

Optimal subgraph structures in scale-free configuration models

Citation for published version (APA):

van der Hofstad, R. W., van Leeuwen, J. S. H., & Stegehuis, C. (2017). Optimal subgraph structures in scale-free configuration models. *arXiv*, (1709.03466), [1709.03466].

Document status and date:

Published: 11/09/2017

Document Version:

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Optimal subgraph structures in scale-free configuration models

Remco van der Hofstad, Johan S.H. van Leeuwen and Clara Stegehuis

Department of Mathematics and Computer Science
Eindhoven University of Technology

September 12, 2017

Abstract

Subgraphs reveal information about the geometry and functionalities of complex networks. For scale-free networks with unbounded degree fluctuations, we count the number of times a small connected graph occurs as a subgraph (motif counting) or as an induced subgraph (graphlet counting). We obtain these results by analyzing the configuration model with degree exponent $\tau \in (2, 3)$ and introducing a novel class of optimization problems. For any given subgraph, the unique optimizer describes the degrees of the nodes that together span the subgraph. We find that every subgraph occurs typically between vertices with specific degree ranges. In this way, we can count and characterize *all* subgraphs. We refrain from double counting in the case of multi-edges, essentially counting the subgraphs in the *erased* configuration model.

1 Introduction

Scale-free networks often have degree distributions that follow power laws with an exponent $\tau \in (2, 3)$ [1, 11, 17, 31], hence with unbounded degree fluctuations. Many networks have been reported to satisfy these conditions, including metabolic networks, the internet and social networks, and hence come with notable characteristics such as a diverging second moment of the degree distribution and the relative commonness of hubs, vertices of extremely high-degrees. These hubs play a dominant role in the network structure and function, creating ultra-small distances, ultra-fast information spreading and resilience against random attacks.

Another scale-free property says that the clustering coefficient (the probability that two neighbors of a node are neighbors themselves) decreases with the node degree [4, 10, 21, 28, 31], again following a power law. This implies that the node degree mitigates the potential for triadic closure, and in particular hubs hardly form triangles, and can be understood by imagining the network to consist of communities of densely connected nodes, these communities being connected to each other through the hubs. Triangle relations then predominantly occur within the communities between low-degree nodes.

All networks are composed of small subgraphs. The triangle is the most studied subgraph, because it not only describes the tendency for local clustering of nodes, but also signals hierarchy and community structure [26]. However, other subgraphs such as bifans or larger cliques are equally important for understanding network organization [2, 29]. Indeed, subgraph counts might vary considerably across different networks [22, 23, 33] and any given network will have a set of statistically significant subgraphs. Statistical relevance can be expressed by comparing the real networks to some mathematically tractable null model. This comparison filters out the effect of the degree sequence and the network size on the motif count. A popular statistic takes the subgraph count, subtracts the expected number of subgraphs in a null model, and divides by the



Figure 1: In the configuration model, there are two triangles from the edge perspective (a) and only one triangle from the vertex perspective (b).

variance in the null model [12, 22, 24]. Such a standardized test statistics sheds light on whether a subgraph is overrepresented in comparison to the null model that serves as the baseline.

This raises the question of what null model to use. A natural candidate is the uniform simple graph with the same degrees as the original network. For $\tau > 3$, when the degree distribution has finite second moment, it is easy to generate such graphs using the configuration model, a random graph model that creates random graphs with any given degree sequence [5, 13]. For $\tau < 3$, however, the configuration model fails to create simple graphs with high probability, and null models usually involve rewiring edges of the original graph. Consequently, the counting of subgraphs remains mathematically intractable, and one needs to resort to algorithms for exhaustive counting of motifs [20, 32], or estimations of the number of motifs by sampling [18] which is computationally expensive.

In this paper we deal with multiple edges and self-loops by excluding double counting. Indeed, we can count subgraphs in two ways: from an edge perspective and from a vertex perspective. Figure 1 illustrates this for the triangle. From the edge perspective, multi-edges may create multiple triangles between one set of vertices. When we count triangles from the vertex perspective however, we count triangles between the same set of vertices only once, the approach we take in this paper. Counting subgraphs in the configuration model then becomes equivalent to counting subgraphs in the *erased* configuration model [8] [13, Chapter 7]. This model is based on the same algorithm as the configuration model, but then followed by the removal of all self-loops and multiple edges.

The erased configuration model is intimately connected with a second popular null model, the rank-1 inhomogeneous random graph or hidden variable model [4, 9]. In this model vertices are characterized by weights that influence the creation of edges between pairs of vertices. The model can be seen as enlarged ensembles of random graphs that can match any given degree distribution in expectation. All topological properties, including correlations and clustering, then become functions of the distribution of the weights and the probability of connecting vertices. The independence between edges makes rank-1 inhomogeneous random graphs highly tractable, while the erased configuration model presents some additional structural dependencies. Nevertheless, we prefer to use the erased configuration model, because of its connection with the configuration model and the larger flexibility in its choice of degree sequence (see [13, Chapter 7] for a discussion). Moreover, we will argue that all results we obtain in this paper for the erased configuration model also hold for the rank-1 inhomogeneous random graph.

We count the number of times a small connected graph H occurs as a subgraph (motif counting) or as an induced subgraph (graphlet counting) in an erased configuration model G with degree exponent $\tau \in (2, 3)$. Let $G = (V, E)$ be a graph, and $H = (V_H, E_H)$ be a small, connected graph. When we count graphlets H , we are interested in $N^{(\text{ind})}(H)$, the number of induced subgraphs of G that are isomorphic to H . For example, when H is a square, we count all sets of 4 vertices that contain a square in the edges between them, but no more edges than that. We also study motifs, where we count $N^{(\text{sub})}(H)$, the number of occurrences of H as a subgraph of G . For example, when H is a square, we count all squares in G , but also all complete graphs of size 4, since they also contain a square. When H is a complete graph, $N^{(\text{ind})}(H) = N^{(\text{sub})}(H)$, otherwise $N^{(\text{ind})}(H) \leq N^{(\text{sub})}(H)$. There is thus a subtle difference between graphlets and motifs.

We find that every small graph H , whether it is a graphlet or motif, occurs typically between

vertices in G with degrees in a very specific range. In this paper we show that many subgraphs consists exclusively of \sqrt{n} -degree vertices, including cliques of all sizes. Hence, in such subgraphs, the hubs (of degree close to the maximal value $n^{1/(\tau-1)}$) are unlikely to act as a vertex of a typical subgraph. Hubs can be part, however, of other subgraphs such as stars. We define optimization problems that find these optimal degree ranges for every motif and graphlet.

The erased configuration model. Given a positive integer n and a *degree sequence*, i.e., a sequence of n positive integers (D_1, D_2, \dots, D_n) , the *configuration model* is a (multi)graph where vertex i has degree D_i . It is defined as follows, see e.g., [6] or [13, Chapter 7]. Given a degree sequence \mathbf{d} with $\sum_{i \in [n]} D_i$ even, we start with d_j free half-edges adjacent to vertex j , for $j = 1, \dots, n$. The configuration model is constructed by successively pairing, uniformly at random, free half-edges into edges, until no free half-edges remain. The wonderful property of the configuration model is that, conditionally on obtaining a simple graph, the resulting graph is a *uniform* graph with the prescribed degrees. This is why the configuration model is often used as a *null model* for real-world networks with given degrees. The *erased configuration model* is the model where all multiple edges are merged and all self-loops are removed.

In this paper, we study the setting where the degree distribution has *infinite variance*. Then the number of self-loops and multiple edges in the configuration model tends to infinity in probability (see e.g., [13, Chapter 7]), so that the configuration model results in a multigraph with high probability and the number of erased edges is large [16] (yet small compared to the total number of edges). In particular, we take the degrees \mathbf{d} to be an i.i.d. sample from of a random variable D such that

$$\mathbb{P}(D = k) = Ck^{-\tau}(1 + o(1)), \quad \text{as } k \rightarrow \infty, \quad (1.1)$$

where $\tau \in (2, 3)$ so that $\mathbb{E}[D^2] = \infty$ and $\mathbb{E}[D] = \mu < \infty$. When this sample constructs a degree sequence such that the sum of the degrees is odd, we add an extra half-edge to the last vertex. This does not affect our computations. In this setting, d_{\max} is of order $n^{1/(\tau-1)}$, where d_{\max} denotes the maximal degree of the degree sequence.

Paper outline. We present our main results in Section 2, including the theorems that characterize all optimal motifs and subgraphs in terms of the solutions to optimization problems. We also apply these theorems to describe the optimal configurations of all subgraphs with 4 and 5 vertices, and present an outlook for further use of our results. We then prove the main theorems for general subgraphs in Section 3 and for \sqrt{n} -optimal subgraphs in Section 4. The proofs of some lemmas introduced along the way are deferred to Sections 5 and 6.

Notation. We use $\xrightarrow{\mathbb{P}}$ for convergence in probability. We say that a sequence of events $(\mathcal{E}_n)_{n \geq 1}$ happens with high probability (w.h.p.) if $\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{E}_n) = 1$. Furthermore, we write $f(n) = o(g(n))$ if $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$, and $f(n) = O(g(n))$ if $|f(n)|/g(n)$ is uniformly bounded, where $(g(n))_{n \geq 1}$ is nonnegative. We write $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ as well as $g(n) = O(f(n))$. We say that $X_n = O_{\mathbb{P}}(g(n))$ for a sequence of random variables $(X_n)_{n \geq 1}$ if $|X_n|/g(n)$ is a tight sequence of random variables, and $X_n = o_{\mathbb{P}}(g(n))$ if $X_n/g(n) \xrightarrow{\mathbb{P}} 0$.

2 Main results

The key insight obtained in this paper is that the creation of subgraphs is crucially affected by the following trade-off, inherently present in power-law networks: On the one hand, hubs contribute substantially to the number of subgraphs, because they are very well connected, and therefore potentially contribute to many graphlets or motifs. On the other hand, hubs are by definition rare. This should be contrasted with lower-degree vertices that occur more frequently, but typically take part in fewer connections and hence fewer subgraphs. Therefore, one may expect every subgraph to consist of a selection of nodes with specific degrees that ‘optimize’ this trade-off and hence maximize the probability that the subgraph occurs.

Write the probability that a motif or graphlet H of k vertices is created between k uniformly chosen vertices as

$$\mathbb{P}(H \text{ present}) = \sum_{\mathbf{D}} \mathbb{P}(H \text{ motif/graphlet on degrees } D_1, \dots, D_k) \mathbb{P}(D_1, \dots, D_k), \quad (2.1)$$

where the sum is over all possible degrees on k vertices $\mathbf{D} = (D_i)_{i \in k}$, and $\mathbb{P}(d_1, \dots, d_k)$ denotes the probability that a randomly chosen set of k vertices has degrees D_1, \dots, D_k . Because of the power-law degree distribution, the last term decreases with D_1, \dots, D_k . The first term in the sum, however, increases with D_1, \dots, D_k , since higher degree vertices are part of more subgraphs. We show that for every subgraph there is a specific range of D_1, \dots, D_k that gives the maximal contribution to (2.1), large enough even to completely ignore all other degree ranges.

We show that there are only four possible ranges of degrees that maximize the term inside the sum in (2.1). These ranges are constant degrees, degrees proportional to $n^{(\tau-2)/(\tau-1)}$, degrees proportional to \sqrt{n} or degrees proportional to $n^{1/(\tau-1)}$. Observe that at this stage, these four ranges are merely an ansatz; rigorous underpinning for these choices comes later. For degrees proportional to $n^{1/(\tau-1)}$, the trade-off between the abundance of low-degree vertices and the connectedness of high-degree vertices is won by the high-degree vertices. Thus, intuitively, vertices in subgraphs that have the largest contribution from the degree proportional to $n^{1/(\tau-1)}$ should have more connections inside the subgraph than other vertices in subgraph. On the other hand, for degrees that are constant or proportional to $n^{(\tau-2)/(\tau-1)}$, the trade-off is ‘won’ by the lower-degree vertices. Intuitively, we therefore expect that these vertices are less well-connected inside subgraph. Vertices with degrees proportional to \sqrt{n} form the middle ground, and are typically connected to vertices with similar degrees. There, the crucial observation is that vertices of degree of order \sqrt{n} are likely, though not certain, to have an edge between them in the erased configuration model.

2.1 An optimization problem

We now present the optimization problems that ask for maximizing the term inside the sum in (2.1), first for motifs and later for graphlets. Let $H = (V_H, E_H)$ be a small, connected graph on $k > 2$ vertices. Denote the set of vertices of H that have degree one by V_1 . Let \mathcal{P} be all partitions of $V_H \setminus V_1$ into three disjoint sets S_1, S_2, S_3 . This partition into S_1, S_2 and S_3 corresponds to these orders of magnitude: S_1 denotes the vertices with degree proportional to $n^{(\tau-2)/(\tau-1)}$, S_2 the ones with degrees proportional to $n^{1/(\tau-1)}$, and S_3 the vertices with degrees proportional to \sqrt{n} . The optimization problem finds the partition of the vertices into these three orders of magnitude such that the contribution to the number of motifs or graphlets is the largest. When a vertex in H has degree 1, its degree in the large graph G is typically small, it does not grow with n . Interestingly, vertices with degrees in these orders of magnitude are the only vertices that contribute to the number of motifs or graphlets, as we will prove later.

