-1} i \cdot d_i(n)/n = \mu + o(1)$ with the limit $\sum_{i=0}^{\infty} i\lambda_i = \mu$. Combining Lemma 4 with Lemma 2 yields the necessary condition to apply Lemma 1, which in turn proves that Property (ii) holds in this setting and thus proves the theorem. □

Lemma 4 *Let $\mathcal{D}$ be a feasible, smooth and sparse asymptotic degree sequence. Fix a constant $c > 1$ and let $F \subseteq [n] \times [n]$ be a set of at most n/c edges chosen uniformly at random. Then it holds that:*

$$\mathbf{Pr}\left[F \subseteq E(\mathscr{C}_n(\mathcal{D}))\right] = O((\theta/n)^{|F|}),$$

where θ is a constant that depends on the sequence $\mathcal{D}$.

Proof: Fix an integer n and create a set of $2m = \sum_{i=0}^{n-1} i \cdot d_i(n)$ vertices for the configuration model. Denote by $\mathscr{M}_n(\mathcal{D})$ a random perfect matching on the vertices of the configuration model. Let F be a set of $k = \lfloor n/c \rfloor$ vertex pairs $\{u_1 v_1, \ldots, u_k v_k\}$ in the configuration model such that no vertex occurs more than once. As such no two potential edges share a vertex. Partition the vertices in F into sets L and R such that every potential edge has exactly one vertex in each of these sets. For calculating the probability that F appears in a random matching, we need only consider the degrees of the endpoints. Therefore for $1 \leq i \leq k$, let $\deg(u_i) = l_i$ and $\deg(v_i) = r_i$ where we assumed that the vertices $u_1, \ldots, u_k$ are in the left set L while the remaining vertices are in R.

We first want to estimate the probability that F appears as an edge set in $\mathscr{M}_n(\mathcal{D})$ under the condition that the degrees of its endpoints is given by the fixed sequence $(l_1, r_1), \ldots, (l_k, r_k)$. We denote this probability by

$$\rho := \mathbf{Pr}\left[F \subseteq \mathscr{M}_n(\mathcal{D}) \,\middle|\, \deg(F) = \bigwedge_{i=1}^{k} (l_i, r_i)\right], \tag{8}$$

where $\deg(F) = \bigwedge_{i=1}^{k}(l_i, r_i)$ represents the event that a randomly chosen matching F of size k has endpoints with degrees (l_i, r_i), $1 \leq i \leq k$. If $e_i \in F$

has endpoints with degrees l_i and r_i, then it can be chosen in $l_i \cdot r_i$ ways. There are $\binom{2m}{m} \cdot m!$ matchings in total and of these there are

$$\prod_{i=1}^{k}(l_i \cdot r_i) \cdot \binom{2m-2k}{m-k} \cdot (m-k)!$$

matchings that generate F. Therefore the expression for ρ is:

$$\rho = \frac{\prod_{i=1}^{k}(l_i \cdot r_i) \cdot \binom{2m-2k}{m-k} \cdot (m-k)!}{\binom{2m}{m} \cdot m!}.$$

Since

$$\mathbf{Pr}\left[F \subseteq \mathscr{M}_n(\mathcal{D}) \wedge \deg(F) = \bigwedge_{i=1}^{k}(l_i, r_i)\right] = \rho \cdot \mathbf{Pr}\left[\deg(F) = \bigwedge_{i=1}^{k}(l_i, r_i)\right],$$

we next consider the second term on the right hand side. This term expresses the probability that a randomly chosen matching with edges $e_1, \ldots, e_k$ is such that edge e_i has as its endpoints with degrees l_i and r_i. A simple upper bound is the following:

$$\mathbf{Pr}\left[\deg(F) = \bigwedge_{i=1}^{k}(l_i, r_i)\right] \le \prod_{i=1}^{k} \frac{d_{l_i}(n) \cdot d_{r_i}(n)}{(n-2k)^2}.$$

Recall that by the sparseness-condition of the degree sequence $\mathcal{D}$, we have that $\sum_{i=0}^{n-1} i \cdot d_i(n)/n = \mu + o(1) := \alpha$, from which it follows that $2m = \alpha n$. Using the union bound, we sum over all possible choices of degrees $l_1, \ldots, l_k, r_1, \ldots, r_k$ to upper bound $\mathbf{Pr}\left[F \subset \mathscr{M}_n(\mathcal{D})\right]$ by:

$$\sum_{(l_1, r_1), \ldots, (l_k, r_k)} \frac{\prod_{i=1}^{k}(l_i \cdot r_i) \cdot \binom{2m-2k}{m-k} \cdot (m-k)!}{\binom{2m}{m} \cdot m!} \cdot \frac{\prod_{i=1}^{k} d_{l_i}(n) \cdot d_{r_i}(n)}{(n-2k)^{2k}}, \quad (9)$$

which simplifies to

$$\frac{m!(2m-2k)!}{(2m)!(m-k)!} \cdot \frac{1}{(n-2k)^{2k}} \cdot \sum_{(l_i, r_i)} \prod_{i=1}^{k}(l_i d_{l_i}(n) \cdot r_i d_{r_i}(n)). \quad (10)$$

Let us next consider the sum-of-products expression from (10) which may be written as:

$$\sum_{l_1} \cdots \sum_{l_k} \sum_{r_1} \cdots \sum_{r_k} l_1 d_{l_1}(n) \cdots l_k d_{l_k}(n) \cdot r_1 d_{r_1}(n) \cdots r_k d_{r_k}(n).$$

17

Since $2m = \sum_{i=0}^{n-1} i d_i(n) = \alpha n$, we may upper-bound the above expression by $(2m)^{2k}$. Using the bound $(n/e)^n < n! < n(n/e)^n$, we thus upper-bound expression (10) by:

$$m \cdot (2m - 2k) \cdot \left(\frac{e}{4}\right)^k \cdot \frac{(m-k)^{m-k}}{m^m} \cdot \frac{1}{(n-2k)^{2k}} \cdot (2m)^{2k}, \qquad (11)$$

which simplifies to

$$2m^2 \cdot e^k \cdot \left(1 - \frac{k}{m}\right)^{m-k} \cdot \left(\frac{m}{n^2 \left(1 - 2k/n\right)^2}\right)^k. \qquad (12)$$

Writing m as αn, gives us:

$$2m^2 \cdot \left(1 - \frac{k}{m}\right)^{m-k} \cdot \left(\frac{\alpha e}{n \left(1 - 2k/n\right)^2}\right)^k = O\left((\theta/n)^k\right), \qquad (13)$$

where we defined θ to be αe. $\qquad \square$

We have now proven that the Configuration Model has bounded expansion w.h.p. for asymptotic degree sequences that are feasible, smooth and sparse. We want to transfer this result to the model which choses simple graphs uniformly at random from Ω_{D_n} for every n. Then the following known result gives a sufficient condition to do so. Let $\mathscr{C}'_n(\mathcal{D})$ denote the random *multigraphs* generated by the configuration model on n nodes.

Proposition 2 ([41]**)** *Let $\mathcal{D}$ be an asymptotic degree sequence. Then*

$$\liminf_{n \to \infty} \mathbf{Pr}\left[\mathscr{C}'_n(\mathcal{D}) \text{ is simple}\right] > 0 \iff \sum_{i \geq 0} i^2 d_i(n) = O\left(\sum_{i \geq 0} i \cdot d_i(n)\right).$$

There exist feasible, smooth and sparse asymptotic degree sequences that converge against the functions in Table 1. Any such sequence will also fulfill the condition stated in Proposition 2, giving the main result of this section.

Theorem 2 *Let $\mathcal{D}$ be a feasible, smooth and sparse asymptotic degree sequence that converges to a normalized version of a function from Table 1 with parameters chosen such that the distribution has finite variance. The random graph model defined by drawing uniformly at random from Ω_{D_n} has bounded expansion a.a.s.*

It is interesting to point out that there is a direct relation between graphs with a power-law degree distribution and scale-free graphs. If we disregard low-degree nodes, sparse graphs are scale-free precisely when their degree distribution follows a power-law [22]. This is not the case if one takes into account low-degree nodes as well [85]. Thus if we make no further assumptions about networks other than scale-freeness, we would expect them to have constant grad.

Name	Definition $f(d)$	Parameters
Power law	$d^{-\gamma}$	$\gamma > 2$
Power law w/ cutoff	$d^{-\gamma}e^{-\lambda d}$	$\gamma > 2, \lambda > 0$
Exponential	$e^{-\lambda d}$	$\lambda > 0$
Stretched exponential	$d^{\beta-1}e^{-\lambda d^{\beta}}$	$\lambda, \beta > 0$
Gaussian	$\exp\left(-\frac{(d-\mu)^2}{2\sigma^2}\right)$	μ, σ
Log-normal	$d^{-1}\exp\left(-\frac{(\log d-\mu)^2}{2\sigma^2}\right)$	μ, σ

Table 1: A selection of established functions used to model degree distributions of complex networks, listed without the necessary normalization factors. Here $f(d)$ is the fraction of nodes which have degree d. These functions were taken from an empirical analysis of degree distributions in real-world networks [21].

3.2.1 The Configuration Model with Households

It has been experimentally established that real-world networks tend to have a constant *average clustering coefficient*, which was defined for a graph G by Watts and Strogatz [90] as

$$C(G) = \frac{1}{n} \sum_{v \in V(G)} C(v) = \frac{1}{n} \sum_{v \in V(G)} \frac{\lambda(v)}{\tau(v)}$$

where $\lambda(v)$ denotes the number of triangles in G that contain v and $\tau(v)$ the number of induced P_3's in which v is the middle vertex.

Graphs generated with the configuration have generally a low clustering coefficient. The *configuration model with household structure* as defined by Ball et al. [8] is a model designed to address this issue, attempting to achieve a given degree distribution and a high clustering coefficient simultaneously. For this variant, one samples a graph with a prescribed degree sequence (that is smooth, feasible and sparse) and then replaces every vertex by a constant-sized "household"-graph, distributing the edges incident to a household graph uniformly to the vertices that form the household in the final graph. It is simple to check that the resulting graph has constant clustering.

If G is the inter-household graph, then the final graph is a subgraph of $G \bullet K_t$, where t is the household size. From Theorem 1 and 3 and the stability of grads under lexicographic products we arrive at the following corollary.

Corollary 1 *The configuration model with households has bounded expansion w.h.p.*

Further insights regarding bounded expansion graph classes and the clustering coefficient can be found in the Appendix.

3.3   The Chung–Lu model

A model closely related to the Configuration Model was proposed (and named after) Chung and Lu [19, 20]. Given a asymptotic degree sequence $\mathcal{D}$, one samples a graph by prescribing per-edge probabilities in such a way that the *expected* degree distribution matches $\mathcal{D}_n$. More precisely, two vertices with expected degrees δ_1 and δ_2 in an n-vertex random graph will be connected by an edge with probability

$$\frac{\delta_1 \delta_2}{\sum_{i=0}^{n-1} \delta_i} = \frac{\delta_1 \delta_2}{\sum_{i=0}^{n-1} i \cdot d_i(n)}$$

In that sense, the model corresponds to inhomogeneous random graphs (for a definition see [12]) with vertices weighted by their expected degree. For an asymptotic degree-sequence $\mathcal{D}$, we denote by $\mathscr{X}(\mathcal{D})$ the random graph model with edge probabilities prescribed as above.

Theorem 3 *Let $\mathcal{D}$ be a smooth and sparse asymptotic degree sequence. Then $\mathscr{X}(\mathcal{D})$ has bounded expansion w.h.p.*

Proof: We again verify the two properties stated in Proposition 1. As in the proof of Theorem 1 Condition (i) directly follows from the sparseness and smoothness-conditions. Combining Lemma 5 with Lemma 2 yields the necessary condition to apply Lemma 1, which in turn proves that Condition (ii) holds in this setting. □

Lemma 5 *Let $\mathcal{D}$ be a smooth and sparse asymptotic degree sequence. Fix a constant $c > 1$ and let $F \subseteq [n] \times [n]$ be a set of at most n/c edges chosen uniformly at random. Then it holds that:*

$$\mathbf{Pr}\left[F \subseteq E(\mathscr{X}_n(\mathcal{D}))\right] = O((\theta/n)^{|F|}),$$

where θ is a constant that depends on the sequence $\mathcal{D}$.