Given a partition \mathcal{P} , let E_{S_i} denote the number of edges in H between vertices in S_i , E_{S_i, S_j} the number of edges between vertices in S_i and S_j and $E_{S_i, 1}$ the number of edges between vertices in V_1 and S_i . We now define the optimization problem for motifs that is equivalent to optimizing the term inside the sum in (2.1)

$$B^{(\text{sub})}(H) = \max_{\mathcal{P}} \left[|S_1| - |S_2| - \frac{2E_{S_1} + E_{S_1, S_3} + E_{S_1, 1} - E_{S_2, 1}}{\tau - 1} \right]. \quad (2.2)$$

The first two terms in the optimization problem give a positive contribution for all vertices in S_1 , vertices with relatively low degree, and a negative contribution for vertices in S_2 having high degrees. Therefore, the first two terms in the optimization problem capture that high-degree vertices are rare, and low-degree vertices abundant. The last term gives a negative contribution for all edges between vertices with relatively low degrees in the motif. This captures the other part of the trade-off: high-degree vertices are much more likely to form edges with other vertices than low degree vertices. Note that $B^{(\text{sub})}(H) \geq 0$, since putting all vertices in S_3 yields zero.

For graphlets, we can define a similar optimization problem

$$B^{(\text{ind})}(H) = \max_{\mathcal{P}^{(\text{ind})}} \left[|S_1| - |S_2| - \frac{2E_{S_1} + E_{S_1, S_3} + E_{S_1, 1} - E_{S_2, 1}}{\tau - 1} \right], \quad (2.3)$$

$$\text{s.t. } (u, v) \in E_H \quad \forall u \in S_2, v \in S_2 \cup S_3, \quad (2.4)$$

where again \mathcal{P} is a partition of $V_H \setminus V_1$ into three sets. The difference with (2.2) for motifs is the extra constraint (2.4) which puts a restriction on the partitions into the three sets that are allowed. The extra constraint arises by the constraint in graphlets that H should be present as an induced subgraph, and ensures that edges that are not present in H are indeed not present in the subgraph. Again, $B^{(\text{ind})}(H) \geq 0$ because putting all vertices in S_3 is a valid solution.

We now have to solve the two above optimization problems, to find the largest contributor to the number of motifs and graphlets. We first show that indeed the optimization problems in (2.2) and (2.3) find the typical degrees of vertices for any motif and any graphlet, and we show what the relation is between the optimization problems and the scaling of the number of motifs and graphlets. We then present sharp asymptotics for a special class of subgraphs, where the optimal contribution to (2.2) or (2.3) comes from $S_3 = V_H$, hence motifs and graphlets where all typical vertices have degrees proportional to \sqrt{n} . For this class, which contains for instance cliques of all sizes, we present sharp asymptotics.

2.2 General subgraphs

Let $S_1^{(\text{sub})}, S_2^{(\text{sub})}, S_3^{(\text{sub})}$ be a maximizer of (2.2). Furthermore, for any $\alpha = (\alpha_1, \dots, \alpha_k)$ such that $\alpha_i \in [0, 1/(\tau - 1)]$, we define

$$M_n^{(\alpha)}(\varepsilon) = \{(u_1, \dots, u_k) : D_{u_i} \in [\varepsilon, 1/\varepsilon](\mu n)^{\alpha_i} \forall i \in [k]\}. \quad (2.5)$$

These are the sets of degrees such that D_1 is proportional to n^{α_1} and D_2 proportional to n^{α_2} and so on. Then, we denote the number of motifs with vertices in $M_n^{(\alpha)}(\varepsilon)$ by $N^{(\text{sub})}(H, M_n^{(\alpha)}(\varepsilon))$. Define the vector $\alpha^{(\text{sub})}$ as

$$\alpha_i^{(\text{sub})} = \begin{cases} (\tau - 2)/(\tau - 1) & i \in S_1^{(\text{sub})}, \\ 1/(\tau - 1) & i \in S_2^{(\text{sub})}, \\ \frac{1}{2} & i \in S_3^{(\text{sub})}, \\ 0 & i \in V_1. \end{cases} \quad (2.6)$$

By the interpretation of S_1, S_2 and S_3 in the optimization problem (2.2), sets of vertices in $M_n^{\alpha^{(\text{sub})}}(\varepsilon)$ intuitively form the largest contribution to the number of motifs. The next theorem shows that this is correct, and computes the scaling of the number of motifs:

Theorem 2.1 (General motifs). *Let H be a motif on k vertices such that the solution to (2.2) is unique. Then, for any $\alpha \neq \alpha^{(\text{sub})}$ and $0 < \varepsilon < 1$,*

$$\frac{N^{(\text{sub})}(H, M_n^{(\alpha)}(\varepsilon))}{N^{(\text{sub})}(H, M_n^{\alpha^{(\text{sub})}}(\varepsilon))} \xrightarrow{\mathbb{P}} 0. \quad (2.7)$$

Furthermore,

$$\frac{N^{(\text{sub})}(H, M_n^{\alpha^{(\text{sub})}}(\varepsilon))}{n^{\frac{3-\tau}{2}(k_{2+} + B^{(\text{sub})}(H)) + \frac{1}{2}k_1}} = f(\varepsilon)\Theta_{\mathbb{P}}(1) \quad (2.8)$$

for some function $f(\varepsilon)$ not depending on n . Here k_{2+} denotes the number of vertices in H of degree at least 2, and k_1 the number of degree one vertices in H .

For graphlets, we can make similar statements. Let $S_1^{(\text{ind})}, S_2^{(\text{ind})}, S_3^{(\text{ind})}$ be a maximizer of (2.3). Let $M_n^{(\alpha)}(\varepsilon)$ be as in (2.5), and define $\alpha^{(\text{ind})}$ as in (2.6), replacing $S_i^{(\text{sub})}$ by $S_i^{(\text{ind})}$. Similarly to the motifs case, vertices in $M_n^{\alpha^{(\text{ind})}}(\varepsilon)$ form the largest contribution to the total number of graphlets, as the next theorem shows.

Theorem 2.2 (General graphlets). *Let H be a connected graph on k vertices such that the solution to (2.3) is unique. Then, for any $\alpha \neq \alpha^{(\text{ind})}$ and $0 < \varepsilon < 1$,*

$$\frac{N^{(\text{ind})}(H, M_n^{(\alpha)}(\varepsilon))}{N^{(\text{sub})}(H, M_n^{(\alpha^{(\text{ind})})}(\varepsilon))} \xrightarrow{\mathbb{P}} 0. \quad (2.9)$$

Furthermore,

$$\frac{N^{(\text{sub})}(H, M_n^{(\alpha^{(\text{ind})})}(\varepsilon))}{n^{\frac{3-\tau}{2}(k_{2+} + B^{(\text{ind})}(H)) + \frac{1}{2}k_1}} = f(\varepsilon)\Theta_{\mathbb{P}}(1) \quad (2.10)$$

for some function $f(\varepsilon)$ not depending on n . Here k_{2+} denotes the number of vertices in H with degree at least 2, and k_1 denotes the number of degree 1 vertices in H .

2.3 Sharp asymptotics for \sqrt{n} subgraphs

Now we study the special class of motifs for which the unique maximum of (2.2) is $S_3 = V_H$. By the above interpretation of S_1 , S_2 and S_3 , we study motifs where the maximum contribution to the number of such motifs comes from vertices that have degrees proportional to \sqrt{n} in G . Examples of motifs that fall into this category are all complete graphs. Bipartite graphs on the other hand, do not fall into the \sqrt{n} -class motifs, since we can use the two parts of the bipartite graph as S_1 and S_2 in such a way that (2.2) results in a non-negative solution. The next theorem gives asymptotics for the number of such motifs:

Theorem 2.3 (Motifs with \sqrt{n} degrees). *Let H be a connected graph on k vertices with minimal degree 2 such that the solution to (2.2) is unique, and $B^{(\text{sub})}(H) = 0$. Then,*

$$\frac{N^{(\text{sub})}(H)}{n^{\frac{k}{2}(3-\tau)}} \xrightarrow{\mathbb{P}} A^{(\text{sub})}(H) < \infty, \quad (2.11)$$

with

$$A^{(\text{sub})}(H) = c^k \mu^{-\frac{k}{2}(\tau-1)} \int_0^\infty \cdots \int_0^\infty (x_1 \cdots x_k)^{-\tau} \prod_{(u,v) \in E_H} (1 - e^{-x_u x_v}) dx_1 \cdots dx_k. \quad (2.12)$$

We now state a similar theorem for graphlets where the unique optimal solution to (2.3) is $S_3 = V_H$. The optimization problem (2.3) is the same as (2.2) with a extra constraint, which is satisfied when $S_3 = V_H$. Therefore, if a for small graph H (2.2) is optimized for $S_3 = V_H$, then (2.3) is also optimized for $S_3 = V_H$. Thus, the graphs H for which Theorem 2.3 can be applied are a subset of the graphs for which Theorem 2.4 can be applied. Therefore, complete graphs fall into the \sqrt{n} -class graphlets as well. Section 2.4 shows which motifs on 4 and 5 vertices belong to the \sqrt{n} class. If the maximum contribution for H comes from \sqrt{n} vertices for counting motifs as well as graphlets, then $N^{(\text{ind})}(H)$ is of the same order of magnitude as $N^{(\text{sub})}(H)$.

Theorem 2.4 (Graphlets with \sqrt{n} degrees). *Let H be a connected graph on k vertices with minimal degree 2 such that the solution to (2.3) is unique, and $B^{(\text{ind})}(H) = 0$. Then,*

$$\frac{N^{(\text{ind})}(H)}{n^{\frac{k}{2}(3-\tau)}} \xrightarrow{\mathbb{P}} A^{(\text{ind})}(H) < \infty, \quad (2.13)$$

with

$$A^{(\text{ind})}(H) = c^k \mu^{-\frac{k}{2}(\tau-1)} \int_0^\infty \cdots \int_0^\infty (x_1 \cdots x_k)^{-\tau} \prod_{(u,v) \in E_H} (1 - e^{-x_u x_v}) \prod_{(u,v) \notin E_H} e^{-x_u x_v} dx_1 \cdots dx_k. \quad (2.14)$$

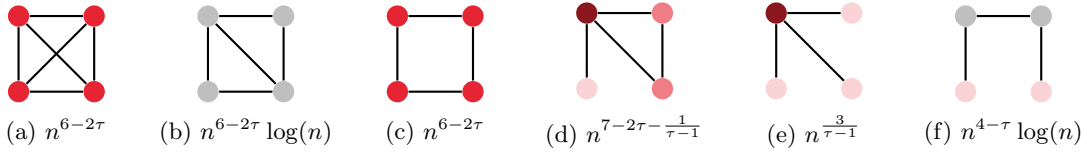


Figure 2: Graphlets. Order of magnitude of $N^{(\text{ind})}(H)$ for all connected graphs on 4 vertices (constants ignored). Vertices with degree one are light pink, vertices in S_1 are darker pink, vertices in S_3 are bright red, and vertices in S_3 are dark red. Vertices where the optimizer is not unique are gray.

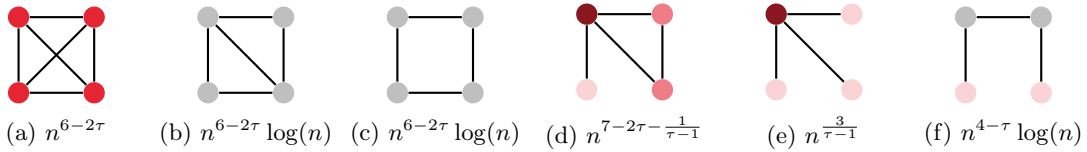


Figure 3: Motifs. Order of magnitude of $N^{(\text{sub})}(H)$ for all connected graphs on 4 vertices (constants ignored). Vertices with degree one are light pink, vertices in S_1 are darker pink, vertices in S_3 are bright red, and vertices in S_3 are dark red. Vertices where the optimizer is not unique are gray.

The difference between counting motifs and counting graphlets is visible in (2.12) and (2.14). In the erased configuration model, the probability that a vertex with degree D_i connects to a vertex with degree D_j can be approximated by $1 - e^{-D_i D_j / L_n}$, where L_n denotes the sum of all degrees. When rescaling, this results in the factors $1 - e^{-x_u x_v}$ in (2.12) for all edges in motif H . When counting graphlets, we count induced subgraphs. Then, we also have to take into account that no other edges than the edges in H are allowed to be present. This gives the extra terms $e^{-x_u x_v}$ in (2.14).