Proof: We first want to estimate the probability that F appears as an edge set in $\mathscr{X}_n(\mathcal{D})$ under the condition that the degrees of its endpoints is given by the fixed sequence $(l_1, r_1), \ldots, (l_k, r_k)$. We denote this probability by:

$$\rho := \mathbf{Pr}\left[F \subset \mathscr{X}_n(\mathcal{D}) \,\middle|\, \deg(F) = \bigwedge_{i=1}^{k} (l_i, r_i)\right], \tag{14}$$

where $\deg(F) = \bigwedge_{i=1}^{k}(l_i, r_i)$ represents the event that a randomly chosen edge set F of size k has as endpoints with degrees (l_i, r_i), $1 \le i \le k$. If $e_i \in F$ has endpoints with degrees l_i and r_i, the probability of it existing in $\mathscr{X}_n(\mathcal{D})$ is given by $l_i r_i / \sum_{i=0}^{n-1} i \cdot d_i(n)$. Defining $D := \sum_{i=0}^{n-1} i \cdot d_i(n)$, we can bound the above probability via

$$\rho \le \frac{1}{D^k} \prod_{i=1}^{k} l_i r_i.$$

Since

$$\mathbf{Pr}\left[F \subset \mathscr{X}_n(\mathcal{D}) \wedge \deg(F) = \bigwedge_{i=1}^{k}(l_i, r_i)\right] = \rho \cdot \mathbf{Pr}\left[\deg(F) = \bigwedge_{i=1}^{k}(l_i, r_i)\right],$$

we next consider the second term on the right hand side. This term expresses the probability that a randomly chosen set of edges $e_1, \ldots, e_k$ is such that edge e_i has as its endpoints with degrees l_i and r_i. A simple upper bound is the following:

$$\mathbf{Pr}\left[\deg(F) = \bigwedge_{i=1}^{k}(l_i, r_i)\right] \le \prod_{i=1}^{k} \frac{d_{l_i}(n) \cdot d_{r_i}(n)}{(n-2k)^2}.$$

Recall that by the sparseness-condition of the degree sequence $\mathcal{D}$, we have that $\sum_{i=0}^{n-1} i \cdot d_i(n)/n = \mu + o(1) := \alpha$, from which it follows that $2m = \alpha n$. Using the union bound, we sum over all possible choices of degrees $l_1, \ldots, l_k, r_1, \ldots, r_k$ to upper bound $\mathbf{Pr}\left[F \subset \mathscr{X}_n(\mathcal{D})\right]$ by:

$$\sum_{(l_1,r_1),\cdots,(l_k,r_k)} \frac{1}{D^k} \prod_{i=1}^{k} l_i r_i \cdot \frac{\prod_{i=1}^{k} d_{l_i}(n) \cdot d_{r_i}(n)}{(n-2k)^{2k}}, \tag{15}$$

which simplifies to

$$\frac{1}{D^k} \cdot \frac{1}{(n-2k)^{2k}} \cdot \sum_{(l_i,r_i)} \prod_{i=1}^{k} (l_i d_{l_i}(n) \cdot r_i d_{r_i}(n)). \tag{16}$$

We again use the bound $(2m)^{2k}$ for the sum-of-products (proven in Lemma 4) and hence bound the whole expression (16) by

$$\frac{1}{D^k} \cdot \frac{(2m)^{2k}}{(n-2k)^{2k}} = \left(\frac{4}{D} \frac{m^2}{n^2(1-\frac{2k}{n})^2}\right)^k = \left(\frac{4}{D} \frac{\alpha^2}{(1-\frac{2k}{n})^2}\right)^k = O\left((4\alpha^2\mu/n)^k\right)$$

and defining $\theta := 4\alpha^2\mu$ proves the statement. $\qquad\square$

Again, Theorem 3 holds in particular for the degree distributions listed in Table 1:

Corollary 2 *Let $\mathcal{D}$ be a smooth and sparse asymptotic degree sequence which converges to a normalized version of a function from Table 1. Then $\mathscr{X}(\mathcal{D})$ has bounded expansion w.h.p.*

3.4 A Generalization of Erdős–Rényi

One of the most well-studied models for random graphs was introduced by Erdős and Rényi. Often denoted $\mathscr{G}(n, p)$, this model has only two input parameters—the number of nodes n and an edge-probability $p \in [0, 1]$. An instance is created by starting with a set of n isolated vertices and adding each of the $\binom{n}{2}$ possible edges independently with probability p. Prior work has shown that the expected vertex degree μ in $\mathscr{G}(n, p)$ is $p(n - 1)$ and the probability p_i that a vertex has degree exactly i is

$$p_i = \binom{n-1}{i} p^i (1-p)^{n-1-i} \approx \frac{\lambda^i e^{-\lambda}}{i!},$$

(where the approximate equality becomes exact as $n \to \infty$), giving a Poisson degree distribution.

This very strict model, in which all the edges have the same probability, can be somewhat relaxed. To model graphs with non-Poisson degree distributions, we will define a new class of graphs consisting of the union of a graph from a random graph class with a bounded-degree graph (which is trivially of bounded expansion on its own.) Specifically, for a fixed $\lambda \in \mathbf{N}$, let $\mathscr{C}^\lambda$ be a graph model such that $\mathbf{Pr}\left[\Delta(\mathscr{C}_n^\lambda) \leq \lambda\right] = 1$ for all $n \geq 0$, i.e. all graphs drawn from $\mathscr{C}^\lambda$ have degree bounded by λ.

We augment this class as follows: fix $\mu \in \mathbf{R}^+$ and let $\hat{p} = (p_i)_{i \in \mathbf{N}}$ be a sequence of functions $p_n \colon [n] \times [n] \to [0, 1]$ such that for $u, v \in [n]$ it holds that $p_n(u, v) = p_n(v, u)$ and $p_n(v, v) = 0$. Now define $\mathscr{G}_n(\mathscr{C}^\lambda, \hat{p})$ to be a random graph constructed at follows:

1. Draw an n-vertex graph G from $\mathscr{C}_n^\lambda$.

2. For $u, v \in [n] \times [n]$, if $uv \notin G$, add uv to G with probability $p_n(u, v)$.

We will call the graph G drawn in the first step the *base graph*. In the Erdős–Rényi model, the base graph G would be the edgeless graph and the edge probabilities p_n constant functions for all n. We will show that $\mathscr{G}_n(\mathscr{C}^\lambda, \hat{p})$ has bounded expansion if for all $n \in N$ and $u, v \in [n] \times [n]$ it holds that $p_n(u, v) \leq \mu/n$ for a fixed constant μ.

This result might seem very intuitive at first: Adding few random edges to a bounded degree graph should not change the structural density too much. However, note that bounded degree graphs can contain arbitrarily dense minors—take, for example, the bounded degree graph obtained from the lexicographic product of an $n \times n$ grid with K_2. We need to show that the random edges added to base graph will not turn such dense minors into dense *shallow* minors.

Let us quickly demonstrate that our result is tight in the sense that a base graph of arbitrary degree does not yield a model with bounded expansion. Take as a simple example the class of graphs consisting of $\sqrt{n}$ stars each of degree $\sim\sqrt{n}$. This graph class has clearly bounded expansion, since the density of any minor is smaller than one. Adding Erdős–Rényi edges to this graph with probability $1/n$, however, will connect a fixed pair of stars with constant probability—hence we create any arbitrarily dense 1-shallow topological minor a.a.s.

We now state the main result of this section.

Theorem 4 *Let $\lambda \in \mathbf{N}$ and $\mu \in \mathbf{R}^+$. Then $\mathscr{G}(\mathscr{C}^\lambda, \hat{p})$ has bounded expansion w.h.p. if for all $n \in \mathbf{N}$ the function p_n is bounded from above by μ/n.*

Proof: As before, to show that $\mathscr{G}(\mathscr{C}^\lambda, \hat{p})$ has bounded expansion, we verify the conditions stated in Proposition 1.

The first condition requires that the fraction of vertices with large degree is vanishingly small. Specifically, in Lemma 7, we prove that the probability there are εn vertices of degree at least $f_2(\varepsilon) = 2\lambda + 8\mu\alpha$, where $\varepsilon = 4e\alpha^{-4\alpha\mu}$, tends to zero as n tends to infinity (so there exists $f_1(\varepsilon)$ such that for all $n > f_1(\varepsilon)$, condition (i) is satisfied with high probability.)

To show that $\mathscr{G}_n(\mathscr{C}^\lambda, \hat{p})$ satisfies condition (ii), we prove all subgraphs whose shallow topological minors are sufficiently dense span a large fraction of the vertices. We start with Lemma 6, which bounds the probability of a creating a uv-path in $\mathscr{G}_n(\mathscr{C}^\lambda, \hat{p})$ that decreases $d(u,v)$ in terms of a function $g(r, \lambda, \mu)$ and n. We combine this with Lemma 1 in Corollary 4 to obtain condition (ii).

Taken together, this shows that with high probability $\mathscr{G}(\mathscr{C}^\lambda, \hat{p})$ has bounded expansion. $\qquad\square$

In the following, let us fix the n-vertex base graph G drawn from $\mathscr{C}_n^\lambda$. We denote the random graph obtained from G by the above process as $\mathscr{G}_n(G, p_n)$.

Lemma 6 *Let G be a graph with maximum degree λ and let $r \in \mathbf{N}$. Fix vertices $u, v \in G$. Then the probability that there exists a path in $\mathscr{G}_n(G, p_n)$, where p_n is bounded by μ/n for some $\mu \in \mathbf{R}^+$, with at most $2r + 1$ edges between u and v that uses at least one edge not in $E(G)$ is at most*

$$\frac{g(r, \lambda, \mu)}{n} := \frac{(2r + 1)e\sqrt{2r} \cdot (\lambda^{4r} \cdot \mu)^{2r+1}}{n}.$$

Proof: Let $\mathcal{P}$ be a u-v-path of length at most $2r + 1$ that uses at exactly $i \geq 1$ random edges. We can view this path as consisting of 'segments' $\mathcal{P}_u, \mathcal{P}_1, \ldots, \mathcal{P}_{i-1}, \mathcal{P}_v$ where each segment is a sequence of (at most $2r$) vertices

that forms a path by itself in the base graph G. If the path has exactly i random edges then there are $i - 1$ internal segments that connect $\mathcal{P}_u$ and $\mathcal{P}_v$ (the segments that contain the endpoints.)

We first wish to estimate an upper bound on the probability that two segments are connected by a random edge. This probability is not simply μ/n, as each segment can have several vertices. Since the base graph has maximum degree λ, and since each segment has at most $2r$ vertices, it is contained in some $2r$-ball in the base graph G. A $2r$-ball in the base graph has at most λ^{2r} vertices and the probability that two such balls are connected by a random edge is at most $\lambda^{4r} \cdot \mu/n$. This is also an upper bound on the probability that two segments, each with at most $2r$ vertices, are connected by a random edge.

To estimate the overall probability, we sum over all values of i. If there are exactly i random edges then our path has $i - 1$ internal segments. The number of ways of choosing these segments is at most $\binom{n}{i-1}$, as we can view this process as selecting $i - 1$ representatives, one for each segment. Moreover, these $i - 1$ internal segments can be ordered in $(i - 1)!$ ways. Therefore an upper bound on the probability that there exists a path from u to v with at most $2r + 1$ edges is given by the expression:

$$\sum_{i=1}^{2r+1} \binom{n}{i - 1} \cdot (i - 1)! \cdot \left(\frac{\lambda^{4r} \cdot \mu}{n} \right)^i. \tag{17}$$

The expression within the sum can be written as:

$$
\begin{aligned}
\binom{n}{i - 1} \cdot (i - 1)! \cdot \left(\frac{\lambda^{4r} \cdot \mu}{n} \right)^i &\le \left(\frac{ne}{i - 1} \right)^{i-1} \cdot (i - 1)! \cdot \left(\frac{\lambda^{4r} \cdot \mu}{n} \right)^i \\
&= \frac{(i - 1)! \cdot e^{i-1}}{(i - 1)^{i-1}} \cdot \frac{(\lambda^{4r} \cdot \mu)^i}{n} \\
&\le \frac{e\sqrt{i - 1} \cdot (\lambda^{4r} \cdot \mu)^i}{n}.
\end{aligned}
$$

The last inequality follows from the fact that $(e^r r!)/r^r \le e\sqrt{r}$ (using Stirling's approximation.) Thus the sum in equation (17) is at most:

$$(2r + 1) \cdot \frac{e\sqrt{2r} \cdot (\lambda^{4r} \cdot \mu)^{2r+1}}{n}.$$

$\square$

Corollary 3 *Let G be a graph with maximum degree λ and let $r \in \mathbf{N}$. Let p_n be bounded from above by μ/n for some $\mu \in \mathbf{R}^+$. The probability that any k*

vertices chosen uniformly at random from $\mathcal{G}_n(G, p_n)$ will be connected by at least $\xi k > \lambda k$ pairwise internally vertex-disjoint paths of length at most $2r+1$ is at most

$$\binom{\binom{k}{2}}{k \cdot (\xi - \lambda)} \cdot \left(\frac{g(r, \lambda, \mu)}{n} \right)^{k \cdot (\xi - \lambda)}.$$

Hence, $\mathcal{G}_n(G, p_n)$ is $(g(r, \lambda, \mu), 0)$-sparse.

Proof: Using the terminology from Definition 1, let $k := |V(H)|$ so that the model of H (which is a subgraph of $\mathcal{G}_n(G, p_n)$) has q nails and at most $q \cdot \xi \cdot (2r + 1)$ subdivision points. Given a set $v_1, \ldots, v_q$ of q vertices in $\mathcal{G}_n(G, p_n)$, let $\mathcal{P}$ denote the set of all paths of length at most $2r + 1$ between any two of these vertices. Since at most $k \cdot \xi$ of these paths could have existed in G, at least $k \cdot (\xi - \lambda)$ of these paths exist in $\mathcal{G}_n(G, p_n)$ because of random edges.

Therefore the probability that we wish to estimate is simply the probability that a set of k vertices chosen uniformly at random have at least $k \cdot (\xi - \lambda)$ new paths of length at most $2r + 1$ between them in $\mathcal{G}_n(G, p_n)$ By Lemma 6, this probability is—for a fixed k—given exactly the one claimed above. $\square$

Corollary 4 *There exist constants n_0, c_0 such that for every r-shallow topological minor H of $\mathcal{G}_n(G, p_n), |G| \geq n_0$ with at most $c_0 \cdot n$ vertices and every $\xi \geq \lambda + 2$ it holds that*

$$\mathbf{Pr}\left[\frac{|E(H)|}{|V(H)|} > \xi \right] \leq 1/n^\xi.$$

Proof: Combining Corollary 3 with Lemma 1 and the functions $\alpha(r) = g(r, \lambda, \mu), \beta(r) = 0$ and gives the above statement when $\xi \geq \lambda + 2$. $\square$

It is left to show that the number of high-degree vertices is small. This was proven for Erdős-Renyi graphs in [68] (Lemma 4.3) and thus carries over to our case, albeit only asymptotically almost surely. Therefore we supplement our own version of the proof with better speed of convergence.

Lemma 7 *For any $0 < \varepsilon < 1$, let α be such that $\varepsilon = 4e\alpha^{-4\alpha\mu}$. Let G be an n-vertex graph with maximal degree λ and let A be a set of $s := 2\lceil (\varepsilon n)/2 \rceil$ vertices in $\mathcal{G}_n(G, p_n)$ of largest degree, where again p_n is bounded from above by μ/n. Then*

$$\lim_{n \to \infty} \mathbf{Pr}\left[\min_{x \in A} \deg(x) \geq 2\lambda + 8\alpha\mu \right] = 0.$$

Proof: Let A be as in the statement of the lemma and let $\delta := \min_{x \in A} \deg(x)$. Then there are at least $s\delta/2$ edges with at least one endpoint in A. If A' is a random subset of A of size $s/2$, then every edge that has an endpoint in A has a probability of at least $1/2$ of having *exactly* one endpoint in A'. Therefore there exists a subset $A' \subseteq A$ of size $s/2$ such that the number of edges crossing the cut $(A', V \setminus A')$ is at least $s\delta/4$. However if $\delta \geq 2\lambda + 8\alpha\mu$, then it implies that there exists $A' \subseteq A$ of size $s/2$ such that the number of edges crossing the cut $(A', V \setminus A')$ is at least $(\lambda + 4\alpha\mu)s/2$.

To show that $\mathbf{Pr}\left[\delta \geq 2\lambda + 8\alpha\mu\right]$ is small, we will show that the probability that there exists a set A' with $s/2$ vertices such that the cut $(A', V \setminus A')$ has at least $(\lambda + 4\alpha\mu)s/2$ edges is $o(1)$. Fix a set A' with $s/2$ vertices. The probability that q edges were added to the cut $(A', V \setminus A')$ is given by:

$$
\mathbf{Pr}\left[|E(A', V \setminus A')| = q\right] \leq \binom{\frac{s}{2} \cdot \left(n - \frac{s}{2}\right)}{q} \cdot \left(\frac{\mu}{n}\right)^q
$$
$$
\leq \left(\frac{ens}{2q}\right)^q \cdot \left(\frac{\mu}{n}\right)^q
$$
$$
\leq \left(\frac{2s\mu}{q}\right)^q.
$$

If the total number of edges crossing the cut is $(\lambda + 4\alpha\mu)s/2$ then q must be at least $2\alpha\mu s$. The probability that at least $2\alpha\mu s$ edges were added to the cut is at most

$$
\left(\frac{2s\mu}{q}\right)^q \leq \left(\frac{2s\mu}{2\alpha\mu s}\right)^q \leq \alpha^{-q} \leq \alpha^{-2\alpha\mu s},
$$

where the last inequality holds since $\alpha > 1$. Therefore the probability that there exists a set A' with $s/2$ vertices such that $|E(A', V \setminus A')| \geq (\lambda + 4\alpha\mu)s/2$ is:

$$
\alpha^{-2\alpha\mu s} \cdot \binom{n}{s/2} < \alpha^{-2\alpha\mu s} \cdot \left(\frac{2en}{s}\right)^{s/2} \leq \left(\frac{e\alpha^{-4\alpha\mu}n}{s/2}\right)^{s/2} \leq \left(\frac{s/4}{s/2}\right)^{s/2} = \frac{1}{2^{s/2}}.
$$

Since $s = \varepsilon n$ (roughly), this probability approaches zero as $n \to \infty$. $\qquad\square$

Finally, since Erdős–Rényi random graphs are a special case of our model (when the bounded degree graph is edgeless, and all the edge-probabilities are equal), we have a strengthening of the previously best known result of Nešetřil and Ossona de Mendez (that bounded expansion holds asymptotically almost surely in the Erdős–Rényi model.) We also note that the result carries over to the *stochastic block model*, which was first studied in mathematical sociology by Holland, Laskey and Leinhardt in 1983 [39] and by Wang and Wong in 1987 [89], if the probabilities involved are small enough.

4 Graph Models without Bounded Expansion

In this section we consider the Kleinberg [43, 44] and Barabási-Albert [6, 9] Models, which, respectively, were designed to replicate "small-world" properties and heavy-tailed (power-law) degree distributions observed in complex networks. We show that both these models (with typical parameters) do not have bounded expansion, and in fact are somewhere dense w.h.p./non-vanishing probability, respectively. This is done by showing the existence of two/one-subdivisions of cliques respectively in the generated graphs with a certain probability.

4.1 The Kleinberg Model

Many social networks exhibit a property that is commonly referred to as the "small-world phenomenon." This property asserts that any two people in a network are likely to be connected by a short chain of acquaintances. This was first observed by Stanley Milgram in a study published in 1967 [61]. Milgram's study suggested that individuals in a social network who only knew the locations of their immediate acquaintances are collectively able to construct short chains between two points in the network. More recently, Kleinberg proposed a family of network models to explain the success of decentralized algorithms in finding short paths in social networks [44].

Kleinberg's model starts with a $n \times n$ grid as the base graph and allows edges to be directed. For a universal constant $p \geq 1$, a node u has a directed edge to every other node within lattice distance p. These are the *local neighbors* of u. For universal constants $q \geq 0$ and $r \geq 0$, node u has q *long range neighbors* chosen independently at random. The ith directed outarc from u has endpoint v with probability $d(u,v)^{-r}/\sum_x d(u,x)^{-r}$.

When $r = 0$, the long-range contacts are uniformly distributed throughout the grid, and one can show that there exist paths between every pair of nodes of length bounded by a polynomial in $\log n$, exponentially smaller than the number of nodes. Kleinberg shows that in this case, the expected delivery time of *every* decentralized algorithm (one that uses only local information) is $\Omega(n^{2/3})$. When $p = q = 1$ and $r = 2$, then short chains continue to exist between the nodes of the network, but here is a decentralized algorithm to transmit a message that takes $O(\log^2 n)$ time in expectation between any two randomly chosen points.

What Kleinberg's model shows is that if the long-range contacts are formed independently of the geometry of the grid, then short chains exist between every pair of nodes, but nodes working with local knowledge are unable to find them. If the long-range contacts are formed by taking into account the

grid structure in a specific way, then short chains exist and nodes working with local knowledge are able to discover them. We show that for those parameters where greedy routing is efficient, not only does the model not have bounded expansion, it is in fact, *somewhere dense* w.h.p.

Theorem 5 *The Kleinberg model with parameters $p = q = 1$ and $r = 2$ is somewhere dense w.h.p.*

Proof: Let Γ_n be an $n \times n$ grid. For $p = q = 1$ and $r = 2$, the probability that a node u has v as its long-range contact is proportional to $d_{\Gamma_n}(u, v)^{-1}$ and the normalizing factor in this case is $O(1/\log n)$. This can be easily seen by summing up $1/d_{\Gamma_n}(u, x)^2$ for all x and noticing that in the grid, there are $4d$ neighbors that are at a distance of d from u.

$$\sum_x \frac{1}{d_{\Gamma_n}(u, x)^2} = \sum_{d=1}^n \frac{4d}{d^2} \approx 4 \log n.$$

To show that the model is somewhere dense, we show that 2-subvidisions of cliques of a certain size $g(n)$ occur with high probability. Later we will see that $g(n) = \Omega(\log \log n)$. To this end, let $\Gamma'_{c \cdot g(n)}$ denote some fixed $c \cdot g(n) \times c \cdot g(n)$ subgrid of Γ_n, where c is some constant that we will fix later. Choose V' and E' to be, respectively, a set of $g(n)$ nodes and a set of $g(n)^2$ edges from the subgrid $\Gamma'_{c \cdot g(n)}$ with the following properties: (i) the endpoints of the edges in E' are different from the nodes in V'; (ii) no two edges in E' share an endpoint.

Given any pair of vertices $u, v \in V'$ and an edge $e \in E'$ with endpoints a, b, the probability that a has u as its long-range neighbor is $\Omega((d_{\Gamma_n}(a, u) \cdot \log n)^{-1})$. Similarly, the probability that b has v as its long-range neighbor is $\Omega((d_{\Gamma_n}(b, v) \cdot \log n)^{-1})$ The probability of both these events happening is

$$\frac{1}{d_{\Gamma_n}(a, u)^2 d_{\Gamma_n}(b, v)^2} \cdot \frac{1}{\log^2 n} \geq \frac{1}{c^4 g(n)^4 \log^2 n}, \tag{18}$$

where we upper-bounded distances $d_{\Gamma_n}(x, y)$ by $c \cdot g(n)$. Thus the probability that there exists a 2-subdivided $g(n)$-clique in $\Gamma'_{c \cdot g(n)}$ is at least:

$$\left(\frac{1}{c^4 g(n)^4 \log^2 n} \right)^{g(n)^2} =: f(n, c). \tag{19}$$

The probability that there does *not* exist a 2-subdivided $g(n)$-clique in $\Gamma'_{c \cdot g(n)}$ is at most $1 - f(n, c)$. Hence the probability that there does not exist a 2-subdivided $g(n)$-clique in *any* $c \cdot g(n) \times c \cdot g(n)$ subgrid is at most

$$(1 - f(n, c))^{\frac{n}{c^2 g(n)^2}} \leq \exp\left(-\frac{n}{c^2 \cdot g(n)^2 \cdot \left(c^4 \cdot g(n)^4 \cdot \log^2 n \right)^{g(n)^2}} \right) := e^{\frac{n}{h(n)}}.$$

28

This follows from the inequality $(1 - x/p)^p \leq e^{-x}$.

Choose $g(n) = \log \log n$ and $c = 3$ (actually any $c \geq 3$ works.) Then it is easy to show that $h(n) < \sqrt{n}$. Thus the probability of a 2-subdivided $(\log \log n)$-clique *not* existing is at most $e^{-\sqrt{n}}$ (which goes to zero as $n \to \infty$), and we conclude that the graph model is somewhere dense. $\qquad\square$

4.2 The Barabási-Albert model

The Barabási-Albert model uses a preferential attachment paradigm to produce graphs with a degree distribution that mimics the heavy-tailed distribution observed in many real-world networks [9]. This model uses a random graph process that works as follows: Start with a small number n_0 of nodes and at every time step, add a new node and link it to $n \leq n_0$ nodes already present in the "system." To model preferential attachment, we assume that the probability with which a new node u is connected to node v already present in the system is proportional to the degree of v, so that $\mathbf{Pr}[u \to v] = \deg(v)/\sum_x \deg(x)$, where the sum in the denominator is over all vertices x that are already in the system. After t time steps, the model leads to a random network with $t + n_0$ vertices and nt edges.

Barabási and Albert suggested that such a network evolves into one in which the fraction $P(d)$ of nodes of degree d is proportional to $d^{-\gamma}$. They observed experimentally that $\gamma = 2.9 \pm 0.1$ and suggested that γ is actually 3. This model was rigorously analyzed by Bollobás, Riordan, Spencer and Tusnády in [14] who showed that it is indeed the case that the fraction of vertices of degree d fall off as d^{-3} as $d \to \infty$. In [13], Bollobás and Riordan showed that the diameter of the graphs generated by this model is asymptotically $\log n / \log \log n$.

We first provide a formal restatement of the Barabási-Albert Model. Note that this is slightly different from the formalization of Bollobás et al. in [14]. We start with a "seed" graph G_0 with n_0 nodes $u_1, \ldots, u_{n_0}$ with degrees $d_1, \ldots, d_{n_0}$. The number of edges in the seed graph is denoted by m_0. At each time step $t = 1, 2, \ldots$, we create a graph G_t by adding a new node v_t and linking it to q nodes in G_{t-1}. These q nodes are picked independently and with a probability that is proportional to their degrees in G_{t-1}. That is, we choose $u \in V(G_{t-1})$ to link to with probability

$$\mathbf{Pr}\left[\{v_t, u\} \in E(G_t)\right] = \frac{\deg_{G_{t-1}}(u)}{2 \cdot |E(G_{t-1})|} = \frac{\deg_{G_{t-1}}(u)}{2(m_0 + q(t-1))}.$$

Note that in this model, all edges between v_t and nodes of G_{t-1} are assumed to be added simultaneously (so that the increasing degrees of nodes which

receive edges from v_t do not influence the probabilities for this time step.) In this restated version, $n_0, d_1, \ldots, d_{n_0}$, and q are the parameters of the model.

Lemma 8 *Given any fixed r, a graph G_n generated by the preferential attachment model with parameters $n_0, d_1, \ldots, d_{n_0}$ and $q \geq 2$ has a 1-subdivided K_r as a subgraph with probability at least $(4(\frac{m_0}{q} + r + r^2))^{-r^2}$, provided $n \geq r + r^2$.*

Proof: Choose any $r \in \mathbf{N}$. We will show that there exists a 1-subdivided K_r in the graph with probability that depends only on m_0, q, and r.

Consider the graph after the first $r + r^2$ time steps. Let $v_1, \ldots, v_r$ be the nodes that were added in the first r time steps and fix two nodes v_i, v_j from among these. The probability that a new node v_k (for any $r + 1 \leq k \leq r^2$) is connected to these two fixed nodes is at least

$$\left(\frac{q}{2(m_0 + q(r + r^2))} \right)^2 =: f(m_0, r, q),$$

where the denominator is the sum of the vertex degrees after $r + r^2$ time steps. Now if v_{r+1} is linked to v_1, v_2 and v_{r+2} is linked to v_1, v_3 and so on such that the nodes added after time step r connect the first r nodes in a pairwise fashion, we would have a 1-subdivided K_r in the graph G_{r+r^2}. The probability of this happening is at least $f(m_0, r, q)^{r^2}$. Thus, for every r, the probability of 1-subdivided K_r existing is non-zero if the graph is large enough. $\qquad \square$

It immediately follows that the Barabási-Albert model (and similar preferential attachment models) is not a.a.s. nowhere denseand in particular does not have bounded expansion a.a.s. We note that this result is more of theoretical interest, since the probabilities involved might be small enough to be irrelevant in practice. As such it would be worthwile to investigate whether the Barabási-Albert model is somewhere dense a.a.s.

5 Experimental Evaluation

Although it has been established that real-world networks are sparse, and tend to have low degeneracy (relative to the size of the network), it is natural to ask whether there is empirical evidence that they satisfy the stronger conditions of bounded expansion. Unfortunately, since bounded expansion itself is a property of graph classes and not single graphs, it is impossible to determine whether individual instances are bounded expansion or not. One natural proxy would be to evaluate the *grad* of these graphs, calculating the maximum density of an r-shallow minor for each $r \in \mathbf{N}$ (obviously stopping when r is the diameter of the network), but it is not known how to find such minors in reasonable time.

In order to get around these difficulties, we calculate upper bounds on χ_{p-1} (the p-centered coloring number), a good proxy, since it and the grad are both related to each other by factors independent of the graph size. This is further justified by the fact that p-centered colorings are directly applicable to algorithm design, where the complexity of such algorithms depends heavily on the number of required colors (as will be shown in Section 6). Since it is very time-consuming to obtain a good p-centered coloring for large p (this is analogous to determining a reasonable bound for the maximum density of an r-shallow minor for large r), we evaluate this property for small values only. This is also roughly the range of p which is relevant to the algorithms presented later in this paper when applied to practical settings.