2.4 Subgraphs on 4 and 5 vertices

We now apply Theorem 2.2 to characterize the optimal subgraph configurations of motifs and graphlets that consist of 4 or 5 vertices. For every partition of the vertices of H into S_1, S_2, S_3 , we compute the contribution to (2.2) and (2.2). In this way, we can find the partitions that maximize (2.2) and (2.3), and check whether this maximum is unique. If the maximum is indeed unique, then we can use Theorems 2.1 and 2.2 to calculate the scaling of the number of such motifs or graphlets. Figures 2 and 4 show the order of magnitude of the number of graphlets on 4 and 5 vertices obtained in this way, together with the optimizing sets of (2.3). Figure 3 shows the order of magnitude of the number of motifs on 4 vertices together with the optimizing sets of (2.2). For example the optimal values of S_1, S_2 and S_3 for the motif in Figure 2(d) show that

$$B^{(\text{ind})}(H) = 2 - 1 + \frac{2 + 0 + 0 - 1}{\tau - 1} = 1 + \frac{1}{\tau - 1}. \quad (2.15)$$

By Theorem 2.2 this shows that the correct scaling of the motif in Figure 2(d) is

$$n^{(3-\tau)/2(4-1/(\tau-1))+1/2} = n^{7-2\tau-\frac{1}{\tau-1}}. \quad (2.16)$$

The scaling of the other motifs and graphlets are computed similarly.

Figures 2 and 3 show the difference between counting motifs or counting graphlets. For example, Figure 2(c) shows that a square occurs $\Theta(n^{6-2\tau})$ times as a graphlet, whereas it occurs $\Theta(n^{6-2\tau} \log(n))$ times as a motif by Figure 3(c). When we count the number of times the square occurs as a motif, we have to add the contributions from the graphlets in Figures 2(a), 2(b) and 2(c), that all contain a square. Then we also see that the square occurs $\Theta(n^{6-2\tau} \log(n))$ times as a motif. The major contribution to the number of square motifs is from graphlet 2(b).

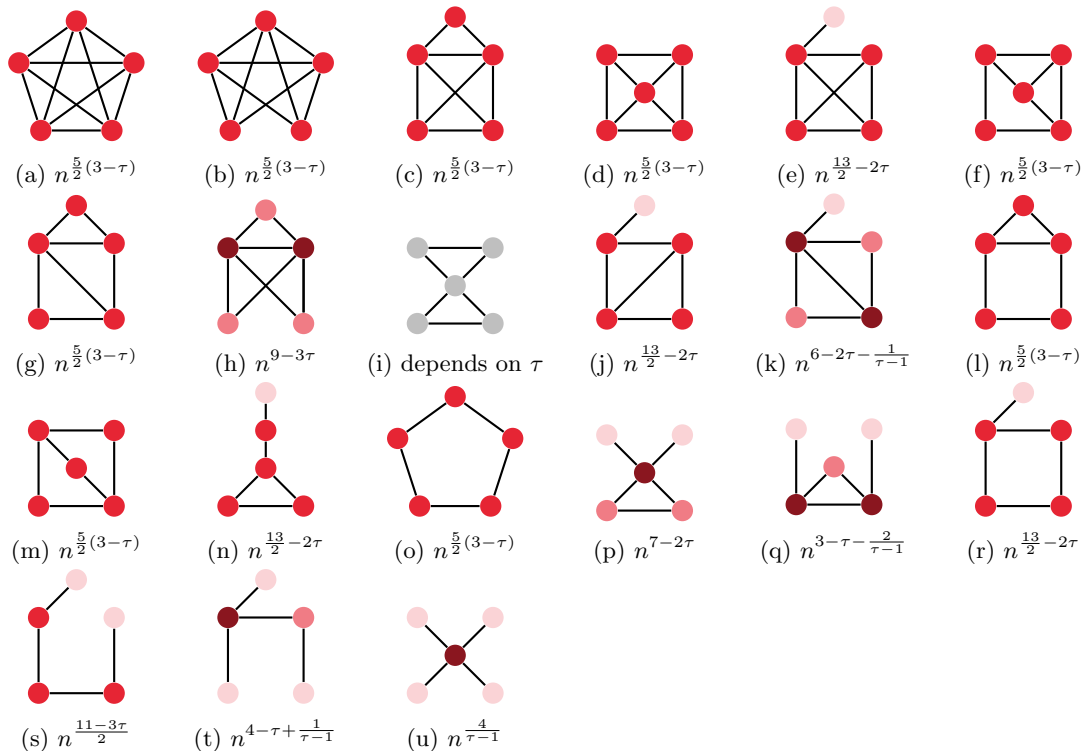


Figure 4: Graphlets. Order of magnitude of $N^{(\text{ind})}(H)$ for all connected graphs on 5 vertices (constants ignored). Vertices with degree one are light pink, vertices in S_1 are darker pink, vertices in S_3 are bright red, and vertices in S_3 are dark red. Vertices where the optimizer is not unique are gray.

This graphlet indeed contains a square, and occurs more frequently than the square occurs as a graphlet. In this manner we can infer the order of magnitude of the number of motifs from the number of graphlets. For this reason, Figure 4 is not shown for motifs. Using only Figure 4, we can argue that the graph in Figure 4(m) occurs $\Theta(n^{5/2(3-\tau)})$ times as a graphlet, but $\Theta(n^{9-3\tau})$ times as a motif. Indeed, the graph in Figure 4(h) contains Figure 4(m) as a subgraph, and occurs more frequently.

Most motifs and graphlets in Figures 2, 3 and 4 satisfy the constraint in Theorem 2.2 that the solution to the optimization problem (2.2) or (2.3) should be unique. However, the gray vertices in Figure 2 and 3 do not have unique optimizers. Still these motifs and graphlets have ranges of degrees that give the major contribution to the number of such graphlets or motifs. The only difference is that these ranges are wider than for the vertices with unique maximizers. For example, for the square motif in Figure 3(c) the major contribution is from vertices where the degrees of vertices at each side of an edge $\{i, j\}$ in the square satisfy $D_i D_j = \Theta(n)$. Note that having all degrees proportional to \sqrt{n} therefore is one of the main contributors to the square motif. However, contributions where the left bottom vertex and the right top vertex have degrees proportional to n^α and the other two vertices have degrees $n^{1-\alpha}$ give an equal contribution for other values of α . Using that $D_i D_j$ follows a power-law distribution with exponent τ with an extra factor $\log(n)$ [30] then gives the extra factor $\log(n)$ in Figure 3(b).

Another motif with gray vertices is the bow tie in Figure 4(i). Unlike the other gray graphlets, this graphlet does satisfy the constraint of Theorem 2.2 that the optimal solution to (2.2) should be unique. However, the optimal solution depends on τ . For τ small, the maximum of (2.3) is uniquely attained at 0, so that for τ small, the major contribution comes from the situation when all vertices are of degree $\Theta(\sqrt{n})$. On the other hand, when $\tau > 7/3$, this function is minimized when S_1 contains all degree 2 vertices, and the middle vertex is in S_1 . This partition gives a

contribution to (2.3) of

$$4 - 1 - \frac{2 \cdot 2}{\tau - 1} = \frac{3\tau - 7}{\tau - 1}, \quad (2.17)$$

which is larger than zero if $\tau > 7/3$. Thus, for τ larger than $7/3$, the major contribution is when the middle vertex has degree $n^{1/(\tau-1)}$, and the other vertices have degrees $n^{(\tau-2)/(\tau-1)}$. When $\tau < 7/3$, the major contribution is from all vertices of degrees \sqrt{n} , so that Theorem 2.4 can be applied.

When the maximal contribution to a graphlet comes from vertices with degrees proportional to \sqrt{n} , then by Theorem 2.4, the number of such graphlets converges to a constant when properly rescaled. When the maximal contribution contains vertices in S_2 and S_1 , this may not be true anymore. For example, counting the claw graphlet of Figure 2(e) is very similar to counting the number of ways to choose three neighbors for every vertex. The only pairs of neighbors we do not count, are neighbors that are connected themselves. This is only a small fraction of the pairs of neighbors [15], thus the number of claws is approximately equal to

$$\sum_{i \in [n]} \frac{1}{6} D_i (D_i - 1) (D_i - 2) \approx \sum_{i \in [n]} D_i^3. \quad (2.18)$$

Since the degrees are i.i.d. samples from a power-law distribution, summing D_i^3 will converge to a stable law when normalized properly. In the \sqrt{n} -degree case, the leading order of the number of motifs or graphlets is constant (see Theorems 2.3 and 2.3). In the regime where vertices of degrees proportional to $n^{1/(\tau-1)}$ contribute, the leading order may contain stable random variables. Thus, the number of graphlets where the optimal solution to (2.3) comes from \sqrt{n} -degree vertices may be less volatile than when the optimal contribution also contains vertices with degrees proportional to $n^{1/(\tau-1)}$.

2.5 Discussion and outlook

Inhomogeneous random graph. All results in this paper are proven for the erased configuration model. An interesting question is whether the results on the number of motifs and graphlets of Theorems 2.1-2.4 only hold for the erased configuration model, or whether they also apply to other models that create simple power-law random graphs. Another possible null model that creates simple graphs with power-law degrees is the rank-1 inhomogeneous random graph [4, 9]. In this model, vertices have weights h_i , where the distribution of the weights is a power-law distribution with exponent $\tau \in (2, 3)$. Then, two vertices are connected with probability $f_n(h_i, h_j)$. Two important connection probability functions are $f_n(h_i, h_j) = \min(h_i h_j / (\mu n), 1)$ [9], or $f_n(h_i, h_j) = 1 - e^{-h_i h_j / (\mu n)}$ [7]. We prove Theorems 2.1-2.4 for the erased configuration model by using the approximation $\mathbb{P}_n(X_{ij} = 1) \approx 1 - e^{-D_i D_j / L_n}$. Therefore, these theorems remain valid when we study the rank-1 inhomogeneous random graph with these connection probabilities instead.

Precise constants. For motifs and graphlets where the most likely degrees are \sqrt{n} , we show in Theorems 2.3 and 2.4 that the rescaled number of motifs or graphlets converges to a constant, $A^{(\text{sub})}(H)$ and $A^{(\text{ind})}(H)$, respectively. The constants $A^{(\text{sub})}(H)$ and $A^{(\text{ind})}(H)$ are in general difficult to compute. It would be useful to have good estimates of these constants to be able to see which types of motifs occur more frequently. Furthermore, it would be interesting to investigate the convergence of motifs that do not satisfy the assumptions of Theorems 2.3 or 2.4. In Section 2.4, we saw that the normalized number of motifs may converge to a stable distribution for some motifs.

Hyperbolic random graph. Another random graph model that creates simple power-law random graphs, is the hyperbolic random graph where nodes are sampled in a disk of radius R , and connected if their hyperbolic distance is at most R [19]. These graphs are very different from the erased configuration model and the rank-1 inhomogeneous random graph, because they contain geometry which creates more clustering. As mentioned before, all complete graphs satisfy the conditions of Theorems 2.3 and 2.4. Thus, a complete graph on k vertices occurs $\Theta(n^{\frac{k}{2}(3-\tau)})$

times as a motif or graphlet in erased configuration models. Interestingly, this is also true for hyperbolic random graphs [3]. From the construction of the proof in [3], we can argue that the largest contribution to the number of cliques in hyperbolic random graphs comes from vertices at radius $R/2$. These vertices have degrees proportional to \sqrt{n} , which is the same as the largest contribution for erased configuration model. It would be interesting to investigate the presence of other types of motifs in hyperbolic random graphs, and see whether these results are similar to the results for the erased configuration model, or if the geometric structure in these graphs makes the largest contribution to the number of motifs or graphlets different. In particular, it would be interesting to see if all other motifs that satisfy the conditions of Theorem 2.3 have the same order of magnitude in the hyperbolic random graph as in the erased configuration model.

Self-averaging or not. Another interesting question relates to the fluctuations of the subgraph counts. When the degree distribution follows a power-law with exponent $\tau \in (2, 3)$, the number of motifs may not be self-averaging [25], that is

$$\limsup_{n \rightarrow \infty} \frac{\text{Var}(N^{(\text{sub})}(H))}{\mathbb{E}[N^{(\text{sub})}(H)]^2} \neq 0. \quad (2.19)$$

One such example is the triangle. By [25], the number of triangles in a network is non self-averaging when τ is close to 3. Still, the triangle motif satisfies the conditions of Theorem 2.3, so that the number of triangles converges in probability to a constant. This indicates that when we generate a configuration model with i.i.d. degrees, most realizations will have a number of triangles that is close to the value that is predicted in Theorem 2.3. Some realizations however, will have a number of triangles that is much larger or smaller than the value predicted in Theorem 2.3, which results in a large variance. Therefore, the number of triangles is non self-averaging. However, when we first fix the degree sequence, Lemmas 5.3 and 5.4 show that the variance from the vertices of degrees \sqrt{n} is small. Similarly, we can show that the other contributions to the variance are small as well. Therefore, the fluctuation in the number of motifs arises from the i.i.d. degree sequence, which was also observed in [25] for the rank-1 inhomogeneous random graph. In particular, when we use the degrees of a real-world model as input in the erased configuration model, the number of motifs is self-averaging. This illustrates the importance of choosing the right null model. A null model with the same degree sequence as the original graph has less variability than a null model where we sample degrees i.i.d. from a power-law distribution with the same exponent as the original degree distribution.

Finding S_1, S_2, S_3 . In Section 2.4, the optimal motif or graphlet structure is found by computing (2.2) and (2.3) for all possible partitions into S_1, S_2, S_3 . For large motifs, this becomes computationally hard. It would be interesting to find other ways to optimize (2.2) and (2.3), by rewriting these optimization problems as linear programs for example.

3 Maximum contribution: proofs of Theorems 2.1 and 2.2

For every motif or graphlet, there is a specific range of degrees that gives a major contribution to the number of motifs. We define two optimization problems that identify these ranges of degrees. In Lemmas 3.2 and 3.3 we show that the optimal solutions to these optimization problems have a highly particular structure. We then use these lemmas to prove Theorems 2.1 and 2.2. We first investigate the dependence of the presence of the edges in the erased configuration model.

3.1 The probability of avoiding a subgraph

We relate $L_n = \sum_i D_i$, the total number of half-edges, to its expected value μn by defining the event

$$J_n = \left\{ |L_n - \mu n| \leq n^{1/(\tau-1)} \right\}. \quad (3.1)$$

By [16], $\mathbb{P}(J_n) \rightarrow 1$ as $n \rightarrow \infty$. When we condition on the degree sequence, we will condition on the event J_n , so that we can write $L_n = \mu n(1 + o(1))$. The presence of the edges that form a motif is not independent. Therefore, we use the following lemma which computes the probability of an edge not being present conditional on other edges not being present:

Lemma 3.1. *Fix $m \in \mathbb{N}$ and $\varepsilon > 0$. Let H be a graph with m edges $(u_i, v_i)_{i \in [m]}$ such that $V(H) \subseteq [n]$ and $(u_{m+1}, v_{m+1}) \notin E(H)$. If $D_{u_i}, D_{v_i} \leq n^{1/(\tau-1)}/\varepsilon$ for $i \in [m+1]$, then under the event J_n ,*

$$\mathbb{P}_n(X_{u_{m+1}, v_{m+1}} = 0 \mid E(H) = 0) = O_{\mathbb{P}} \left(e^{-\frac{D_{u_{m+1}} D_{v_{m+1}}}{2L_n}} \right). \quad (3.2)$$

Furthermore, when $D_{u_{m+1}} D_{v_{m+1}} \leq n/\varepsilon$,

$$\mathbb{P}_n(X_{u_{m+1}, v_{m+1}} = 0 \mid E(H) = 0) = e^{-\frac{D_{u_{m+1}} D_{v_{m+1}}}{L_n}} \left(1 + o_{\mathbb{P}} \left(\frac{D_{u_{m+1}} D_{v_{m+1}}}{L_n} n^{-(\tau-2)/(\tau-1)} \right) \right). \quad (3.3)$$

Proof. If H is empty, the claim is proven by [14, Eq (4.9)], which states that for two vertices i and j with $D_i > D_j$,

$$\mathbb{P}_n(X_{i,j} = 0) = e^{-D_i D_j / L_n} + O \left(\frac{D_i^2 D_j}{L_n^2} \right), \quad (3.4)$$

and that by [14, Eq. (4.5)]

$$\mathbb{P}_n(X_{i,j} = 0) = O \left(e^{-D_i D_j / 2L_n} \right). \quad (3.5)$$

Thus we may assume that H is not empty. Order the vertices of H in such a way that u_{m+1}, v_{m+1} are the last vertices in the ordering (if they are present in H at all). We denote the vertices in this order by $w_1, \dots, w_{|V(H)|}$. We now pair the half-edges of the erased configuration model G in this order. Thus, first we pair all half-edges adjacent to w_1 . Since we condition on H not being present, no edge from w_1 is allowed to pair to any of its neighbors in H . After that, we pair all remaining half-edges from w_2 , conditionally on these half-edges not connecting to one of the neighbors of w_2 in H , and so on. We continue until all edges of H have at least one incident vertex that has already been paired. Then, if we pair the rest of the half-edges, we know that H will not be present. Let B denote the number of vertices we have to pair before this happens. Note that in this way, we do not have to pair half-edges adjacent to u_{m+1} or to v_{m+1} (if they are present in H), since they are last in the ordering, and they are not neighbors of each other in H . Thus, all neighbors of u_{m+1} and v_{m+1} in H have already been paired before arriving at u_{m+1} or v_{m+1} . Let $\mathcal{F}_{\leq s} = \sigma((X_{w_i, j})_{i \leq s, j \in [n]})$ be the information about the pairings that have been constructed up to time s .

After B pairings, we denote $\tilde{L}_n = L_n - 2 \sum_{i \in [B]} (D_{w_i} - X_{w_i, w_i})$ and $D_{\tilde{u}_{m+1}} = D_{u_{m+1}} - \sum_{i \in [B]} X_{i, u_{m+1}}$, and we define $D_{\tilde{v}_{m+1}}$ similarly. Note that these quantities are all known in $\mathcal{F}_{\leq B}$. Then, the probability that u_{m+1} does not pair to v_{m+1} is the probability that \tilde{u}_{m+1} does not connect to \tilde{v}_{m+1} in a configuration model with \tilde{L}_n half-edges. Thus,

$$\mathbb{P}_n(X_{u_{m+1}, v_{m+1}} = 0 \mid \mathcal{F}_{\leq B}) = e^{-D_{\tilde{u}_{m+1}} D_{\tilde{v}_{m+1}} / \tilde{L}_n} + O \left(\frac{D_{\tilde{u}_{m+1}}^2 D_{\tilde{v}_{m+1}}}{\tilde{L}_n^2} \right), \quad (3.6)$$

where we have assumed w.l.o.g. that $D_{\tilde{u}_{m+1}} \geq D_{\tilde{v}_{m+1}}$. Since we are under the event J_n from (3.1), $\tilde{L}_n = L_n(1 + o(1))$. Now, we show that $D_{\tilde{u}_{m+1}} = D_{u_{m+1}}(1 + o_{\mathbb{P}}(n^{-(\tau-2)/(\tau-1)}))$. When we pair the half-edges adjacent to w_i , the probability that the j th half-edge pairs to u_{m+1} can be bounded as

$$\mathbb{P}_n(j\text{th half-edge pairs to } u_{m+1}) \leq \frac{D_{u_{m+1}}}{L_n - 2j - 3 - 2 \sum_{s \in [i-1]} D_{w_s}} \leq K \frac{D_{u_{m+1}}}{L_n}, \quad (3.7)$$

for some $K > 0$. We have to pair at most D_{w_i} half-edges, since some of the half-edges incident to w_i may have been used already in previous pairings. Thus, we can stochastically dominate

$X_{w_i, u_{m+1}}$ by Y_i , where $Y_i \sim \text{Bin}(D_{w_i}, KD_{u_{m+1}}/L_n)$. Then,

$$X_{u_{m+1}, w_i} = O_{\mathbb{P}} \left(\frac{D_{w_i} D_{u_{m+1}}}{L_n} \right). \quad (3.8)$$

Then,

$$D_{\bar{u}_{m+1}} = D_{u_{m+1}} \left(1 - \sum_{i \in [B]} O_{\mathbb{P}} \left(\frac{D_{w_i}}{L_n} \right) \right) = D_{u_{m+1}} (1 + o_{\mathbb{P}}(n^{-\frac{\tau-2}{\tau-1}})), \quad (3.9)$$

Similarly, $D_{\bar{v}_{m+1}} = D_{v_{m+1}} (1 + o_{\mathbb{P}}(n^{-(\tau-2)/(\tau-1)}))$. Then, when $D_{u_{m+1}} D_{v_{m+1}} = O(n)$, (3.6) becomes

$$\begin{aligned} \mathbb{P}_n (X_{u_{m+1}, v_{m+1}} = 0 \mid \mathcal{F}_{B+1}) &= e^{-D_{u_{m+1}} D_{v_{m+1}} (1 + o_{\mathbb{P}}(n^{-(\tau-2)/(\tau-1)})) / L_n} + O \left(\frac{D_{u_{m+1}}^2 D_{v_{m+1}}}{L_n^2} \right) \\ &= e^{-D_{u_{m+1}} D_{v_{m+1}} / L_n} \left(1 + o_{\mathbb{P}} \left(\frac{D_{u_{m+1}} D_{v_{m+1}}}{L_n} n^{-(\tau-2)/(\tau-1)} \right) \right). \end{aligned} \quad (3.10)$$

By (3.5)

$$\mathbb{P}_n (X_{u_{m+1}, v_{m+1}} = 0 \mid \mathcal{F}_{B+1}) \leq e^{-D_{\bar{u}_{m+1}} D_{\bar{v}_{m+1}} / 2\bar{L}_n} = O_{\mathbb{P}} \left(e^{-\frac{D_{u_{m+1}} D_{v_{m+1}}}{2L_n}} \right). \quad (3.11)$$

which proves the lemma. \square

3.2 Motifs

Assume that $D_i = \Theta(n^{\alpha_i})$ for some $\alpha_i \in [0, 1/(\tau-1)]$ for all i . Then, when $\alpha_i + \alpha_j < 1$, by (3.4)

$$\mathbb{P}_n (X_{ij} = 1) = \left(1 - e^{-\Theta(n^{\alpha_i + \alpha_j - 1})} \right) (1 + o(1)) = \Theta(n^{\alpha_i + \alpha_j - 1}). \quad (3.12)$$

When $\alpha_i + \alpha_j > 1$, by (3.5)

$$\mathbb{P}_n (X_{ij} = 1) = 1 - O \left(e^{-n^{\alpha_i + \alpha_j - 1}/2} \right). \quad (3.13)$$

For vertices i and j denote $w_{ij} = n^{\alpha_i + \alpha_j - 1 - (\tau-2)/(\tau-1)}$. By Lemma 3.1, for any set of m edges,

$$\begin{aligned} \mathbb{P}_n (X_{u_1, v_1} = \dots = X_{u_m, v_m} = 0) &= \prod_{\alpha_{u_i} + \alpha_{v_i} < 1} (1 + o_{\mathbb{P}}(w_{u_i, v_i})) \Theta \left(1 - n^{\alpha_{u_i} + \alpha_{v_i} - 1} \right) \\ &\times \prod_{\alpha_{u_i} + \alpha_{v_i} < 1} O_{\mathbb{P}} \left(e^{-n^{(\alpha_{v_i} + \alpha_{u_i} - 1)/2}} \right). \end{aligned} \quad (3.14)$$

Let H be a motif on k vertices labeled as $1, \dots, k$ and edges $\{u_1, v_1\}, \dots, \{u_m, v_m\}$. Then, we can write the probability that motif H is present on a specified subset of vertices (i_1, \dots, i_k) as

$$\begin{aligned} \mathbb{P}_n (H_{(i_1, \dots, i_k)} \text{ present}) &= 1 - \sum_{l=1}^m \mathbb{P}_n (X_{i_{u_l}, i_{v_l}} = 0) + \sum_{l \neq j} \mathbb{P}_n (X_{i_{u_l}, i_{v_l}} = X_{i_{u_j}, i_{v_j}} = 0) \dots \\ &- \sum_{l \neq j \neq w} \mathbb{P}_n (X_{i_{u_l}, i_{v_l}} = X_{i_{u_j}, i_{v_j}} = X_{i_{u_w}, i_{v_w}} = 0) + \dots \\ &+ (-1)^m \mathbb{P}_n (X_{i_{u_1}, i_{v_1}} = \dots = X_{i_{u_m}, i_{v_m}} = 0) \\ &= \Theta_{\mathbb{P}} \left(\prod_{(i, j) \in E(H): \alpha_i + \alpha_j < 1} n^{\alpha_i + \alpha_j - 1} \right), \end{aligned} \quad (3.15)$$

where we used that for $\alpha_i + \alpha_j < 1$

$$1 - (1 - n^{\alpha_i + \alpha_j - 1})(1 + o_{\mathbb{P}}(w_{ij})) = \Theta_{\mathbb{P}}(n^{\alpha_i + \alpha_j - 1}), \quad (3.16)$$

and that for $\alpha_i + \alpha_j > 1$

$$1 - O_{\mathbb{P}}\left(e^{-n^{(\alpha_i + \alpha_j - 1)/2}}\right) = 1 + o_{\mathbb{P}}(1). \quad (3.17)$$