To obtain upper bounds on χ_{p-1}, we implemented the transitive- fraternal augmentation procedure of Nešetřil and Ossona de Mendez [66]. Our theoretical results predict that graphs generated with the configuration model for typical degree distributions of complex networks will likely have bounded expansion. We thus compared the results for χ_3 of this procedure between real networks and networks generated using the configuration model for the same degree distributions. We chose χ_3 since it is relatively easy to compute for large networks but still heavily influenced by one-subdivisions of cliques. The results of this experiment can be found in Figure 1.

For almost all networks the bound for the real-world network is either smaller or comparable to the values for the synthetic graphs. The one exception is "power", which is not surprising since this network has a relatively complex (grid-like) structure over low degree nodes.

We furthermore extended the algorithm based on transitive-fraternal augmentations with simple heuristic improvements (e.g. giving high degree nodes their own private color, merging color classes where possible), ran it to find p-centered colorings on a small corpus of well-known complex networks and verified the results. The results of the best colorings we were able to achieve with this relatively simple methos can be found in Table 2.

The results show that some networks clearly have a moderately growing grad; in particular the larger networks *Netscience*, both *Cpan-* networks and *Diseasome*. Other networks, like *Twittercrawl*, have such quickly growing p-centered coloring numbers that we did not invest the time to determine the value for larger p. Since graphs of bounded crossing number (as infrastructure networks tend to be) and bounded degree have bounded expansion, we are not surprised at the small number of colors needed by *Power* and *Hex*. Finally some networks, like *CondMat* and *Hep-th*, start out reasonably well for small p but show a sudden jump at $p = 3$. At present we do not know whether this is an artifact of the procedure we use to obtain the coloring or whether the networks have indeed dense minors from a certain depth on. As shown in
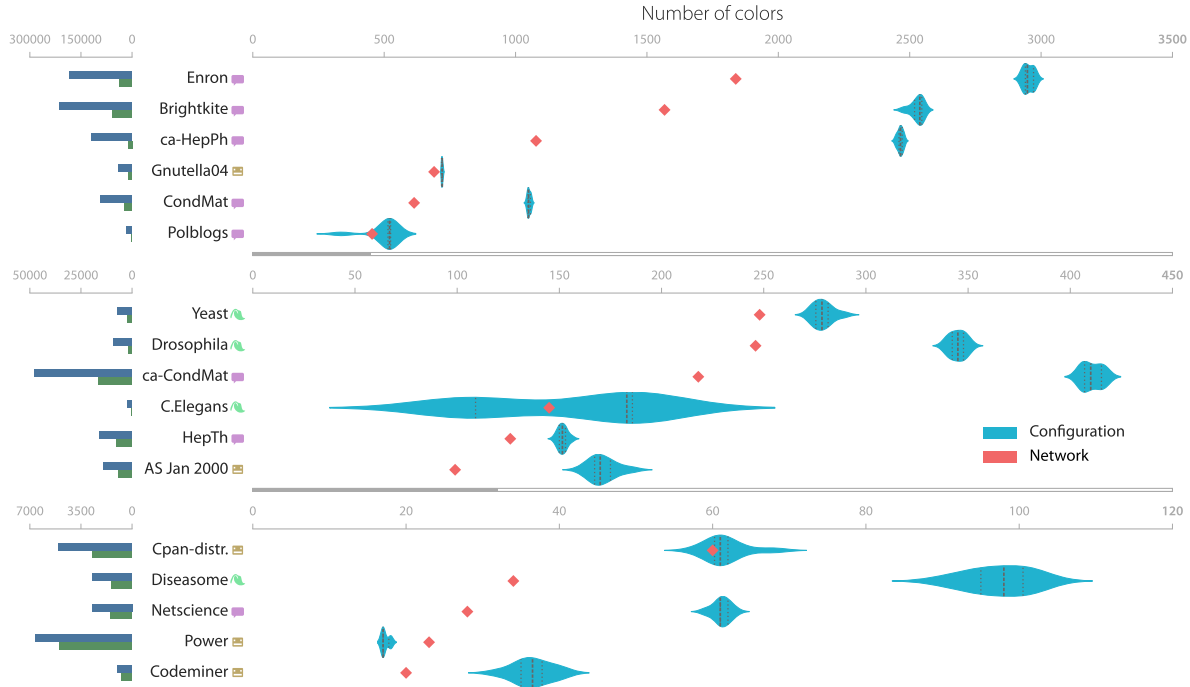
Figure 1: Comparison of 4-centered coloring numbers on sets of small (bottom), medium (middle) and large (top) real-world networks (red diamonds) and comparable synthetic graphs (blue violins); axes are scaled for each set. Violins show distribution over 10 random instances with matching degree distribution (generated with the configuration model); median/quartiles are marked with dashed lines. The number of nodes (green) and edges (blue) are indicated to the left of the network name. Icons indicate application domain (violet: social/collaboration, green: biological, brown: infrastructure/technology).

Section 4, some complex network models predict such an occurrence already for depth at most two. The growth behavior for small p as depicted in Table 2 might therefore serve as a property to distinguish types of networks, meriting future research.

In the last column, we provide upper bounds for the treedepth of these graphs. Notice that the number of colors needed for a p-centered coloring will be fewer than the treedepth of the the graph for any $p \leq n$: Given a treedepth decomposition of a graph of depth t we can color every node by its depth in the treedepth decomposition. Since the treedepth is a hereditary property, the graph induced by any subset of nodes will have treedepth at most t. Notice that the simple coloring algorithm we used sometimes colors the graph with more than t colors. This is a good indication that a better coloring algorithm or heuristic exists.

We also note that the subgraph isomorphism algorithm presented in Section 6 should be directly applicable to some of these graphs, given their

Network	Vertices	Edges	2	3	4	5	6	∞
					p			
Karate [92]	35	78	6	7	9	9	10	8
Dolphins [55]	62	159	7	11	17	18	19	24
Lesmiserables [46]	77	254	10	15	16	16	16	16
Polbooks [48,69]	105	441	8	16	22	29	31	30
Word-adjacencies [73]	112	425	8	18	27	35	41	48
Football [35]	115	613	9	22	33	49	64	69
Airlines [2]	235	1297	13	28	39	52	67	64
Sp-data-school [84]	238	5539	31	123	151	174	179	171
C.Elegans [90]	306	2148	12	41	74	103	148	153
Hex-grid	331	930	4	9	21	22	26	69
Codeminer [2]	724	1017	5	11	17	20	21	51
Cpan-authors [1]	839	2212	9	29	40	44	62	224
Diseasome [36]	1419	2738	12	17	22	25	30	30
Polblogs [3]	1491	16715	30	118	286	354	392	603
Netscience [73]	1589	2742	20	20	28	28	28	20
Drosophila [86]	1781	8911	12	65	137	188	263	395
Yeast [16]	2284	6646	12	38	178	254	431	408
Cpan-distr. [1]	2719	5016	5	14	32	42	56	224
Twittercrawl [2]	3656	154824	89	561	1206	1285	–	–
Power [90]	4941	6594	6	12	20	21	34	95
AS Jan 2000 [50]	6474	13895	12	32	103	316	317	357
Hep-th [70]	7610	15751	24	25	137	137	–	558
Gnutella04 [50,81]	10876	39994	8	49	705	–	–	–
ca-HepPh [50]	12008	118489	239	296	1002	–	–	–
CondMat [70]	16264	47594	18	47	255	1839	–	1310
ca-CondMat [50]	23133	93497	26	89	665	–	–	–
Enron [45,53]	36692	183831	27	214	–	–	–	–
Brightkite [18]	58228	214078	39	193	–	–	–	–

Table 2: Number of colors in p-centered colorings computed on real world networks and upper bounds of their respective treedepth. These networks were mostly taken from the datasets found in [2,51,69].

comparatively low treedepth.

Finally, we argue that for practical purposes the definition of low treedepth colorings is too strict. Remember that per definition, in a p-centered coloring every graphs induced on $i < p$ colors has treedepth $\leq i$. Relaxing this latter condition, we arrive at the following variation of Proposition 3:

Observation 1 *Let $\mathcal{G}$ be a graph class of bounded expansion. There exists*

function f and g such that for every $G \in \mathcal{G}$, $p \in \mathbf{N}$, the graph G can be colored with $f(r)$ colors so that any $i < p$ color classes induce a graph of treedepth $\leq g(i)$ in G.

Obviously this follows from Proposition 3, taking g as the identity. However, algorithms based on low treedepth colorings would run faster if g allows for a larger margin, provided we can decrease the number of colors f. This is owed to the fact that for large number of colors, iterating through all $< p$-sized subsets will be the deciding factor in the running time.

6 Algorithms

In this section, we present efficient algorithms for several important problems arising in the study complex networks which exploit bounded expansion. Note that in both cases, only the algorithm's running time relies on the grad being small, not its correctness. The problems that we discuss revolve around the themes of *subgraph counting* and *centrality estimation.*

Computing the frequency of small fixed pattern graphs inside a network is the key algorithmic challenge in using *network motifs* and *graphlet degree distributions* to analyze network data (both of which are described in more detail in Section 6.1). We present a parameterized algorithm for counting the number of subgraphs with at most h vertices with a running time of $8^h \cdot t^h \cdot n$, where t is the treedepth of the input graph. In a graph class of bounded expansion, we use this algorithm in conjunction with p-centered colorings.

Another topic of interest in complex networks is estimating the relative importance of a vertex in the network (for example, how influential a person is inside a social network, which roads are busiest in a road-network, or which location is most attractive for business). The typical approach is to define/select an appropriate *centrality measure* (see [47] for a survey of common measures). We focus on the *closeness centrality*, which was introduced by Sabidussi [82], and related extensions. These measures are related in the sense that computing them requires knowledge of all the pairwise distances between the vertices of the network, which even in sparse networks takes time $O(n^2)$ to compute [15]. We introduce localized variants of the closeness-based centrality measures and design a linear-time algorithm to compute them in graphs of bounded expansion, together with experimental data that suggests that these measures are able to recover the topmost central vertices quite well.

Before we begin, we formally state the relationship between bounded expansion and low-treedepth colorings, and give a result of Nešetřil and Ossona de Mendez that we will make use of.

Proposition 3 (Low treedepth colorings [66]) *Let $\mathcal{G}$ be a graph class of bounded expansion. There exists a function f such that for every $G \in \mathcal{G}$, $r \in \mathbf{N}$, the graph G can be colored with $f(r)$ colors so that any $i < r$ color classes induce a graph of treedepth $\leq i$ in G. This coloring can be computed in linear time.*

Proposition 4 (Truncated distances [66]) *Let G be a graph of bounded expansion. For every r one can compute in linear time a directed graph $\vec{G}_r$ with in-degree bounded by $f(r)$ – for some function f – on the same vertex set as G and an arc labeling $\omega\colon \vec{E}(\vec{G}_r) \to \mathbf{N}$ such that for every pair $u, v \in G$ with $d_G(u, v) \leq r$ one of the following holds:*

(i) $uv \in \vec{G}_r$ and $\omega(uv) = d_G(u, v)$;

(ii) $vu \in \vec{G}_r$ and $\omega(vu) = d_G(u, v)$;

(iii) there exists $w \in N^-_{\vec{G}_r}(u) \cap N^-_{\vec{G}_r}(v)$ such that $\omega(wu) + \omega(wv) = d_G(u, v)$.

6.1 Counting graphlets and subgraphs

In the following we highlight three domain-specific applications of computing the frequency of small fixed pattern graphs inside a network. In particular, the concept of *network motifs* and *graphlets* has proven very useful in the area of computational biology.

A *network motif* is a subgraph (not necessarily induced and possibly labeled) that appears with a significantly higher frequency in a real-world network than one would expect by pure chance. Introduced in [62] under the hypothesis that such frequently occurring structures have a functional significance, motifs have been identified in a plethora of different domains— including protein-protein-interaction networks [5], brain networks [83] and electronic circuits [40]. We point the interested reader to the surveys of Kaiser, Ribeiro and Silva [80] and Masoudi-Nejad, Schreiber, and Kashani [57] for a more extensive overview.