The degrees are i.i.d. samples from a power-law distribution. Therefore,

$$\mathbb{P}(D_1 \in [\varepsilon, 1/\varepsilon](\mu n)^{\alpha}) = \int_{\varepsilon(\mu n)^{\alpha}}^{1/\varepsilon(\mu n)^{\alpha}} cx^{-\tau} dx = K(\varepsilon)(\mu n)^{\alpha(1-\tau)} \quad (3.18)$$

for some constant $K(\varepsilon)$ not depending on n . The number of vertices with degrees in $[\varepsilon, 1/\varepsilon](\mu n)^{\alpha}$ is $\text{Binomial}(n, K(\varepsilon)(\mu n)^{\alpha(1-\tau)})$. Therefore, the number of vertices with degrees in $[\varepsilon, 1/\varepsilon](\mu n)^{\alpha}$ is $\Theta(n^{(1-\tau)\alpha+1})$ with high probability. Let $M_n^{(\alpha)}$ be as in (2.5). Then,

$$\# \text{ sets of vertices with degrees in } M_n^{(\alpha)} = \Theta_{\mathbb{P}}\left(n^{k+(1-\tau)\sum_i \alpha_i}\right). \quad (3.19)$$

Combining (3.15) and (3.19) yields that

$$N^{(\text{sub})}(H, M_n^{(\alpha)}(\varepsilon)) = \Theta_{\mathbb{P}}\left(n^{k+(1-\tau)\sum_i \alpha_i} \prod_{(i,j) \in E_H: \alpha_i + \alpha_j < 1} n^{\alpha_i + \alpha_j - 1}\right). \quad (3.20)$$

The maximum contribution is obtained for α_i that maximize

$$\begin{aligned} & \max(1 - \tau) \sum_i \alpha_i + \sum_{(i,j) \in E_H: \alpha_i + \alpha_j < 1} \alpha_i + \alpha_j - 1 \\ & \text{s.t. } \alpha_i \in [0, \frac{1}{\tau-1}] \forall i. \end{aligned} \quad (3.21)$$

The following lemma shows that this optimization problem attains its maximum for highly specific values of α :

Lemma 3.2 (Maximum contribution to motifs). *Let H be a connected graph on k vertices. If the solution to (3.21) is unique, then the optimal solution satisfies $\alpha_i \in \{0, \frac{\tau-2}{\tau-1}, \frac{1}{2}, \frac{1}{\tau-1}\}$ for all i . If it is not unique, then there exist at least 2 optimal solutions with $\alpha_i \in \{0, \frac{\tau-2}{\tau-1}, \frac{1}{2}, \frac{1}{\tau-1}\}$ for all i . In any optimal solution $\alpha_i = 0$ if and only if vertex i has degree one in H .*

Proof. Defining $\beta_i = \alpha_i - \frac{1}{2}$ yields for (3.21)

$$\max \frac{1-\tau}{2}k + (1-\tau) \sum_i \beta_i + \sum_{(i,j) \in E_H: \beta_i + \beta_j < 0} \beta_i + \beta_j, \quad (3.22)$$

over all possible values of $\beta_i \in [-\frac{1}{2}, \frac{3-\tau}{2(\tau-1)}]$. Then, we have to prove that $\beta_i \in \{-\frac{1}{2}, \frac{\tau-3}{2(\tau-1)}, 0, \frac{3-\tau}{2(\tau-1)}\}$ for all i in the optimal solution. Note that (3.22) is a piecewise linear function in β . Therefore, if (3.22) has a unique maximum, it must be attained at the boundary for β_i or at a border of one of the linear sections. Thus, any unique optimal value of β_i satisfies $\beta_i = -\frac{1}{2}$, $\beta_i = \frac{\tau-3}{2(\tau-1)}$ or $\beta_i + \beta_j = 0$ for some j . We ignore the constant factor of $(1-\tau)\frac{k}{2}$ in (3.22), since it does not influence the optimal β values. Rewriting (3.22) without the constant factor yields

$$\max \sum_i \beta_i (1 - \tau + \# \text{ edges to } j \text{ with } \beta_j < -\beta_i). \quad (3.23)$$

The proof of the lemma then consists of three steps.

Step 1. Show that $\beta_i = -\frac{1}{2}$ if and only if vertex i has degree 1 in H in any optimal solution.

Step 2. Show that any unique solution does not have vertices i with $|\beta_i| \in (0, \frac{3-\tau}{2(\tau-1)})$.

Step 3. Show that any optimal solution that is not unique can be transformed into two different optimal solutions with $\beta_i \in \{-\frac{1}{2}, \frac{\tau-3}{2(\tau-1)}, 0, \frac{3-\tau}{2(\tau-1)}\}$ for all i .

Step 1. Let i be a vertex of degree 1 in H , and j be the neighbor of i . Let N_j denote the number of edges in H from j to other vertices v not equal to i with $\beta_v < -\beta_j$. The contribution from vertices i and j to (3.23) is

$$\beta_j(1 - \tau + N_j) + \beta_i(1 - \tau + \mathbb{1}_{\{\beta_i > -\beta_j\}}) + \beta_j \mathbb{1}_{\{\beta_i < -\beta_j\}}. \quad (3.24)$$

For any value of $\beta_j \in [-\frac{1}{2}, \frac{3-\tau}{2(\tau-1)}]$, this contribution is maximized when choosing $\beta_i = -\frac{1}{2}$. Thus, $\beta_i = -\frac{1}{2}$ in the optimal solution if the degree of vertex i is one.

Let i be a vertex with $d_i \geq 2$ in H , and suppose $\beta_i = -\frac{1}{2}$. Because the maximal value of $\beta = \frac{3-\tau}{2(\tau-1)}$, the contribution to (3.23) is

$$-\frac{1}{2}(1 - \tau + d_i) < 0. \quad (3.25)$$

Increasing β_i to $\frac{\tau-3}{2(\tau-1)}$ then gives a higher contribution. Thus, any vertex i with degree at least 2 in H must have $\beta_i = \frac{3-\tau}{2(\tau-1)}$ or $\beta_i + \beta_j = 0$ for some j in an optimal solution. In particular, this means that $\beta_i \geq \frac{\tau-3}{2(\tau-1)}$ when $d_i \geq 2$.

Step 2. Now we show that when the solution to (3.23) is unique, it is never optimal to have $|\beta| \in (0, \frac{3-\tau}{2(\tau-1)})$. Note that vertices with β in this range must be vertices of degree at least 2. Let

$$\tilde{\beta} = \min_{i: |\beta_i| > 0} |\beta_i|, \quad (3.26)$$

and let $\hat{\beta}$ be the second smallest positive $|\beta|$. If $\tilde{\beta} = \frac{3-\tau}{2(\tau-1)}$, then we are finished, so assume that $\tilde{\beta} < \frac{3-\tau}{2(\tau-1)}$. Then, there exist $N_{\tilde{\beta}^-}$ vertices with their β value equal to $-\tilde{\beta}$, and $N_{\tilde{\beta}^+}$ vertices with value $\tilde{\beta}$, where $N_{\tilde{\beta}^+} + N_{\tilde{\beta}^-} \geq 1$. Furthermore, let $E_{\tilde{\beta}^-}$ denote the number of edges from vertices with value $-\tilde{\beta}$ to other vertices j such that $\beta_j < \tilde{\beta}$, and $E_{\tilde{\beta}^+}$ the number of edges from vertices with value $\tilde{\beta}$ to other vertices j such that $\beta_j < -\tilde{\beta}$. Then, the contribution from these vertices to (3.23) is

$$\tilde{\beta} \left((1 - \tau) \left(N_{\tilde{\beta}^+} - N_{\tilde{\beta}^-} \right) + E_{\tilde{\beta}^+} - E_{\tilde{\beta}^-} \right). \quad (3.27)$$

If this contribution is smaller than zero, then we can decrease $\tilde{\beta}$ to 0. This does not change the contribution of the other vertices by the choice of $\tilde{\beta}$, thus this increases the optimal value, which is impossible. On the other hand, when this contribution is larger than zero, we can increase $\tilde{\beta}$ to $\hat{\beta}$. Again, this does not change the other contributions, so this would again increase the optimal value. Thus, this contribution must equal zero. Then, changing $\tilde{\beta}$ to 0 does not change the optimal value, so the solution is not unique. Thus, if the optimal solution is unique, $\tilde{\beta} = \frac{3-\tau}{2(\tau-1)}$. This shows that $\beta_i \in \{\frac{\tau-3}{2(\tau-1)}, 0, \frac{3-\tau}{2(\tau-1)}\}$ for all i .

Step 3. If the solution to (3.23) is not unique, then by the same argument that leads to (3.27), there exist $\hat{\beta}_1, \dots, \hat{\beta}_s > 0$ for some $s \geq 1$ such that

$$\hat{\beta}_j \left((1 - \tau) \left(N_{\hat{\beta}_j^+} - N_{\hat{\beta}_j^-} \right) + E_{\hat{\beta}_j^+} - E_{\hat{\beta}_j^-} \right) = 0 \quad \forall j \in [s]. \quad (3.28)$$

Here we use the same notation as in (3.27). All other values of β must either be $0, \frac{3-\tau}{2(\tau-1)}$ or $\frac{\tau-3}{2(\tau-1)}$ by the argument in Step 3. Thus, setting all $\hat{\beta}_j$ to zero does not change the value of the solution, and setting all $\hat{\beta}_j$ to $\frac{3-\tau}{2(\tau-1)}$ also does not change the value of the solution. Thus, if the solution to (3.23) is not unique, at least 2 solutions exist with $\beta_i \in \{\frac{\tau-3}{2(\tau-1)}, 0, \frac{3-\tau}{2(\tau-1)}\}$ for all i . \square

Proof of Theorem 2.1. Let $\alpha^{(\text{sub})}$ be the unique optimizer of (3.21). Then, by (3.20), for any $\alpha \neq \alpha^{(\text{sub})}$, then

$$\frac{N^{(\text{sub})}(H, M_n^{(\alpha)}(\varepsilon))}{N^{(\text{sub})}(H, M_n^{(\alpha^{(\text{sub})})}(\varepsilon))} = \Theta_{\mathbb{P}}(n^{-\eta}) \quad (3.29)$$

for some $\eta > 0$. By Lemma 3.2, the maximal value of (3.21) is attained by partitioning $V_H \setminus V_1$ into the sets S_1, S_2, S_3 such that vertices in S_1 have $\alpha_i = \frac{\tau-2}{\tau-1}$, vertices in S_2 have $\alpha_i = \frac{1}{\tau-1}$, vertices in S_3 have $\alpha_i = \frac{1}{2}$ and vertices in V_1 have $\alpha_i = 0$. Then, the edges with $\alpha_i + \alpha_j < 1$ are edges inside S_1 , edges between S_1 and S_3 and edges from degree 1 vertices. If we denote the number of edges inside S_1 by E_{S_1} , the number of edges between S_1 and S_3 by E_{S_1, S_3} and the number of edges between V_1 and S_i by $E_{S_i, 1}$, then we can rewrite (3.21) as

$$\begin{aligned} \max_{\mathcal{P}} (1-\tau) \left(\frac{\tau-2}{\tau-1} |S_1| + \frac{1}{\tau-1} |S_2| + \frac{1}{2} |S_3| \right) + \frac{\tau-3}{\tau-1} E_{S_1} + \frac{\tau-3}{2(\tau-1)} E_{S_1, S_3} \\ - \frac{E_{S_1, 1}}{\tau-1} - \frac{\tau-2}{\tau-1} E_{S_2, 1} - \frac{1}{2} E_{S_3, 1}, \end{aligned} \quad (3.30)$$

over all partitions \mathcal{P} of the vertices of H into S_1, S_2, S_3 . Using that $|S_3| = k - |S_1| - |S_2| - k_1$, $E_{S_3, 1} = k_1 - E_{S_1, 1} - E_{S_2, 1}$, where $k_1 = |V_1|$ and extracting a factor $(3-\tau)/2$ results in

$$\max_{\mathcal{P}} \frac{1-\tau}{2} k + \frac{(3-\tau)}{2} \left(|S_1| - |S_2| + \frac{\tau-2}{3-\tau} k_1 - \frac{2E_{S_1} + E_{S_1, S_3}}{\tau-1} - \frac{E_{S_1, 1} - E_{S_2, 1}}{\tau-1} \right), \quad (3.31)$$

Since k and k_1 are fixed and $3-\tau > 0$, we need to maximize

$$B^{(\text{sub})}(H) = \max_{\mathcal{P}} \left[|S_1| - |S_2| - \frac{2E_{S_1} + E_{S_1, S_3} + E_{S_1, 1} - E_{S_2, 1}}{\tau-1} \right]. \quad (3.32)$$

Furthermore, by Lemma 3.2, the optimal value of (3.32) is unique if and only if the solution to (3.21) is unique. Combining this with (3.29) proves the first part of the theorem. By (3.20), the contribution of the maximum is then given by

$$n^{\frac{3-\tau}{2}(k+B^{(\text{sub})}(H)) + \frac{\tau-2}{2}k_1} = n^{\frac{3-\tau}{2}(k_2+B^{(\text{sub})}(H)) + \frac{1}{2}k_1}, \quad (3.33)$$

which proves the second part of the theorem. \square

3.3 Graphlets

For graphlets we can define a similar optimization problem as (3.22). When $\alpha_i + \alpha_j < 1$, (3.4) results in

$$\mathbb{P}_n(X_{ij} = 0) = e^{-\Theta(n^{\alpha_i + \alpha_j - 1})}(1 + o(1)) = 1 + o(1), \quad (3.34)$$

whereas for $\alpha_i + \alpha_j > 1$, (3.13) yields

$$\mathbb{P}_n(X_{ij} = 0) = o(1). \quad (3.35)$$

Similar to (3.15), we can write the probability that H occurs as an induced subgraph on (v_1, \dots, v_k) as

$$\mathbb{P}_n(H \text{ induced on } (v_1, \dots, v_k)) = \Theta_{\mathbb{P}} \left(\prod_{(v_i, v_j) \in E_H: \alpha_i + \alpha_j < 1} n^{\alpha_i + \alpha_j - 1} \prod_{(i, j) \notin E_H: \alpha_i + \alpha_j > 1} o(e^{-n^{\alpha_i + \alpha_j - 1/2}}) \right). \quad (3.36)$$

Thus, the probability that H is an induced subgraph on (v_1, \dots, v_k) is exponentially small in n when two vertices i and j with $\alpha_i + \alpha_j > 1$ are not connected in H . Then the corresponding optimization problem to (3.21) for graphlets becomes

$$\begin{aligned} \max(1-\tau) \sum_i \alpha_i + \sum_{(i, j) \in E_H: \alpha_i + \alpha_j < 1} \alpha_i + \alpha_j - 1, \\ \text{s.t. } \alpha_i + \alpha_j \leq 1 \quad \forall (i, j) \notin E_H. \end{aligned} \quad (3.37)$$

Lemma 3.3 (Maximum contribution to graphlets). *Let H be connected graph on k vertices. If the solution to (3.38) is unique, then the optimal solution satisfies $\alpha_i \in \{0, \frac{\tau-2}{\tau-1}, \frac{1}{2}, \frac{1}{\tau-1}\}$ for all i . If it is not unique, then there exist at least 2 optimal solutions with $\alpha_i \in \{0, \frac{\tau-2}{\tau-1}, \frac{1}{2}, \frac{1}{\tau-1}\}$ for all i . In any optimal solution $\alpha_i = 0$ if and only if vertex i has degree one in H .*

Proof. This proof is highly similar to the proof for motifs. First, we define again $\beta_i = \alpha_i - \frac{1}{2}$, so that (3.37) becomes

$$\begin{aligned} \max \quad & \frac{1-\tau}{2}k + (1-\tau) \sum_i \beta_i + \sum_{(i,j) \in E_H: \beta_i + \beta_j < 0} \beta_i + \beta_j, \\ \text{s.t.} \quad & \beta_i + \beta_j \leq 0 \quad \forall (i,j) \notin E_H. \end{aligned} \quad (3.38)$$

The proof of Step 1 from Lemma 3.2 then also hold for graphlets. Now we prove that if the optimal solution to (3.38) is unique, it satisfies $\beta_i \in \{-\frac{1}{2}, \frac{\tau-3}{2(\tau-1)}, 0, \frac{3-\tau}{2(\tau-1)}\}$ for all i . We take $\tilde{\beta}$ as in (3.26), and $\hat{\beta}$ as the second smallest positive absolute value of β . When $\tilde{\beta} = \frac{3-\tau}{2(\tau-1)}$ we are finished, so we assume that $\tilde{\beta} < \frac{3-\tau}{2(\tau-1)}$. Again, the contribution of the vertices with $|\beta_i| = \tilde{\beta}$ is as in (3.27). By increasing $\tilde{\beta}$ to $\hat{\beta}$ or by decreasing it to zero, the constraints on $\beta_i + \beta_j$ are still satisfied for all (i,j) . Thus, we can use the same argument as in Lemma 3.2 to conclude that $\beta_i \in \{\frac{\tau-3}{2(\tau-1)}, 0, \frac{3-\tau}{2(\tau-1)}\}$ for all i with $d_i \geq 2$. A similar argument as in Step 3 of Lemma 3.2 then shows that if the solution to (3.38) is not unique, it can be transformed into two optimal solutions that satisfy $\beta_i \in \{-\frac{1}{2}, \frac{\tau-3}{2(\tau-1)}, 0, \frac{3-\tau}{2(\tau-1)}\}$ for all i with degree at least 2. \square

Proof of Theorem 2.2. This proof is highly similar to the proof of Theorem 2.1. \square

4 Proof of Theorems 2.3 and 2.4

We prove Theorems 2.3 and 2.4 using the following lemmas. We define

$$W_n^k(\varepsilon) = \{(u_1, \dots, u_k) : D_{u_i} \in [\varepsilon, 1/\varepsilon] \sqrt{\mu n} \quad \forall i \in [k]\}. \quad (4.1)$$

Then, we denote the number of motifs or graphlets H with all degrees in $W_n^k(\varepsilon)$ by $N^{(\text{sub})}(H, W_n^k(\varepsilon))$ and $N^{(\text{ind})}(H, W_n^k(\varepsilon))$ respectively.

Lemma 4.1 (Convergence of major contribution to motifs). *Let H be a connected graph on $k > 2$ vertices such that (2.2) is uniquely optimized at 0. Then,*

(i) *The number of motifs with vertices in $W_n^k(\varepsilon)$ satisfies*

$$\frac{N^{(\text{sub})}(H, W_n^k(\varepsilon))}{n^{\frac{k}{2}(3-\tau)}} \xrightarrow{\mathbb{P}} c^k \mu^{-\frac{k}{2}(\tau-1)} \int_{\varepsilon}^{1/\varepsilon} \dots \int_{\varepsilon}^{1/\varepsilon} (x_1 \dots x_t)^{-\tau} \prod_{(i,j) \in E_H} (1 - e^{-x_i x_j}) dx_1 \dots dx_k. \quad (4.2)$$

(ii) *Furthermore,*

$$A^{(\text{sub})}(H) = \int_0^{\infty} \dots \int_0^{\infty} (x_1 \dots x_t)^{-\tau} \prod_{(i,j) \in E_H} (1 - e^{-x_i x_j}) dx_1 \dots dx_k < \infty. \quad (4.3)$$

Lemma 4.2 (Convergence of major contribution to graphlets). *Let H be a connected graph on $k > 2$ vertices such that (2.3) is uniquely optimized at 0. Then,*

(i) *The number of graphlets with vertices in $W_n^k(\varepsilon)$ satisfies*

$$\frac{N^{(\text{ind})}(H, W_n^k(\varepsilon))}{n^{\frac{k}{2}(3-\tau)}} \xrightarrow{\mathbb{P}} c^k \mu^{-\frac{k}{2}(\tau-1)} \int_{\varepsilon}^{1/\varepsilon} \dots \int_{\varepsilon}^{1/\varepsilon} (x_1 \dots x_t)^{-\tau} \prod_{(i,j) \in E_H} (1 - e^{-x_i x_j}) \prod_{(i,j) \notin E_H} e^{-x_i x_j} dx_1 \dots dx_k. \quad (4.4)$$

(ii) Furthermore,

$$A^{(\text{ind})}(H) = \int_0^\infty \cdots \int_0^\infty (x_1 \cdots x_t)^{-\tau} \prod_{(i,j) \in E_H} (1 - e^{-x_i x_j}) \prod_{(i,j) \notin E_H} e^{-x_i x_j} dx_1 \cdots dx_k < \infty. \quad (4.5)$$

The proof of these lemmas can be found in Section 5. The proof of the first parts of Lemmas 4.1 and 4.2 can be found in Section 5. In fact, when we adjust the normalization, the first part of Lemmas 4.1 and 4.2 holds for all motifs of graphlets that have a unique optimizer for (2.2) or (2.3) respectively. For ease of notation we only prove the lemmas for unique optimizers of (2.2) and (2.3) at 0. The second part of the lemmas is proven in Section 6. We now prove Theorems 2.3 and 2.4 using these lemmas.

Proof of Theorem 2.1. We first study the expected number of motifs with vertices outside $W_n^k(\varepsilon)$. First, we investigate the expected number of motifs where vertex 1 has degree smaller than $\varepsilon\sqrt{\mu n}$. Because $\mathbb{P}_n(X_{ij} = 1) \leq \min(D_i D_j / L_n, 1)$, this contribution can be bounded as

$$\begin{aligned} \mathbb{E}[N(H) \mid D_1 < \varepsilon\sqrt{\mu n}] &\leq n^k \int_1^{\varepsilon\sqrt{\mu n}} \int_1^\infty \cdots \int_1^\infty (x_1 \cdots x_k)^{-\tau} \prod_{(i,j) \in E_H} \min\left(\frac{x_i x_j}{\mu n}, 1\right) dx_1 \cdots dx_k \\ &= n^k (\mu n)^{\frac{k}{2}(1-\tau)} \int_0^\varepsilon \int_0^\infty \cdots \int_0^\infty (t_1 \cdots t_k)^{-\tau} \prod_{(i,j) \in E_H} \min(t_i t_j, 1) dt_1 \cdots dt_k \\ &\leq K^{|E_H|} n^{\frac{k}{2}(3-\tau)} \mu^{\frac{k}{2}(1-\tau)} \int_0^\varepsilon \int_0^\infty \cdots \int_0^\infty (t_1 \cdots t_k)^{-\tau} \prod_{(i,j) \in E_H} (1 - e^{-t_i t_j}) dt_1 \cdots dt_k \\ &= O\left(n^{\frac{k}{2}(3-\tau)}\right) h_1(\varepsilon), \end{aligned} \quad (4.6)$$

where we used that $\min(1, x) \leq K(1 - e^{-x})$ for some $K > 0$, and $h_1(\varepsilon)$ is a function of ε . By Lemma 4.1(ii), $h(\varepsilon) \rightarrow 0$ as ε tends to zero. We can bound the situation where one of the other vertices has degree smaller than $\varepsilon\sqrt{n}$, or where one of the vertices has degree larger than \sqrt{n}/ε similarly. This results in

$$\mathbb{E}[N(H, \bar{W}_n^k(\varepsilon))] = O\left(n^{\frac{k}{2}(3-\tau)}\right) h(\varepsilon), \quad (4.7)$$

for some function $h(\varepsilon)$ not depending on n such that $h(\varepsilon) \rightarrow 0$ when $\varepsilon \rightarrow 0$. Then, by the Markov inequality,

$$N(H, \bar{W}_n^k(\varepsilon)) = h(\varepsilon) O_p\left(n^{\frac{k}{2}(3-\tau)}\right). \quad (4.8)$$

Combining this with Lemma 4.1(i) gives

$$\frac{N(H)}{n^{\frac{k}{2}(3-\tau)}} \xrightarrow{\mathbb{P}} c^k \mu^{-\frac{k}{2}(\tau-1)} \int_\varepsilon^{1/\varepsilon} \cdots \int_\varepsilon^{1/\varepsilon} (x_1, \cdots, x_k)^{-\tau} \prod_{(u,v) \in E_H} (1 - e^{-x_u x_v}) dx_1 \cdots dx_k + O_p(h(\varepsilon)). \quad (4.9)$$

Then letting $\varepsilon \rightarrow 0$ proves the theorem. \square

Proof of Theorem 2.2. This proof is similar to the proof of Theorem 2.1, and uses Lemma 4.2. \square

5 Proof of Lemmas 4.1(i) and 4.2(i)

We now prove Lemma 4.1. We first condition on the degree sequence, and compute the expected value and the variance of the number of motifs conditioned on the degrees in Lemmas 5.1 and 5.3, and for graphlets in Lemmas 5.2 and 5.4. Then we take the i.i.d. degrees into account in Lemma 5.5 for motifs and 5.6. Together, these lemmas prove Lemma 4.1.

5.1 Conditional expectation

In this section, we study the expectation of the number of motifs, conditioned on the degrees. Let H be a motif on k vertices, labeled as $1, \dots, k$, and m edges. We denote the edges as $e_1 = \{u_1, v_1\}, \dots, e_m = \{u_m, v_m\}$.