Graphlets are a related concept, though their application is in an entirely different scope. While motifs are used to identify and explain local structure in networks, graphlets are used to 'fingerprint' them. Pržulj [79] introduced the *graphlet degree distribution* as a way of measuring network similarity. To compute it, one enumerates all connected graphs up to a fixed size (five in the original paper) and computes for each vertex of the target graph how often it appears in a subgraph isomorphic to one of those patterns. Since some graphlets exhibit higher symmetry than others, the computation takes into
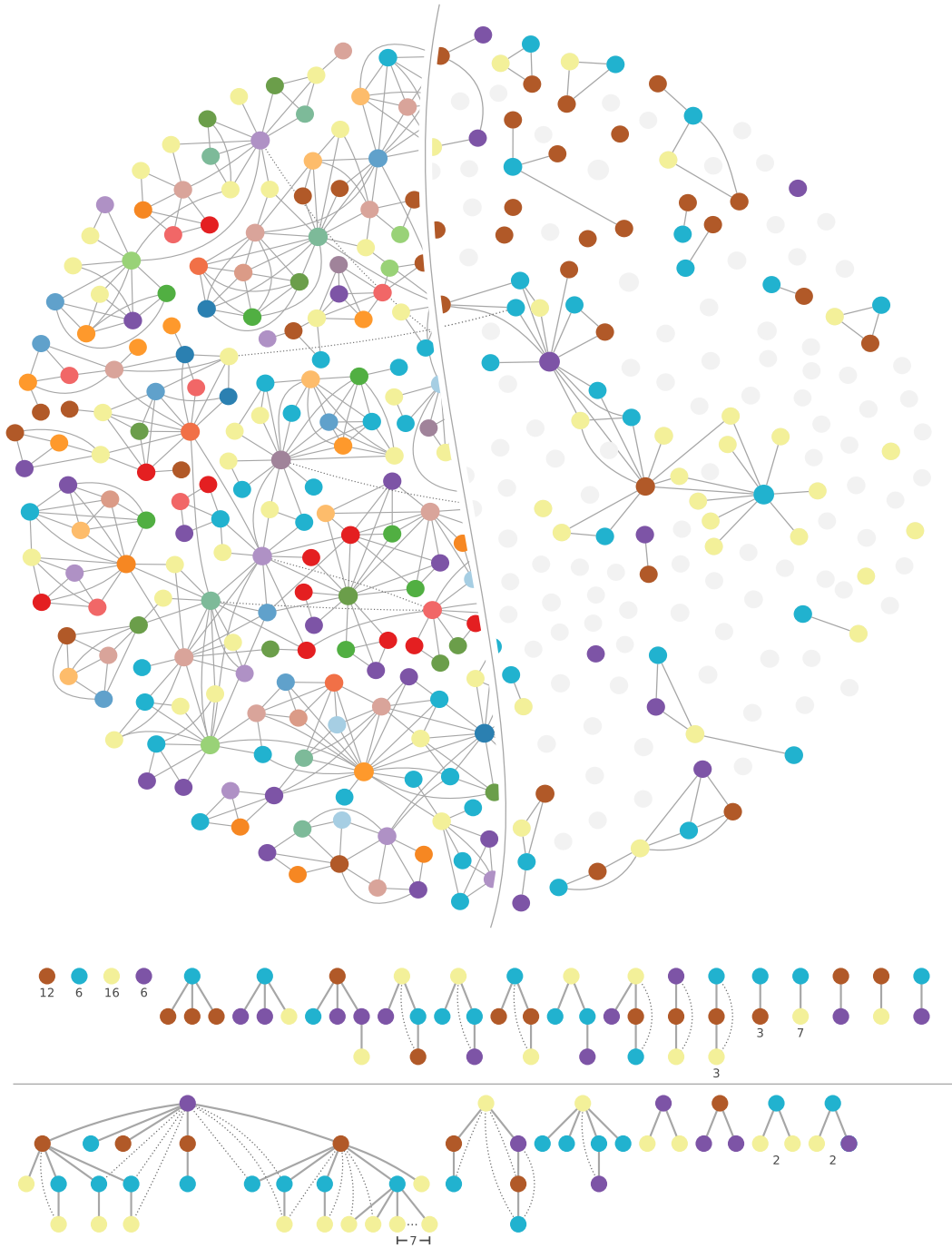
Figure 2: A 5-centered coloring (using 21 colors) of a real-world social network, Newman's Network Science [73] (giant component shown), which represents co-authorships between researchers in the field of network science as of 2006. The right half is restricted to a a subgraph formed by 4 color classes. Below, the corresponding representation by trees of depth ≤ 4 (with multiplicities noted).

account all possible automorphisms. The degree distribution then describes for each graphlet G_i, how many vertices of the target graph are contained in $0, 1, 2, \ldots$ subgraphs isomorphic to G_i—more precisely, in how many orbits of the respective automorphism groups it appears in. Note that if the set of graphlets only contains the single-edge graph this computation yields exactly the classical degree distribution.

The application of this distribution is two-fold: On the one hand, it can be used to measure similarity of multiple networks, in particular, networks depicting biological data [38]. On the other hand, the local structure around a vertex can reveal domain-specific functions. This is the case for protein-protein interaction networks where local structure correlates with biological activity [60], which has been applied to identify cancer genes [59] and construct phylogenetic trees [49]. Graphlets have further been used in analysis of workplace dynamics [87], photo cropping [17] and DoS attack detection [78].

A third application of subgraph counting was given by Ugander et al. [88]: their empirical analysis and subsequent modeling of social networks revealed that there is an inherent bias towards the occurrence of certain subgraphs. Thus the frequencies of constant-sized subgraphs seems an important indicator for the social domain, similar to the role of graphlet frequencies in biological networks.

In Theorem 18.9 from [67] it was shown that for a graph class of bounded expansion counting the number of satisfying assignments of a fixed boolean query is possible in linear time on a labeled graph. This implies that (labeled) graphlet and motif counting are linear time on a graph classes of bounded expansion. This result is achieved by using the algorithm presented in Lemma 17.3 in [67] to count the number of satisfying assignments of a fixed boolean query parameterized by treedepth. The running time of this algorithm is $O(2^{ht} \cdot ht \cdot n)$, where h is the number of nodes in the graphlet or motif and t is the depth of a treedepth decomposition of the graph. We provide an algorithm with a running time of $O(8^h \cdot t^h \cdot h^2 \cdot n)$. This achieves a better running time when used to count on graphs of bounded expansion, since then, as will be explained later, h will equal t, i.e. a constant. We discuss how with a small modification, this algorithm can also count how many times a node appears as a specific node of a specific graphlet or motif.

The tool of choice for applying a counting algorithm designed for bounded-treedepth graphs to a class of bounded expansion is a p-centered coloring: to compute the frequency of a given pattern of size k, we compute a $(k + 1)$-centered coloring of the input graph in linear time as per Proposition 3. We then enumerate all possible choice of $i < k$ colors and count the frequency of the pattern graph in the graph induced by those colors classes. As this induced subgraph has bounded treedepth, we can focus on counting a fixed

subgraph inside a target graph of treedepth at most k. It is then easy to compute the frequency for the original graph using inclusion-exclusion on the color classes.

Central to the dynamic programming we will use to count isomorphisms is the following notion of a *k-pattern* which is very similar to the well-known notion of *boundaried graphs*.

Definition 10 *A k-pattern of the graph H is a triple $M = (W, X, \pi)$ where $X \subseteq W \subseteq V(H)$ and $|X| \leq k$ and $\pi \colon X \to [k]$ is an injective function. We will call the set X the* boundary *of M. For a given k-pattern M we denote the underlying graph by $H[M] = H[W]$, the vertex set by $V(M) = W$, the boundary by $\mathrm{bd}(M) = X$ and the mapping by π^M.*

We denote by $\mathcal{P}_k(H)$ the set of all k-patterns of H. Note that every k-pattern (W, X, π) is also a $(k + 1)$-pattern. In the following we denote by $|H| = |V(H)|$.

Lemma 9 *Let H be a graph. Then $|\mathcal{P}_k(H)| \leq 2^{2|H| + |H| \log k} \leq 2^{2|H|} \cdot k^{|H|}$.*

Proof: The graph H has $2^{|H|}$ possible vertex subsets, each with at most $2^{|H|}$ possible choices for a boundary. The number of ways an injective mapping for a boundary of size $b \leq |H|$ into $[k]$ can be chosen is bounded by $k^{|H|} = 2^{|H| \log k}$. In total the size of $\mathcal{P}_k(H)$ is always less than $2^{2|H| + |H| \log k}$. $\qquad \square$

The following definition show how k-patterns will be used structural, namely by gluing them together or by demoting a boundary-vertex to a simple vertex. These operations will later be used for dynamic programming.

Definition 11 (k-pattern join) *Let G be a graph and $M_1 = (W_1, X_1, \pi_1)$, $M_2 = (W_2, X_2, \pi_2)$ k-patterns of G. Then the two patterns are* compatible *if $W_1 \cap W_2 = X_1 = X_2$ and for all $v \in X_1$ it holds that $\pi_1(v) = \pi_2(v)$. Their join is defined as the k-pattern $M_1 \oplus M_2 = (W_1 \cup W_2, X_1, \pi_1)$.*

Definition 12 (k-pattern forget) *Let G be a graph, let $M = (W, X, \pi)$ be a k-pattern of G and $i \in [k]$. Then the* forget operation *is the k-pattern*

$$M \ominus i = \begin{cases} (W, X \setminus \pi^{-1}(i), \pi|_{X \setminus \pi^{-1}(i)}) & \text{if } \pi^{-1}(i) \neq \varnothing \\ (W, X, \pi) & \text{otherwise} \end{cases}$$

Structurally, the k-pattern's boundaries will represent vertices from the path of the root vertex to the currently considered vertex in the treedepth decomposition, while the remaining vertices of the pattern represent vertices somewhere below it. The following two notations help expressing these properties.

Definition 13 (Subtree and root path) *Let T be a treedepth decomposition of G rooted at $r \in G$ and let $v \in V(G)$ be a vertex. Then the* subtree *of v is the subtree T_v of T rooted at v. The* root path *of v is the unique path P_v from the root r to v in T. We let $P_v[i]$ denote the ith vertex of the path (starting at the root), so that $P_v[1] = r$ and $P_v[|P_v|] = v$.*

We can now state the main lemma. The proof contains the description of the dynamic programming which works bottom-up on the vertices of the given treedepth decomposition (i.e. starting at the leaves and working towards the root of the decomposition).

Lemma 10 *Let H be a fixed graph on h vertices. Given a graph G on n vertices and a treedepth decomposition T of height t, one can compute the number of isomorphisms from H to induced subgraphs of G in time $O(8^h \cdot t^h \cdot h^2 \cdot n)$ and space $O(4^h \cdot t^h \cdot ht \cdot \log n)$.*

Proof: We provide the following induction that easily lends itself to dynamic programming over T. Denote by $M_H = (V(H), \varnothing, \varepsilon)$ the trivial t-pattern of H, here $\varepsilon \colon \varnothing \to \varnothing$ denotes the null function. Consider a set of vertices $v_1, v_2, \ldots, v_\ell \in G$ with a common parent v in T with respective subtrees T_{v_i} and root paths P_{v_i} for $1 \le i \le \ell$. Note that the root paths $P_{v_1}, \ldots, P_{v_\ell}$ all have the same length k and share the path P_v as a common prefix.

Let M_1 be a fixed t-pattern of H. We define the mapping $\psi_v^{M_1} \colon \mathrm{bd}(M_1) \to P_v$ via $\psi_v^{M_1}(v) = P_v[\pi^{M_1}(v)]$ which takes the pattern's boundary and maps it to the vertices of the root-path.

For patterns M_1 that satisfy that for all $u \in \mathrm{bd}(M_1)$, $\pi^{M_1}[u] \le l$, we denote by $f[v_1, \ldots, v_\ell][M_1]$ the number of isomorphisms $\phi_1 \colon V(M_1) \to V(G)$ such that

(i) $\phi_1|_{\mathrm{bd}(M_1)} = \psi_v^{M_1}$

(ii) $\phi_1(V(M_1) \setminus \mathrm{bd}(M_1)) \subseteq G[V(T_{v_1} \cup \cdots \cup T_{v_\ell})]$.

In other words we charge subgraphs to patterns whose boundaries lie on the shared root-path P_v, such that the labeling of the boundary coincides with the numbering induced by P_v while the rest of the pattern is contained entirely in the subtree below v.

Let r be the root of the treedepth decomposition. By the above definition, $f[r][M_H]$ counts exactly the number of isomorphisms of H into subgraphs of G.

We will show now how we can compute $f[r][M_H]$ recursively. For a leaf $v \in T$ and a t-pattern $M_1 = (W_1, X_1, \pi_1) \in \mathcal{P}_k(H)$ we compute $f[v][M_1]$ as follows: Defined the value $p_v^{M_1}$ to be 1 if the function $\psi \colon W_1 \to P_v$ defined as

$\psi(w) = P_v[\pi_1[w]]$ is an isomorphism from $H[W_1]$ to $G[\psi(W)]$ and 0 otherwise. In particular, $p_v^{M_1}$ will be zero if $W_1 \neq X_1$ or $|W_1| > P_v$. Then for the leaf v, we compute

$$f[v][M_1] = \sum_{M_2 \ominus |P_v| = M_1} p_v^{M_2}$$

where $M_2 \in \mathcal{P}_t(H)$.

The following recursive definitions show how $f[\cdot][M_1]$ can be computed for all inner vertices of T.

$$f[v][M_1] = \sum_{M_2 \ominus |P_v| = M_1} f[v_1, \ldots, v_\ell][M_2] \qquad \text{(forget)}$$

$$f[v_1, \ldots, v_{j-1}, v_j][M_1] = \sum_{M_2 \oplus M_3 = M_1} f[v_1, \ldots, v_{j-1}][M_2] \cdot f[v_j][M_3] \qquad \text{(join)}$$

where $M_2, M_3 \in \mathcal{P}_t(H)$.

We need to prove that the table f correctly reflects the number of isomorphisms to subgraphs satisfying properties i and ii.

Consider the *join*-case first: fix a pattern $M_1 \in \mathcal{P}_t(H)$. By induction, the entries $f[v_1, \ldots, v_{j-1}][\cdot]$ and $f[v_j][\cdot]$ correspond to the number of isomorphisms to subgraphs that satisfy properties i and ii with the node tuples $v_1, \ldots, v_{j-1}$ and v_j, respectively. We need to show that $f[v_1, \ldots, v_j][M_1]$ as computed above counts the number of isomorphisms from $H[M_1]$ to subgraph of G such that $\phi_1|_{\text{bd}(M_1)} = \psi_v^{M_1}$ and $\phi_1(V(M_1) \setminus \text{bd}(M_1)) \subseteq G[V(T_{v_1} \cup \cdots \cup T_{v_j})]$.

Consider the set Φ_1 of all isomorphisms from $H[M_1]$ to subgraphs of G satisfying properties i and ii for the vertex tuple $v_1, \ldots, v_j$. For any vertex subset $R \subseteq V(M_1) \setminus \text{bd}(M_1)$, define the slice $\Phi_1(R) \subseteq \Phi_1$ as those isomorphisms ϕ that satisfy $\phi^{-1}(\phi(V(H)) \cap T_{v_j}) = R$. Let $L = (V(M_1) \setminus \text{bd}(M_1)) \setminus R$ and define the patterns $M_L = (L \cup \text{bd}(M_1), \text{bd}(M_1), \pi^{M_1})$ and $M_R = (R \cup \text{bd}(M_1), \text{bd}(M_1), \pi^{M_1})$. Then by induction $|\Phi_1(R)| = f[v_1, \ldots, v_{j-1}][M_L] \cdot f[v_j][M_R]$. Since $M_1 = M_L \oplus M_R$ and clearly $M_L, M_R \in \mathcal{P}_t(H)$, the sum computes exactly $\sum_{R \subseteq V(M_1) \setminus \text{bd}(M_1)} |\phi_1(R)| = |\phi_1|$.