Lemma 5.1 (Conditional expectation of motifs). *Let H be a motif such that (2.2) has a unique maximum, attained at 0. Then, under the event J_n as defined in (3.1)*

$$\mathbb{E}_n [N^{(\text{sub})}(H, W_n^k(\varepsilon))] = \sum_{(i_1, \dots, i_k) \in W_n^k(\varepsilon)} \prod_{(j, k) \in E_H} (1 - e^{-D_{i_j} D_{i_k} / L_n}) (1 + o_{\mathbb{P}}(1)). \quad (5.1)$$

Proof. We can use (3.15) to show that the expected value satisfies

$$\begin{aligned} \mathbb{E}_n [N^{(\text{sub})}(H, W_n^k(\varepsilon))] &= \sum_{(i_1, \dots, i_k) \in W_n^k(\varepsilon)} \mathbb{P}_n (H_{(i_1, \dots, i_k)} \text{ present}) \\ &= (1 + o_{\mathbb{P}}(1)) \sum_{(i_1, \dots, i_k) \in W_n^k(\varepsilon)} \prod_{l=1}^m \left(1 - \mathbb{P}_n (X_{i_{u_l}, i_{v_l}} = 0)\right). \end{aligned} \quad (5.2)$$

Because $D_i D_j = O(n)$ and $L_n = \mu n(1 + o(1))$ under J_n , by (3.4)

$$\mathbb{P}_n (X_{ij} = 1) = 1 - e^{-D_i D_j / L_n} + O\left(\frac{D_i^2 D_j}{L_n^2}\right) = (1 + o(1)) \left(1 - e^{-D_i D_j / L_n}\right). \quad (5.3)$$

This results in

$$\mathbb{E}_n [N^{(\text{sub})}(H, W_n^k(\varepsilon))] = (1 + o_{\mathbb{P}}(1)) \sum_{(i_1, \dots, i_k) \in W_n^k(\varepsilon)} \prod_{(j, k) \in E_H} (1 - e^{-D_{i_j} D_{i_k} / L_n}). \quad (5.4)$$

□

Now we prove a similar lemma for graphlets.

Lemma 5.2 (Conditional expectation of graphlets). *Let H be a graphlet such that (2.3) has a unique maximum attained at 0. Then, under the event J_n as defined in (3.1),*

$$\mathbb{E}_n [N^{(\text{ind})}(H, W_n^k(\varepsilon))] = \sum_{(i_1, \dots, i_k) \in W_n^k(\varepsilon)} \prod_{(j, k) \in E_H} (1 - e^{-D_{i_j} D_{i_k} / L_n}) \prod_{(s, t) \notin E_H} e^{-D_{i_s} D_{i_t} / L_n} (1 + o(1)). \quad (5.5)$$

Proof. This proof follows similar lines as the proof of Lemma 5.1. For $(i, j) \notin E_H$, we now also use (see (5.3))

$$\mathbb{P}_n (X_{ij} = 0) = e^{-D_i D_j / L_n} + O\left(\frac{D_i^2 D_j}{L_n^2}\right) = e^{-D_i D_j / L_n} (1 + o(1)), \quad (5.6)$$

since $D_i D_j = O(n)$ when $(i, j) \notin E_H$ by (2.3) and $L_n = \Theta(n)$ under the event J_n . □

5.2 Conditional variance

In this section, we still condition on the degrees. The following lemma shows that the variance of the number of motifs is small compared to the expected value computed in the previous section:

Lemma 5.3 (Conditional variance for motifs). *Let H be a motif such that (2.2) has a unique maximum attained at 0. Then, under the event J_n as defined in (3.1)*

$$\frac{\text{Var}_n (N^{(\text{sub})}(H, W_n^k(\varepsilon)))}{\mathbb{E}_n [N^{(\text{sub})}(H, W_n^k(\varepsilon))]^2} \xrightarrow{\mathbb{P}} 0. \quad (5.7)$$

Proof. By Theorem 2.1,

$$\mathbb{E}_n [N^{(\text{sub})}(H, W_n^k(\varepsilon))]^2 = \Theta(n^{(3-\tau)k}), \quad (5.8)$$

where the last step follows from the fact that for \sqrt{n} motifs, $B^{(\text{sub})}(H) = 0$ and $k_1 = 0$. Thus, we need to prove that the variance is small compared to $n^{(3-\tau)k}$. Denote $\mathbf{i} = (i_1, \dots, i_k)$ and $\mathbf{j} = (j_1, \dots, j_k)$. We write the variance as

$$\text{Var}_n (N(H, W_n^k(\varepsilon))) = \sum_{\mathbf{i} \in W_n^k(\varepsilon)} \sum_{\mathbf{j} \in W_n^k(\varepsilon)} \mathbb{P}_n (H_{\mathbf{i}}, H_{\mathbf{j}} \text{ present}) - \mathbb{P}_n (H_{\mathbf{i}} \text{ present}) \mathbb{P}_n (H_{\mathbf{j}} \text{ present}). \quad (5.9)$$

This splits into various cases, depending on the overlap of \mathbf{i} and \mathbf{j} . When \mathbf{i} and \mathbf{j} do not overlap, by (3.15)

$$\begin{aligned} & \sum_{\mathbf{i} \in W_n^k(\varepsilon)} \sum_{\mathbf{j} \in W_n^k(\varepsilon)} \mathbb{P}_n (H_{\mathbf{i}}, H_{\mathbf{j}} \text{ present}) - \mathbb{P}_n (H_{\mathbf{i}} \text{ present}) \mathbb{P}_n (H_{\mathbf{j}} \text{ present}) \\ &= \sum_{\mathbf{i} \in W_n^k(\varepsilon)} \sum_{\mathbf{j} \in W_n^k(\varepsilon)} (1 + o_{\mathbb{P}}(1)) \prod_{l=1}^m (1 - \mathbb{P}_n (X_{i_{u_l}, i_{v_l}} = 0)) \prod_{l=1}^m (1 - \mathbb{P}_n (X_{j_{u_l}, j_{v_l}} = 0)) \\ & \quad - (1 + o_{\mathbb{P}}(1)) \prod_{l=1}^m (1 - \mathbb{P}_n (X_{i_{u_l}, i_{v_l}} = 0)) \prod_{l=1}^m (1 - \mathbb{P}_n (X_{j_{u_l}, j_{v_l}} = 0)) \\ &= \mathbb{E}_n [N^{(\text{sub})}(H, W_n^k(\varepsilon))]^2 o_{\mathbb{P}}(1). \end{aligned} \quad (5.10)$$

The other contributions are when \mathbf{i} and \mathbf{j} overlap. In this situation, we use the bound $\mathbb{P}_n (X_{ij} = 1) \leq 1$. When \mathbf{i} and \mathbf{j} overlap on $s \geq 1$ vertices, we bound the contribution to (5.9) as

$$\sum_{\mathbf{i}, \mathbf{j} \in W_n^k(\varepsilon): |\{\mathbf{i}, \mathbf{j}\}| = 2k-s} \mathbb{P}_n (H_{\mathbf{i}}, H_{\mathbf{j}} \text{ present}) \leq |\{i : D_i \in \sqrt{\mu n}[\varepsilon, 1/\varepsilon]\}|^{2k-s} = O_{\mathbb{P}} \left(n^{\frac{(3-\tau)(2k-s)}{2}} \right), \quad (5.11)$$

which is smaller than $n^{(3-\tau)k}$, as required. \square

A similar lemma holds for graphlets:

Lemma 5.4 (Conditional variance for graphlets). *Let H be a graphlet such that (2.3) has a unique maximum, attained at 0. Then, under the event J_n as defined in (3.1)*

$$\frac{\text{Var}_n (N^{(\text{ind})}(H, W_n^k(\varepsilon)))}{\mathbb{E}_n [N^{(\text{ind})}(H, W_n^k(\varepsilon))]^2} \xrightarrow{\mathbb{P}} 0. \quad (5.12)$$

Proof. This proof is highly similar to the proof of Lemma 5.3. By Theorem 2.2 we have to prove that the variance is small compared to $n^{k(3-\tau)}$, as in the previous theorem. Therefore, the bound (5.11) is also sufficient for graphlets. We only need to derive a straightforward generalization of (5.10), where we also include the probability that edges that are not in H should not be present in the induced subgraph. \square

5.3 Convergence of conditional expectation

We now consider the convergence of the expectation of the number of subgraphs conditioned on the degrees.

Lemma 5.5 (Convergence of conditional expectation of \sqrt{n} motifs). *Let H be a motif such that (2.2) has a unique maximizer, and the maximum is attained at 0. Then,*

$$\frac{\mathbb{E}_n [N^{(\text{sub})}(H, W_n^k(\varepsilon))]}{n^{\frac{k}{2}(3-\tau)}} \xrightarrow{\mathbb{P}} c^k \mu^{-\frac{k}{2}(\tau-1)} \int_{\varepsilon}^{1/\varepsilon} \dots \int_{\varepsilon}^{1/\varepsilon} (x_1 \dots x_k)^{-\tau} \prod_{(u,v) \in E_H} (1 - e^{-x_u x_v}) dx_1 \dots dx_k. \quad (5.13)$$

Proof. Let $|E_H| = m$ and denote the edges of H by $(u_1, v_1), \dots, (u_m, v_m)$. Define

$$g(t_1, \dots, t_k) := \prod_{(u,v) \in E_H} (1 - e^{-t_u t_v}). \quad (5.14)$$

Taylor expanding $1 - e^{-xy}$ on $[\varepsilon, 1/\varepsilon]$ yields

$$1 - e^{-xy} = \sum_{i=1}^s \frac{(xy)^i}{i!} (-1)^i + O\left(\frac{\varepsilon^{-s}}{(s+1)!}\right). \quad (5.15)$$

Since g is a bounded function on $F = [\varepsilon, 1/\varepsilon]^m$, for any $\eta > 0$, we can find s_1, \dots, s_m such that

$$\begin{aligned} g(t_1, \dots, t_k) &= \sum_{i_1=1}^{s_1} \dots \sum_{i_m=1}^{s_m} \left((-1)^{i_1} \frac{t_{u_1}^{i_1} t_{v_1}^{i_1}}{i_1!} \dots (-1)^{i_m} \frac{t_{u_m}^{i_m} t_{v_m}^{i_m}}{i_m!} \right) + O(\eta) \\ &= \sum_{i_1=1}^{s_1} \dots \sum_{i_m=1}^{s_m} \left(\frac{(-1)^{i_1 + \dots + i_m}}{i_1! \dots i_m!} t_1^{\gamma_1} t_2^{\gamma_2} \dots t_k^{\gamma_k} \right) + O(\eta), \end{aligned} \quad (5.16)$$

where

$$\gamma_j := \gamma_j(i_1, \dots, i_k) = \sum_l i_l \mathbb{1}_{\{u_l=j \text{ or } v_l=j\}}. \quad (5.17)$$

Let $M^{(n)}$ denote the random measure

$$M^{(n)}([a, b]) = \frac{(\mu n)^{\frac{1}{2}(\tau-1)}}{n} \sum_{i=1}^n \mathbb{1}_{\{D_i \in \sqrt{\mu n}[a, b]\}}. \quad (5.18)$$

Because the number of vertices with degrees in a certain interval $[a, b]$ is binomially distributed,

$$\begin{aligned} M^{(n)}([a, b]) &= (\mu n)^{\frac{1}{2}(\tau-1)} \mathbb{P}_n \left(\mathcal{D}_u \in (\mu n)^{\frac{1}{2}} [a, b] \right) \xrightarrow{\mathbb{P}} (\mu n)^{\frac{1}{2}(\tau-1)} \int_{a\sqrt{\mu n}}^{b\sqrt{\mu n}} cx^{-\tau} dx \\ &= c \int_a^b x^{-\tau} dx := \lambda([a, b]). \end{aligned} \quad (5.19)$$

Let $N^{(n)}$ denote the product measure $M^{(n)} \times M^{(n)} \times \dots \times M^{(n)}$ (k times). Then, choosing $\eta = \varepsilon^{k+1}$ in (5.16) together with Lemma 5.1 yields

$$\begin{aligned} \frac{\mathbb{E}_n [N^{(\text{sub})}(H, W_n^k(\varepsilon))]}{n^{\frac{k}{2}(3-\tau)} \mu^{\frac{k}{2}(1-\tau)}} &= \int_F g(t_1, \dots, t_k) dN^{(n)}(t_1, \dots, t_k) \\ &= \int_F \sum_{i_1=1}^{s_1} \dots \sum_{i_m=1}^{s_m} \left(\frac{(-1)^{i_1 + \dots + i_m}}{i_1! \dots i_m!} t_1^{\gamma_1} t_2^{\gamma_2} \dots t_k^{\gamma_k} \right) + O(\varepsilon^{k+1}) dN^{(n)}(t_1, \dots, t_k) \\ &= \sum_{i_1=1}^{s_1} \dots \sum_{i_m=1}^{s_m} \frac{(-1)^{i_1 + \dots + i_m}}{i_1! \dots i_m!} \int_{\varepsilon}^{1/\varepsilon} t_1^{\gamma_1} dM^{(n)}(t_1) \dots \int_{\varepsilon}^{1/\varepsilon} t_k^{\gamma_k} dM^{(n)}(t_k) + O(\varepsilon). \end{aligned} \quad (5.20)$$