Next, consider the *forget*-case. Again, fix $M_1 \in \mathcal{P}_t(H)$ and let u be the parent of v in T. Let Φ_1 be the set of those isomorphisms from $H[M_1]$ to subgraphs of G for which $\phi_1|_{\text{bd}(M_1)} = \psi_u^{M_1}$ and $\phi_1(V(M_1) \setminus \text{bd}(M_1)) \subseteq G[V(T_v)]$. We partition Φ_1 into $\Phi_1 = \Phi_{1,v} \cup \Phi_{1,\bar{v}}$ where $\Phi_{1,v}$ contains those isomorphisms ϕ for which $\phi^{-1}(v) \neq \varnothing$ and $\Phi_{1,\bar{v}}$ the rest. Since $|\Phi_{1,\bar{v}}| = f[v_1, \ldots, v_\ell][M_1]$ we focus on $\Phi_{1,v}$ in the following. For $w \in V(M_1) \setminus \text{bd}(M_1)$, define $\Phi_{1,v}(w)$ as the set of those isomorphisms ϕ for which $\phi(w) = v$. Clearly, $\{\Phi_{1,v}(w) \mid w \in V(M_1) \setminus \text{bd}(M_1)\}$ is a partition of $\Phi_{1,v}$. Define the pattern $M_w = (V(M_1), \text{bd}(M_1) \cup \{w\}, \pi_w^{M_1})$ where $\pi_w^{M_1}$ is π^{M_1} augmented with the

value $\pi_w^{M_1}(v) = |P_v|$. Note that by construction $M_1 = M_w \ominus |P_v|$. By induction, $|\Phi_{1,v}(w)| = f[v_1, \ldots, v_\ell][M_w]$ and therefore

$$|\Phi_1| = |\Phi_{1,\bar{v}}| + \sum_{w \in V(M_1) \setminus \mathrm{bd}(M_1)} |\Phi_{1,v}(w)| = \sum_{M_2 \ominus |P_v|} f[v_1, \ldots, v_\ell][M_2]$$

It remains to prove the claimed running time. Initialization of f for a leaf takes time $O(|\mathcal{P}_t(H)|h^2)$ since we need to test whether the function ψ defined above is an isomorphism for each pattern $M_1 \in \mathcal{P}_t(H)$.

For the other vertices, a forget operation can be achieved in time $O(|\mathcal{P}_t(H)|)$ per vertex by enumerating all t-patterns, performing the forget operation and looking up the count of the resulting pattern in the previous table.

A join operation needs time $O(|\mathcal{P}_t(H)| \cdot h \cdot 2^h)$ per vertex, since for a given pattern M_1 those patterns M_2, M_3 with $M_1 = M_2 \oplus M_3$ are uniquely determined by partitions of the set $V(M_1) \setminus \mathrm{bd}(M_1)$.

In total the running time of the whole algorithm is $O(|\mathcal{P}_t(H)| \cdot 2^h \cdot h^2 \cdot n)$. Note that we only have to keep at most $O(t)$ tables in memory, each of which contains the occurrence of up to $|P_t(H)|$ patterns stored in numbers up to n^h. Thus in total the space complexity is $O(|\mathcal{P}_t(H)| \cdot t \cdot \log(n^h)) = O(|\mathcal{P}_t(H)| \cdot ht \cdot \log n)$. $\qquad \square$

To count the occurrences of H as an induced subgraph instead the number of subgraph isomorphisms, one can simply determine the number of automorphisms of H in time $2^{O(\sqrt{h \log h})}$ [7, 58] and divide the total count by this value (since this preprocessing time is dominated by our running time we will not mention it in the following). Counting isomorphism to non-induced subgraphs can be done in the same time and space by changing the initialization on the leaves, such that it checks for an subgraph instead of an induced subgraph. Dividing again by the number of automorphisms gives the number of subgraphs. By allowing the mapping of the patterns to map several nodes to the same value, we can use them to represent homomorphisms. Testing the leaves accordingly the same algorithm can be used to count the number of homomorphisms from H to subgraphs of G. By keeping all tables in memory, thus sacrificing the logarithmic space complexity, and using backtracking we can also label every node with the number of times it appears as a certain vertex of H.

From these observations and Lemma 9 we arrive at the following theorem:

Theorem 6 *Given a graph H on h vertices, a graph G on n vertices and a treedepth decomposition of G of height t, one can compute the number of isomorphisms from H to subgraphs of G, homomorphisms from H to subgraphs of G, or (induced) subgraphs of G isomorphic to H in time $O(8^h \cdot t^h \cdot h^2 \cdot n)$ and space $O(4^h \cdot t^h \cdot ht \cdot \log n)$.*

Note that for graphs of unbounded treedepth the running time of the algorithm degenerates to $O(8^h \cdot h^2 \cdot n^{h+1})$, which is comparable to the running time of $2^{O(\sqrt{h \log h})} \cdot n^h$ of the trivial counting algorithm.

Theorem 7 *Given a graph H and a a graph G belonging to a class of bounded expansion, there exists an algorithm to count the appearances of H as a subgraph of G in time*

$$O\left(\binom{f(h)}{h} \cdot 8^h \cdot h^{h+2} \cdot n\right)$$

where f is a function depending only on the graph class.

This immediately extends to nowhere dense classes, which have low treedepth-colorings with at most n^ε colors (for sufficiently large graphs) for any $\varepsilon > 0$. Choosing the graphs large enough and setting $\varepsilon' = \varepsilon/h$, we can bound the term $\binom{f(h)}{h}$ by $n^{\varepsilon' \cdot h} = n^\varepsilon$.

Theorem 8 *Let $\mathcal{G}$ be a nowhere-dense class and let H be a graph. For every $\varepsilon > 0$ there exists $N_\varepsilon \in \mathbf{N}$, such that for any graph $G \in \mathcal{G}, |G| > N_\varepsilon$ there exists an algorithm to count the appearances of H as a subgraph of G in time*

$$O\left(8^h \cdot h^{h+2} n^{1+\varepsilon}\right)$$

where f is a function depending only on the graph class.

Finally, we would like to point out that this counting algorithm is trivially parallelizable.

6.2 Localized Centrality

Centrality is a notion used to ascribe the relative importance of a vertex in the network. A centrality measure is a real-valued function that assigns each vertex of the network some value with the understanding that higher values correspond to more central vertices. Depending on the application, "central"vertices need not be those with high degree (for example, a cut-vertex may have high centrality as it is the only way for information to flow between two large subgraphs). There have been a wide variety of centrality scores introduced in the literature, including degree centrality, closeness centrality, eigenvector centrality, betweenness centrality and others. For a comprehensive introduction to centrality measures in social networks see, for instance, [30, 31]. There are several recent articles devoted to the topic of centrality measures in general [47, 77]. In this section, we consider localized variants of measures

Measure	Definition	Localized
Closeness [82]	$c_C(v) = (\sum\limits_{u \in V(G)} d(v,u))^{-1}$	$c_C^r(v) = (\sum\limits_{u \in N^r(v)} d(v,u))^{-1}$
Harmonic [75]	$c_H(v) = \sum\limits_{u \in V(G)} d(v,u)^{-1}$	$c_H^r(v) = \sum\limits_{u \in N^r(v)} d(v,u)^{-1}$
Lin's index [54]	$c_L(v) = \dfrac{\lvert \{v \mid d(v,v) < \infty\} \rvert^2}{\sum_{u \in V(G):d(v,u)<\infty} d(v,u)}$	$c_L^r(v) = \dfrac{\lvert N^r[v] \rvert^2}{\sum_{u \in N^r[v]} d(v,u)}$

Table 3: Distance-based centrality measures with localized variants that can be computed in linear time on graphs of bounded expansion.

similar to the *closeness centrality* introduced by Sabidussi [82] (see Table 3). The global versions of all these measures require one to compute the distance between all vertex pairs in the network, a sub-routine where the fastest known algorithm (due to Brandes [15]) is $O(n(n+m))$ which in the context of sparse networks reduces to quadratic time.

In these localized variants, we compute the measure of a vertex with respect to its rth neighborhood rather than with respect to the whole graph. We give linear time algorithms for computing these measures on graphs of bounded expansion for every constant r. As the value of r increases, the value of the measure computed for a vertex approaches its unlocalized variant at the expense of an increase in running time.[3] The measures in question and their localized variants are listed in Table 3.

One natural question is the utility of localized variants (and their accuracy in reflecting the global measure). We remark that Marsden demonstrated that for some networks, calculating the measure for a vertex v inside its closed neighborhood $G[N[v]]$ can be used as a viable substitute for the full measure [56]. In the context of computer networks, Pantazopoulos et al. [76] consider local variants which lend themselves to distributed computing and found a close correlation to the full measures on a sample of networks. We can show experimentally that our variants reliably capture the top ten percent in (arbitrarily) selected networks of our real-world corpus. To that end, we compare the top 10 percent as identified by our localized variants to those 10 percent identified by the respective full centrality measure. The results in Figure 3 suggest that already a value of r equal to about half the diameter yields very good results across all three measures. Note that we do not compare rankings; but rather only the difference between the sets of the

---

[3]For general graphs, we can compute these localized variants in time $O(n(n+m))$, by performing a breadth-first search from every vertex, for instance. We do not know whether a better running time is possible.
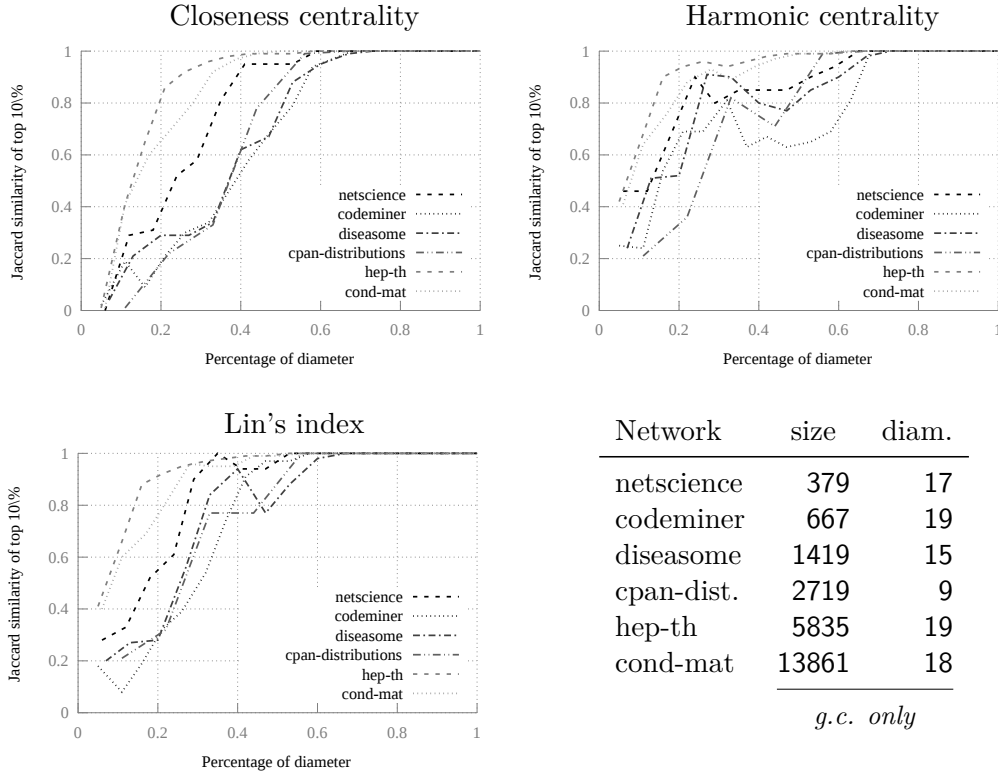
Figure 3: Quality of localized centrality measures in terms of similarity of the top ten percent vertices against the full centrality measure. Measurements were taken only in the giant component of the networks displayed in the table.

identified top vertices—our experiments showed that rank ordering is not preserved reliably. Furthermore, there seems to be a slight positive tendency towards the localized version being better in larger networks, though it is hard to draw any conclusions on a small experiment like that.

While the localized Lin's index and the localized harmonic closeness work as depicted in Table 3, the localized closeness needs a small normalization tweak in order to yield good results: this is achieved by treating the $(r+1)$st neighborhood of very vertex as if it would contain all remaining vertices and adding this value accordingly (this obviously does not change the running time of the algorithm).

Let G be a graph from a class of bounded expansion and let $\vec{G}_r$ be the directed graph with in-degree bounded by $f(r)$, for some function f, that is obtained from G as by Proposition 4 (given at the beginning of Section 6. In the following, we let $N_r^-(v)$ denote the in-neighborhood of the vertex v in the directed graph $\vec{G}_r$. We assume that the vertices of G are ordered so

44

that every vertex set has a unique representation as a tuple and, by slight abuse of notation, we use both representations interchangeably. For $v \in V(G)$ and $A = (a_1, a_2, \ldots, a_p) \subseteq N_r^-(v)$, define the *distance vector* from v to A as $\text{dist}(v, A) := (\omega(a_1 v), \omega(a_2 v), \ldots, \omega(a_p v))$, where ω denotes the arc-labeling from Proposition 4. Since $a_i \in N_r^-(v)$, $\omega(a_i, v) = d_G(a_i, v)$.

Definition 14 *Let $v \in \vec{G}_r$, $\varnothing \neq X \subseteq N_r^-(v)$, $\alpha : V(G) \to \mathbf{R}$ a vertex weighting and let $\bar{d} \in [r]^{|X|}$ be a distance vector. We define*

$$N(v, X, \bar{d}) := \{v \neq u \in V(G) \mid N_r^-(v) \cap N_r^-(u) = X \text{ and } \text{dist}(u, X) = \bar{d}\}$$

as those vertices whose in-neighborhood in $\vec{G}_r$ overlap with the in-neighborhood of v in exactly X and whose distance-vector to X is exactly $\bar{d}$. Then the query-function c_α is defined as

$$c_\alpha(v, X, \bar{d}) := \sum_{u \in N(v, X, \bar{d})} \alpha(u).$$

Lemma 11 *Given $\vec{G}_r$, one can compute a data structure in time $O(n)$ such that queries $c_\alpha(v, X, \bar{d})$ as in Definition 14 can be answered in constant time.*

Proof: We define an auxiliary dictionary R indexed by vertex sets $X \subseteq N^-r(v)$, for some vertex v. At each entry $v \in \vec{G}_r$, we will store another dictionary indexed by distance vectors which in turn stores a simple counter. We initialize R as follows: for every $v \in \vec{G}_r, X \subseteq N_r^-(v)$ and every distance vector $\bar{d} \in [r]^{|X|}$, set $R[X][\bar{d}] = 0$. Note that in total, R contains only $O(n)$ entries since all in-neighborhoods in $\vec{G}_r$ have constant size. We can implement R as a hash-map to achieve the desired (expected) constant-time for insertion and look-up, though this would yield a randomized algorithm. A possible way to implement R on a RAM deterministically is the following: We store the key $X = \{x_1, x_2, \ldots, x_p\}$ at address $x_1 + n \cdot x_2 + \ldots + n^p \cdot x_p$. This uses addresses up to size n^c, for some constant c, but since we only insert $O(n)$ keys the setup takes only linear time. Our later queries to R will be restricted to keys that are guaranteed to be contained in the dictionary, thus we will never visit a register that has not been initialized.

Now, for every $v \in \vec{G}_r$, every $X \subseteq N_r^-(v)$, increment the counter $R[X][\text{dist}(v, X)]$ by $\alpha(v)$. We now claim that queries of the form $c_\alpha(v, X, \bar{d})$ can be computed using inclusion-exclusion as follows:

$$c_\alpha(v, X, \bar{d}) = \sum_{X \subseteq Y \subseteq N_r^-(v)} (-1)^{|Y \setminus X|} \sum_{\bar{d}' : \bar{d}'|_X = \bar{d}} R[Y][\bar{d}'].$$

The computation of the sum clearly takes constant time[4]. We now prove that it indeed computes the quantity $c_\alpha(v, X, \bar{d})$.

Consider a vertex $u \in \vec{G}_r$, such that $N_r^-(u) \cap N_r^-(v) = X$ and $\mathrm{dist}(u, X) = \bar{d}$. We argue that $\alpha(u)$ is counted once by the above sum: $\alpha(u)$ is not counted by any $R[Y][\cdot]$ with $Y \supsetneq X$, therefore only the term where $X = Y$ counts $\alpha(u)$ and does so exactly once. It remains to be shown that the weight of vertices that do not conform with Definition 14 are either not counted by the sum or cancel out.

Consider a vertex $w \in \vec{G}_r$ with such that $\mathrm{dist}(w, X) \neq \bar{d}$. The weight of such a vertex is not counted by the above sum, since $\alpha(w)$ is only counted in entries of R that do not occur as summands.

Finally, consider a vertex $w' \in \vec{G}_r$ with $N_r^-(w') \cap N^-r(v) = Z$ where $X \subsetneq Z \subseteq N_r^-(v)$ and such that $\mathrm{dist}(w', X) = \bar{d}$. The weight of this vertex is counted in each term of

$$\sum_{X \subseteq Y \subseteq Z} (-1)^{|Y \setminus X|} R[Y][\mathrm{dist}(w', Z)|_Y]$$

since

$$\sum_{X \subseteq Y \subseteq Z} (-1)^{|Y \setminus X|} = \sum_{0 \leq k \leq n} (-1)^k \binom{n}{k} = 0$$

we know that the signs cancel out and thus $\alpha(w')$ does not contribute to $c_\alpha(v, X, \bar{d})$. Hence the above sum computes exactly the query $c_\alpha(v, X, \bar{d})$. $\square$

Theorem 9 *Let $\mathcal{G}$ be a graph class of bounded expansion, $G \in \mathcal{G}$ a graph and $r \in \mathbf{N}$ an integer. Then one can compute the quantities $\alpha_d(v) = \sum_{w \in N^d(v)} \alpha(w)$ for all $v \in G, d \leq r$ in linear time.*

Proof: By Theorem 4 we can compute $\vec{G}_r$ in linear time, thus we can employ Lemma 11 to answer queries as defined in Definition 14 in constant time. To compute the quantity $\alpha_d(v)$ for all $0 < d \leq r$ and $v \in V(G)$, we proceed as follows. Initialize an array C by setting $C[v][d] = 0$ for every $v \in G, 1 < d \leq r$.

Now for every $v \in V(G)$, every $X \subseteq N_r^-(v)$ and every distance vector $\bar{d} \in [r]^{|X|}$, update C via

$$C[v][\min(\bar{d} + \mathrm{dist}(v, X))] \leftarrow C[v][\min(\bar{d} + \mathrm{dist}(v, X))] + c_\alpha(v, X, \bar{d})$$

and then apply the correction

$$C[v][\min(\mathrm{dist}(v, X) + \mathrm{dist}(v, X))] \leftarrow C[v][\min(\mathrm{dist}(v, X) + \mathrm{dist}(v, X))] - 1$$

---

[4]We tacitly assume that the weights α only assign numbers polynomially bounded by the size of the graph.

in both cases with the convention that we dismiss entries where $\min(\bar{d} + \operatorname{dist}(v, X)) > r$. The second case corrects for the fact that the query $c_\alpha(v, N^-(v), \operatorname{dist}(v, N^-(v)))$ will count the vertex v itself.

At this point, $C[v][d]$ contains the sum of weights of vertices u for which $\min(\operatorname{dist}(v, X) + \operatorname{dist}(u, X)) = d$ where $X = N_r^-(v) \cap N_r^-(u) \neq \varnothing$. This follows directly from the definition of c_α.

By Theorem 4, every pair of vertices of distance $< r$ in G either is connected by an arc or they share a common in-neighbor in $\vec{G}_r$. Accordingly, we update the values of C as follows: for every $uv \in \vec{E}(\vec{G}_r)$

- if $N_r^-(u) \cap N_r^-(v) = \varnothing$, the weights of the vertices u and v were not counted in $C[v][\cdot], C[u][\cdot]$ respectively, thus we update C via

$$C[v][\omega(uv)] \leftarrow C[v][\omega(uv)] + \alpha(u)$$
$$C[u][\omega(uv)] \leftarrow C[u][\omega(uv)] + \alpha(v)$$

- if $X = N_r^-(u) \cap N_r^-(v) \neq \varnothing$, the weights of the vertices u and v were counted in $C[v][d']$ and $C[u][d']$ for $d' = \min(\operatorname{dist}(u, X) + \operatorname{dist}(v, X))$, respectively. Since d' might be larger than $\omega(uv)$ (but cannot be smaller), we update C via

$$C[v][d'] \leftarrow C[v][d'] - \alpha(u)$$
$$C[u][d'] \leftarrow C[u][d'] - \alpha(v)$$
$$C[v][\omega(uv)] \leftarrow C[v][\omega(uv)] + \alpha(u)$$
$$C[u][\omega(uv)] \leftarrow C[u][\omega(uv)] + \alpha(v)$$

where we again ignore the update of $C[\cdot][d']$ if $d' > r$.

Note that this procedure is problematic if both uv and vu are present in the graph, since then the this correction would (wrongly) be applied twice. The simple solution is that in the case of both arcs being present we only apply the above update for that arc where the start vertex is smaller than the end vertex, i.e. to uv if $u < v$ and vu otherwise.

At this point, $C[v][d]$ contains the sum of weights of vertices u for which either

- the value $d = \min(\operatorname{dist}(v, X) + \operatorname{dist}(u, X))$ where $X = N_r^-(v) \cap N_r^-(u) \neq \varnothing$ and $uv \notin \vec{E}(\vec{G}_r)$,

- or $d = \omega(uv)$ and $uv \in \vec{E}(\vec{G}_r)$.

Thus by Theorem 4 we have that $C[v][d] = \alpha_d(v)$ for $d < r$ and $v \in G$. Since all of the above operations take time linear in $|V(G)|$, the claim follows. $\square$

If we take $\alpha(\cdot) = 1$, the above algorithm counts exactly the sizes of the dth neighborhoods of each vertex, for $d < r$. Thus it can be used to compute the r-centric centrality measures presented in Table 3.

Corollary 5 *Let $\mathcal{G}$ be a graph class of bounded expansion, $G \in \mathcal{G}$ a graph and $r \in \mathbf{N}$ an integer. Then the r-centric closeness, harmonic centrality and Lin's index can be computed for all vertices of G in total time $O(|V(G)|)$.*

7   Conclusion and Open Problems

We propose unifying structural graph algorithms with complex network analysis by searching for observable structural properties that satisfy the litmus test of enabling efficient algorithms for network analysis. We presented theoretical and empirical results that support our hypothesis that complex networks are structurally sparse in a well-defined and robust sense. Efficient algorithms are known for networks of bounded expansion [27, 37, 67], and we show that for key network analysis problems these algorithms can be even further improved. On the theoretical side, we show that several random graph models of complex networks exhibit bounded expansion with high probability, although not all do—suggesting an interesting dichotomy of networks. On the experimental side, we confirm these mathematical results, and show that many real-world complex networks additionally appear to exhibit bounded expansion as measured using specialized colorings. This new approach enables fast algorithms to analyze features including communities, centrality, and motifs while more broadly providing a rigorous framework for a deeper understanding of real-world networks and related models.

There are a plethora of random graph models specifically designed to mimic properties of complex networks. Which of these models exhibit structural sparsity (and which ones do not)? There is also room for debate about how to establish that a model will generate graphs with certain properties *in practice*. Asymptotic behavior is only a proxy, although we took care to provide details on the speed of convergence in our proofs where if possible. As exemplified by the relatively weak result about the Barabási-Albert model, the practical implications are sometimes difficult to judge.

On the algorithmic side, there are several key challenges remaining. Does there exist a better algorithm/heuristic to obtain low treedepth colorings, in particular taking into consideration the special structure of complex networks? Does a good coloring algorithm exist that provides a trade-off between the number of colors and the treedepth of subgraphs induced by few color classes? Can we compute or approximate lower bounds for either χ_p or $\widetilde{\nabla}_r$ with reasonable margins of error? Both would likely improve our current empirical

understanding of the grad of networks. We would also like to investigate whether the grad for small depths is a reliable measure to differentiate networks; both our empirical and theoretical results seem to indicate so.

Finally, algorithms exploiting low grad should be tested extensively via computational experiments, to ascertain the feasibility of applying these techniques to real-world networks.

References

[1] CPAN explorer. http://cpan-explorer.org/. Accessed: July 2009.

[2] Datasets - gephi:wiki. https://wiki.gephi.org/index.php/Datasets. Accessed: 2015-01-27.

[3] L. A. Adamic and N. Glance. The political blogosphere and the 2004 US election. In *Proceedings of the WWW-2005 Workshop on the Weblogging Ecosystem*, 2005.

[4] A. Adcock, B. D. Sullivan, and M. Mahoney. Tree-like structure in large social and information networks. In *ICDM*, pages 1–10, 2013.

[5] I. Albert and R. Albert. Conserved network motifs allow protein-protein interaction prediction. *Bioinformatics*, 20(18):3346–3352, 2004.

[6] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–97, Jan. 2002.

[7] L. Babai, W. M. Kantor, and E. M. Luks. Computational complexity and the classification of finite simple groups. In *Proc. of 24th FOCS*, pages 162–171. IEEE, 1983.

[8] F. Ball, D. Sirl, and P. Trapman. Threshold behaviour and final outcome of an epidemic on a random network with household structure. *Adv. in Appl. Probab.*, 41(3):765–796, 09 2009.

[9] A. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[10] E. A. Bender and E. R. Canfield. The asymptotic number of labeled graphs with given degree sequences. *Jour. Comb. Theory, Series A*, 24(3):296–307, 1978.

[11] B. Bollobás. *Random Graphs.* London Math. Soc. Monographs. Academic Press, 1985.

[12] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.

[13] B. Bollobás and O. Riordan. The diameter of a scale-free random graph. *Combinatorica*, 18(3):279–290, 2001.

[14] B. Bollobás, O. Riordan, J. Spencer, and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Structures and Algorithms*, 24(1):5–34, 2004.

[15] U. Brandes. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.

[16] D. Bu, Y. Zhao, L. Cai, H. Xue, X. Zhu, H. Lu, J. Zhang, S. Sun, L. Ling, N. Zhang, et al. Topological structure analysis of the protein–protein interaction network in budding yeast. *Nucleic acids research*, 31(9):2443–2450, 2003.

[17] J. Bu, C. Chen, X. Liu, M. Song, L. Zhang, and Q. Zhao. Probabilistic graphlet transfer for photo cropping. *Trans. Image Processing*, 22(2):802–815, 2013.

[18] E. Cho, S. A. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1082–1090. ACM, 2011.

[19] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 99(25):15879–15882, 2002.

[20] F. Chung and L. Lu. Connected components in random graphs with given expected degree sequences. *Annals of combinatorics*, 6(2):125–145, 2002.

[21] A. Clauset, C. R. Shalizi, and M. E. Newman. Power-law distributions in empirical data. *SIAM review*, 51(4):661–703, 2009.

[22] J. Cooper and L. Lu. Graphs with asymptotically invariant degree sequences under restriction. *Internet Mathematics*, 7(1):67–80, 2011.

[23] E. D. Demaine, F. V. Fomin, M. Hajiaghayi, and D. M. Thilikos. Subexponential parameterized algorithms on graphs of bounded-genus and H-minor-free graphs. *J. ACM*, 52(6):866–893, 2005.

[24] E. D. Demaine and M. Hajiaghayi. The bidimensionality theory and its algorithmic applications. *The Computer Journal*, 51(3):292–302, 2008.

[25] E. D. Demaine, M. Hajiaghayi, and K. Kawarabayashi. Decomposition, approximation, and coloring of odd-minor-free graphs. In *Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2010)*, pages 329–344, Austin, Texas, January 17–19 2010.

[26] E. D. Demaine, M. Hajiaghayi, and K. Kawarabayashi. Contraction decomposition in H-minor-free graphs and algorithmic applications. In *Proc. of 43rd STOC*, pages 441–450, June 6–8 2011.

[27] Z. Dvořák, D. Kráľ, and R. Thomas. Testing first-order properties for subclasses of sparse graphs. *J. ACM*, 60(5):36, 2013.

[28] M. Farrell, T. Goodrich, N. Lemons, F. Reidl, F. Sánchez Villaamil, and B. D. Sullivan. Hyperbolicity, degeneracy, and expansion of random intersection graphs. *arXiv preprint arXiv:1409.8196*, 2014.

[29] F. V. Fomin, D. Lokshtanov, S. Saurabh, and D. M. Thilikos. Bidimensionality and kernels. In *Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 503–510, Austin, Texas, 2010.

[30] L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1977.

[31] L. C. Freeman. Centrality in social networks: Conceptual clarification. *Social Networks*, 1:215–239, 1978/79.

[32] S. Gago and D. Schlatter. Bounded expansion in web graphs. *Comment. Math. Univ. Carolin*, 50(2):181–190, 2009.

[33] J. Gajarský, P. Hliněný, J. Obdržálek, S. Ordyniak, F. Reidl, P. Rossmanith, F. Sánchez Villaamil, and S. Sikdar. Kernelization using structural parameters on sparse graph classes. In *Proc. of 21st ESA*, pages 529–540, 2013.

[34] Y. Gao. Treewidth of Erdős–Rényi random graphs, random intersection graphs, and scale-free random graphs. *Disc. Appl. Math.*, 160(4):566–578, 2012.

[35] M. Girvan and M. E. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.

[36] K.-I. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal, and A.-L. Barabási. The human disease network. *Proceedings of the National Academy of Sciences*, 104(21):8685–8690, 2007.

[37] M. Grohe, S. Kreutzer, and S. Siebertz. Deciding first-order properties of nowhere dense graphs. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC)*, pages 89–98. ACM, 2014.

[38] W. Hayes, K. Sun, and N. Pržulj. Graphlet-based measures are suitable for biological network comparison. *Bioinformatics*, 29(4):483–491, 2013.

[39] P. W. Holland, K. B. Laskey, and S. Leinhardt. Stochastic blockmodels: First steps. *Social networks*, 5(2):109–137, 1983.

[40] S. Itzkovitz, R. Levitt, N. Kashtan, R. Milo, M. Itzkovitz, and U. Alon. Coarsegraining and self-dissimilarity of complex networks. *Physical Review E*, 71(1):016127, 2005.

[41] S. Janson. The probability that a random multigraph is simple. II. *Journal of Applied Probability*, 51:123–137, 2014.

[42] E. J. Kim, A. Langer, C. Paul, F. Reidl, P. Rossmanith, I. Sau, and S. Sikdar. Linear kernels and single-exponential algorithms via protrusion-decompositions. In *Proc. of 40th ICALP*, pages 613–624, 2013.

[43] J. Kleinberg. Navigation in a small world. *Nature*, 406(6798):845–845, 2000.

[44] J. Kleinberg. The small-world phenomenon: An algorithmic perspective. In *Proc. of 32nd STOC*, STOC '00, pages 163–170. ACM, 2000.

[45] B. Klimt and Y. Yang. Introducing the enron corpus. In *CEAS*, 2004.

[46] D. E. Knuth. Stanford GraphBase: A platform for combinatorial algorithms. In *Proceedings of the Fourth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 41–43, 1993.

[47] D. Koschützki, K. A. Lehmann, L. Peeters, S. Richter, D. Tenfelde-Podehl, and O. Zlotowski. Centrality indices. In *Network analysis*, volume 3418 of *Lecture Notes in Computer Science*, pages 16–61. Springer, 2005.

[48] V. Krebs. V. Krebs' Website. http://www.orgnet.com/. Accessed: 2015-01-26.

[49] O. Kuchaiev, T. Milenković, V. Memišević, W. Hayes, and N. Pržulj. Topological network alignment uncovers biological function and phylogeny. *J. R. Soc. Interface*, 7(50):1341–1354, 2010.

[50] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 1(1):2, 2007.

[51] J. Leskovec and A. Krevl. SNAP Datasets: Stanford large network dataset collection. http://snap.stanford.edu/data, June 2014.

[52] J. Leskovec, K. Lang, A. Dasgupta, and M. Mahoney. Statistical properties of community structure in large social and information networks. In *WWW '08: Proceedings of the 17th International Conference on World Wide Web*, pages 695–704, 2008.

[53] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney. Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics*, 6(1):29–123, 2009.

[54] N. Lin. *Foundations of Social Research*. McGraw-Hill New York, 1976.

[55] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson. The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations. *Behavioral Ecology and Sociobiology*, 54(4):396–405, 2003.

[56] P. V. Marsden. Egocentric and sociocentric measures of network centrality. *Social Networks*, 24(4):407–422, 2002.

[57] A. Masoudi-Nejad, F. Schreiber, and Z. R. M. Kashani. Building blocks of biological networks: a review on major network motif discovery algorithms. *IET Systems Biology*, 6:164–174, 2012.

[58] R. Mathon. A note on the graph isomorphism counting problem. *Inform. Proc. Lett.*, 8(3):131–136, 1979.

[59] T. Milenković, V. Memišević, A. K. Ganesan, and N. Pržulj. Systems-level cancer gene identification from protein interaction network topology applied to melanogenesis-related functional genomics data. *J. R. Soc. Interface*, 7(44):423–437, 2010.

[60] T. Milenković and N. Pržulj. Uncovering biological network function via graphlet degree signatures. *Cancer Informatics*, 6(6):257, 2008.

[61] S. Milgram. The small-world problem. *Psychology Today*, 61(1), 1967.

[62] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network motifs: simple building blocks of complex networks. *Science*, 298(5594):824–827, 2002.

[63] M. Molloy and B. A. Reed. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, 6(2/3):161–180, 1995.

[64] M. Molloy and B. A. Reed. The size of the giant component of a random graph with a given degree sequence. *Combin., Probab. Comput.*, 7(3):295–305, 1998.

[65] J. Nešetřil and P. Ossona de Mendez. Grad and classes with bounded expansion I. decompositions. *Eur. Jour. Comb.*, 29(3):760–776, 2008.

[66] J. Nešetřil and P. Ossona de Mendez. Grad and classes with bounded expansion II. algorithmic aspects. *Eur. Jour. Comb.*, 29(3):777–791, 2008.

[67] J. Nešetřil and P. Ossona de Mendez. *Sparsity: Graphs, Structures, and Algorithms*, volume 28 of *Algorithms & Combinatorics*. Springer, 2012.

[68] J. Nešetřil, P. Ossona de Mendez, and D. R. Wood. Characterisations and examples of graph classes with bounded expansion. *Eur. Jour. Comb.*, 33(3):350–373, 2012.

[69] M. E. Newman. Newman's network data. http://www-personal.umich.edu/~mejn/netdata/. Accessed: 2015-01-26.

[70] M. E. Newman. The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2):404–409, 2001.

[71] M. E. Newman. Random graphs as models of networks. In S. Bornholdt and H. G. Schuster, editors, *Handbook of Graphs and Networks*. Wiley-VCH, 2003.

[72] M. E. Newman. The structure and function of complex networks. *SIAM Reviews*, 45(2):167–256, 2003.

[73] M. E. Newman. Finding community structure in networks using the eigenvectors of matrices. *Physical review E*, 74(3):036104, 2006.

[74] M. E. Newman, S. H. Strogatz, and D. J. Watts. Random graphs with arbitrary degree distributions and their applications. *Physical Review E*, 64(2), 2001.

[75] T. Opsahl, F. Agneessens, and J. Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, 2010.

[76] P. Pantazopoulos, M. Karaliopoulos, and I. Stavrakakis. On the local approximations of node centrality in internet router-level topologies. In *Proc. of 7th IFIP IWSOS*, pages 1–12, 2013.

[77] N. Perra and S. Fortunato. Spectral centrality measures in complex networks. *Physical Review E*, 78, 2008.

[78] P. Pongpaibool, S. Pukkawanna, and V. Visoottiviseth. Lightweight detection of DoS attacks. In *Proc. of 15th ICON*, pages 77–82. IEEE, 2007.

[79] N. Pržulj. Biological network comparison using graphlet degree distribution. *Bioinformatics*, 23(2):177–183, 2007.

[80] P. Ribeiro, F. Silva, and M. Kaiser. Strategies for network motifs discovery. In *Proc. of 5th E-Science*, pages 80–87. IEEE, 2009.

[81] M. Ripeanu, A. Iamnitchi, and I. Foster. Mapping the Gnutella network. *IEEE internet computing*, 6(1):50–57, 2002.

[82] G. Sabidussi. The centrality index of a graph. *Psychometrika*, 31(4):581–603, 1966.

[83] O. Sporns and R. Kötter. Motifs in brain networks. *PLoS Biology*, 2(11):369, 2004.

[84] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina, and P. Vanhems. High-resolution measurements of face-to-face contact patterns in a primary school. *PLOS ONE*, 6(8):e23176, 08 2011.

[85] M. P. Stumpf, C. Wiuf, and R. M. May. Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proceedings of the National Academy of Sciences of the United States of America*, 102(12):4221–4224, 2005.

[86] S. Takemura, A. Bharioke, Z. Lu, A. Nern, S. Vitaladevuni, P. K. Rivlin, W. T. Katz, D. J. Olbris, S. M. Plaza, P. Winston, T. Zhao, J. A. Horne, R. D. Fetter, S. Takemura, K. Blazek, L.-A. Chang, O. Ogundeyi, M. A. Saunders, V. Shapiro, C. Sigmund, G. M. Rubin, L. K. Scheffer, I. A. Meinertzhagen, and D. B. Chklovskii. A visual motion detection circuit suggested by Drosophila connectomics. *Nature*, 500(7461):175–181, Aug. 2013.

[87] S. Teso, J. Staiano, B. Lepri, A. Passerini, and F. Pianesi. Ego-centric graphlets for personality and affective states recognition. In *Proc. of SocialCom 2013*, pages 874–877. IEEE, 2013.

[88] J. Ugander, L. Backstrom, and J. Kleinberg. Subgraph frequencies: Mapping the empirical and extremal geography of large graph collections. In *Proc. of the 22nd WWW*, pages 1307–1318. IW3C2, 2013.

[89] Y. J. Wang and G. Y. Wong. Stochastic blockmodels for directed graphs. *Journal of the American Statistical Association*, 82(397):8–19, 1987.

[90] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, 1998.

[91] N. C. Wormald. *Some problems in the enumeration of labelled graphs*. PhD thesis, Newcastle University, 1978.

[92] W. W. Zachary. An information flow model for conflict and fission in small groups. *Journal of anthropological research*, pages 452–473, 1977.

[93] T. Zhou, G. Yan, and B.-H. Wang. Maximal planar networks with large clustering coefficient and power-law degree distribution. *Physical Review E.*, 71(4):046141, 2005.

Appendix

Clustering and bounded expansion

One could wonder if having a graph class with constant clustering coefficient is an impediment to the graph class having bounded expansion. This does not have to be the case, as exemplified by our result on the configuration model with households and the *random appolonian networks* model, which generates planar networks with a power-law degree distribution and constant clustering coefficient in the limit [93].

It is easy to show something even stronger for graph classes with bounded expansion, namely that we can "force" any constant coefficient. Note that if G is a graph with bounded average degree d, we can obtain a constant clustering as follows: for every vertex in the graph, add d^2 edges to its neighborhood. To see that this results in a constant clustering, note that every vertex with degree at most d now has a clique as its neighborhood—thus in the resulting

graph G', the local clustering of such a vertex is necessarily equal to 1. Since the number of vertices that have degree at most the total average degree makes up a constant fraction of the whole vertex set, the resulting clustering coefficient will be bounded from below by a constant. However, we need to show that such an operation will not destroy the property of having bounded expansion.

Lemma 12 *Let $\mathcal{G}$ be a graph class. For any constant t, let $\mathcal{C}_t(\mathcal{G})$ be the class of graphs obtained from any graph of $\mathcal{G}$ by adding up to t edges to the neighborhood of every vertex. Then the grad of $\mathcal{C}_t(\mathcal{G})$ depends only on the grad of $\mathcal{G}$ and t. In particular, if $\mathcal{G}$ has bounded expansion, so does $\mathcal{C}_t(\mathcal{G})$.*

Proof: Take any graph $G \in \mathcal{G}$ and for each $v \in G$, let E_v be a set of at most t edges to be added to the neighborhood of v. Let G' be the graph obtained from G by adding the edge set $\bigcup_{v \in G} E_v$. We demonstrate how to create G' through operations that increase the grad of the graph only by some function depending on the old grad and some constants.

Let $d = 2\nabla_0(G)$ denote the degeneracy of G. We compute a proper coloring of G with $d + 1$ colors and obtain a partition $V_1, \ldots, V_{d+1}$ of the vertex set. This also naturally partitions the edge set $\bigcup_{v \in G} E_v$ into sets $E_1, \ldots, E_{d+1}$, where $E_i = \bigcup_{v \in V_i} E_v$ for $1 \leq i \leq d + 1$.

Let us demonstrate how the edges from a single set E_i can be added to G through operations that do not increase the grad too much; this sequence of operations can then be simply executed for every $1 \leq i \leq d + 1$ to obtain enough edges for the graph to have constant clustering. Thus, fix one set E_i in the following. We first take the lexicographic product $G \bullet K_{t+1}$. For each vertex $v \in G$, let us denote the 'copies' of v in $G \bullet K_{t+1}$ by the set $C_v := \{(v, u) \mid u \in K_{t+1}\}$.

By Lemma 5.1 and Lemma 5.2 in [65], this increases the grad of the graph only by some function of t and the old grad. Now, consider some vertex $v \in V_i$ in whose neighborhood the edge set $E_v \subseteq E_i$ should be added. To this end, for every edge $ab \in E_v$ we take some vertex $v' \in C_v$ and contract it into an edge between C_a and C_b (by the choice of v', those vertices are exactly determined.) Since $|E_v| \leq t$, we can do this without exhausting C_v. Furthermore, since the vertices of V_i are independent, no vertex will both create *and* receive an edge. Therefore, after contracting each sets $C_v, v \in G$ in the remaining graph—i.e. excluding the vertices that wzere contracted into edges—we obtain exactly the graph G with the edges E_i added to it. This is, by construction, a 1-shallow minor of $G \bullet K_{t+1}$. Thus this last operation cannot increase the grad arbitrarily, but only by a function of t.

Applying the above steps of operations for each edge set E_i yields the graph G with the edges $\bigcup_{v \in G} E_v$ added to it. Since d is bounded by the grad

of G, the resulting graph's grad only depends on t and the grad of G. $\qquad\square$

We therefore conclude that clustering is not at all orthogonal to bounded expansion.