As in [27, Eq. (55)] for any γ

$$\int_{\varepsilon}^{1/\varepsilon} x^{\gamma} dM^{(n)}(x) \xrightarrow{\mathbb{P}} \int_{\varepsilon}^{1/\varepsilon} x^{\gamma} d\lambda(x). \quad (5.21)$$

Combining this with (5.20) results in

$$\begin{aligned}
\frac{\mathbb{E}_n [N^{(\text{sub})}(H, W_n^k(\varepsilon))]}{f^{(\text{sub})}(n, H)\mu^{\frac{k}{2}(1-\tau)}} &\xrightarrow{\mathbb{P}} \sum_{i_1=1}^{s_1} \cdots \sum_{i_m=1}^{s_m} \frac{(-1)^{i_1+\cdots+i_m}}{i_1! \cdots i_m!} \int_{\varepsilon}^{1/\varepsilon} t_1^{\alpha_1} d\lambda(t_1) \cdots \int_{\varepsilon}^{1/\varepsilon} t_k^{\alpha_k} d\lambda(t_k) + O(\varepsilon) \\
&= \int_F \sum_{i_1=1}^{s_1} \cdots \sum_{i_m=1}^{s_m} \frac{(-1)^{i_1+\cdots+i_m}}{i_1! \cdots i_m!} t_1^{\alpha_1} \cdots t_k^{\alpha_k} d\lambda(t_1) \cdots d\lambda(t_k) + O(\varepsilon) \\
&= \int_F g(t_1, \dots, t_k) d\lambda(t_1) \cdots d\lambda(t_k) + O(\varepsilon).
\end{aligned} \tag{5.22}$$

Then, by (5.19)

$$\frac{\mathbb{E}_n [N^{(\text{sub})}(H, W_n^k(\varepsilon))]}{f^{(\text{sub})}(n, H)} \xrightarrow{\mathbb{P}} c^k \mu^{-\frac{k}{2}(\tau-1)} \int_{\varepsilon}^{1/\varepsilon} \cdots \int_{\varepsilon}^{1/\varepsilon} (t_1 \cdots t_k)^{-\tau} g(t_1, \dots, t_k) dt_1 \cdots dt_k, \tag{5.23}$$

which proves the claim. \square

Lemma 5.6 (Convergence of conditional expectation of graphlets). *Let H be a motif such that (2.3) has a unique maximizer, and the maximum is attained at 0. Then,*

$$\begin{aligned}
\frac{\mathbb{E}_n [N^{(\text{ind})}(H, W_n^k(\varepsilon))]}{n^{\frac{k}{2}(3-\tau)}} &\xrightarrow{\mathbb{P}} c^k \mu^{-\frac{k}{2}(\tau-1)} \int_{\varepsilon}^{1/\varepsilon} \cdots \int_{\varepsilon}^{1/\varepsilon} (x_1, \dots, x_k)^{-\tau} \prod_{(u,v) \in E_H} (1 - e^{-x_u x_v}) \\
&\quad \times \prod_{(s,t) \notin E_H} e^{-x_s x_t} dx_1 \cdots dx_k.
\end{aligned} \tag{5.24}$$

Proof. This proof is highly similar to the proof of Lemma 5.5. The only difference is that we include the exponentials over the edges that are not present in H in the function g in (5.14). \square

6 Proof of Lemmas 4.1(ii) and 4.2(ii)

In this section we first show that if (2.2) uniquely attains its maximum for $S_3 = V_H$, then $A^{(\text{sub})}(H)$ as defined in (2.12) is finite. Similarly, if (2.3) uniquely attains its maximum for $S_3 = V_H$, then $A^{(\text{ind})}(H)$ as defined in (2.14) is finite.

Lemma 6.1. *If H is a connected graph on k vertices with minimum degree 2 such that the maximum in (2.2) is uniquely attained at zero, then $A^{(\text{sub})}(H) < \infty$.*

Proof. Since $1 - e^{-x} \leq \min(x, 1)$ for all $x \in [0, \infty)$, it is sufficient to show that

$$\int_0^{\infty} \cdots \int_0^{\infty} (x_1 \cdots x_k)^{-\tau} \prod_{(u,v) \in E_H} \min(1, x_u x_v) dx_1 \cdots dx_k < \infty. \tag{6.1}$$

This integral consists of multiple regions. The first region is where $x_1, \dots, x_k \geq 1$. Then, since $\tau \in (2, 3)$, this integral results in

$$\int_1^{\infty} \cdots \int_1^{\infty} (x_1 \cdots x_k)^{-\tau} dx_1 \cdots dx_k < \infty. \tag{6.2}$$

The second region is where $x_1, \dots, x_k \in [0, 1]$. Since the minimal degree of H is 2, this integral can be bounded as

$$\int_0^1 \cdots \int_0^1 (x_1 \cdots x_k)^{-\tau} \prod_{(u,v) \in E_H} x_u x_v dx_1 \cdots dx_k \leq \int_0^1 \cdots \int_0^1 (x_1 \cdots x_k)^{2-\tau} dx_1 \cdots dx_k < \infty. \tag{6.3}$$

The other regions can be described by a set S such that the integral runs from 1 to ∞ for all $i \in S$, and from 0 to 1 for all $i \in \bar{S} = V_H \setminus S$. Then, $\min(x_i x_j, 1) = x_i x_j$ when $i, j \notin S$, and $\min(x_i, x_j) = 1$ when $i, j \in S$. W.l.o.g. assume $S = \{1, \dots, t\}$ for some $t \geq 1$. For any $W \subseteq V_H$, we denote by $d_{i(W)}$ the number of edges from of vertex i to vertices in W . Then, the contribution to (6.1) can be written as

$$\int_1^\infty \cdots \int_1^\infty (x_1 \cdots x_t)^{-\tau} \int_0^1 \cdots \int_0^1 \prod_{i=t+1}^k \left(x_i^{-\tau+d_{i(\bar{S})}} \prod_{j \in S: (i,j) \in E_H} \min(x_i x_j, 1) \right) dx_k \cdots dx_1. \quad (6.4)$$

Again, this integral consists of multiple regions, depending on whether $x_u x_v < 1$. Suppose a vertex $u \in \bar{S}$ is connected to vertices $1, 2, \dots, l \in S$, such that $x_1 < x_2 < \cdots < x_l$. Then, the integral over x_u in (6.4) equals

$$\begin{aligned} & \int_0^1 x_u^{-\tau+d_{u(\bar{S})}} \min(x_u x_1, 1) \min(x_u x_2, 1) \cdots \min(x_u x_l, 1) dx_u \\ &= \int_0^{1/x_l} x_u^{-\tau+l+d_{u(\bar{S})}} x_1 \cdots x_l dx_u + \int_{1/x_l}^{1/x_{l-1}} x_u^{-\tau+l-1+d_{u(\bar{S})}} x_1 \cdots x_{l-1} dx_u + \cdots + \int_{1/x_1}^1 x_u^{-\tau+d_{u(\bar{S})}} dx_u \\ &= C_1 x_1 x_2 \cdots x_l^{\tau-l-d_{u(\bar{S})}} + C_2 (x_1 \cdots x_{l-1}^{\tau-l-d_{u(\bar{S})}+1} - x_1 x_2 \cdots x_l^{\tau-l-d_{u(\bar{S})}}) + \cdots + C_l (1 - x_1^{\tau-d_{u(\bar{S})}+1}), \end{aligned} \quad (6.5)$$

for some constants C_1, \dots, C_l . Since $d_{u(\bar{S})} + l - \tau = d_u - \tau > -1$, we know that $C_1 > 0$. Since $1 < x_1 < x_2 < \cdots < x_l$, the largest contribution to (6.5) is the integral from 0 to $1/x_l$. Thus, the largest contribution to (6.4) is when all minima are attained at $x_i x_j$. We now compute this largest contribution when $x_1 < x_2 < \cdots < x_t$. We let

$$u_j = \max\{i \mid i \in S, (i, j) \in E_H\} \quad (6.6)$$

for all $j \in \bar{S}$ such that $d_{j_S} \geq 1$. Furthermore, let

$$f(j) = \begin{cases} 1/x_{u_j} & \text{if } d_{j_S} \geq 1, \\ 1 & \text{else,} \end{cases} \quad (6.7)$$

for all $j \in \bar{S}$. Thus, the ranges of the integrals in (6.4) such that all minima are attained by $x_i x_j$, are from 0 to $f(j)$ for all $j \in \bar{S}$. Then, the contribution to (6.4) where all minima are attained by $x_u x_v$ equals

$$\begin{aligned} & \int_1^\infty \int_{x_1}^\infty \cdots \int_{x_{t-1}}^\infty \prod_{i=1}^t x_i^{-\tau+d_{i(\bar{S})}} \prod_{j=t+1}^k \left(\int_0^{f(j)} x_j^{-\tau+d_j} dx_j \right) dx_t \cdots dx_1 \\ &= C \int_1^\infty \int_{x_1}^\infty \cdots \int_{x_{t-1}}^\infty \prod_{i=1}^t x_i^{-\tau+d_{i(\bar{S})}} \prod_{j=t+1}^k f(j)^{-\tau+1+d_j} dx_t \cdots dx_1, \end{aligned} \quad (6.8)$$

for some constant $C > 0$. Let $W_i = \{i \in \bar{S} : u_j = i\}$ for $i \in [t]$ (for an illustration, see Figure 5). Then (6.8) results in

$$C \int_1^\infty \int_{x_1}^\infty \cdots \int_{x_{t-1}}^\infty \prod_{i=1}^t x_i^{-\tau+d_{i(\bar{S})}+(\tau-1)|W_i|-2E_{W_i}-E_{W_i, \bar{W}_i}} dx_t \cdots dx_1, \quad (6.9)$$

where E_{W_i} denotes the number of edges inside W_i and E_{W_i, \bar{W}_i} denotes the number of edges between W_i and \bar{W}_i . We now want to show that

$$-\tau + d_{t_{\bar{S}}} + (\tau - 1) |W_t| - 2E_{W_t} - E_{W_t, \bar{W}_t} < -1, \quad (6.10)$$

so that the integral in (6.9) over x_t is finite. Note that by definition of (6.6) and $W_t, d_{t_{\bar{S}}} = d_{t_{W_t}}$ (see also Figure 5). By setting $S_2 = \{t\}$ in (2.2), $S_1 = W_t$ and $S_3 = V_H \setminus (S_1 \cup S_2)$, we have

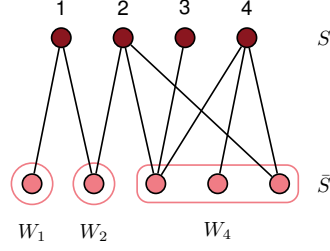


Figure 5: Illustration of the sets W_i . For clarity of the picture, edges in S and in \bar{S} are not displayed.

$E_{S_1} = E_{W_t}$ and $E_{S_1, S_3} = E_{W_t, \bar{W}_t} - d_{t_{W_t}} = E_{W_t, \bar{W}_t} - d_{t_{\bar{S}}}$. Because (2.2) is uniquely optimized at 0 for $S_3 = V_H$,

$$|W_t| - 1 - \frac{2E_{W_t} - E_{W_t, \bar{W}_t} + d_{t_{\bar{S}}}}{\tau - 1} < 0, \quad (6.11)$$

or

$$-\tau + (\tau - 1)|W_t| - 2E_{W_t} - E_{W_t, \bar{W}_t} + d_{t_{\bar{S}}} < -1. \quad (6.12)$$

Thus, the inner integral is finite, and evaluates to

$$x_{t-1}^{1-2\tau+d_{t-1, \bar{S}}+(\tau-1)|W_t \cup W_{t-1}|-2E_{W_t \cup W_{t-1}}-E_{W_t \cup W_{t-1}, W_t \cup \bar{W}_{t-1}}}. \quad (6.13)$$

Then, choosing $S_2 = \{t, t-1\}$, $S_1 = W_t \cup W_{t-1}$ and $S_3 = V_H \setminus (S_1 \cup S_2)$, we can again prove using (2.2) that this integral is finite. We can continue this process until we arrive at the integral over x_1 and show that this final integral is finite. Since the ordering $x_1 < x_2 < \dots < x_t$ was arbitrary, the integral over any ordering of x_1, \dots, x_t is finite. \square

Lemma 6.2. *If H is a connected graph on k vertices with minimum degree 2 such that (2.3) has a unique maximizer attained at 0, then $A^{(\text{ind})}(H) < \infty$.*

Proof. Because $1 - e^{-x} \leq \min(1, x)$, it suffices to show that

$$\int_0^\infty \dots \int_0^\infty (x_1 \dots x_k)^{-\tau} \prod_{(u,v) \in E_H} \min(1, x_u x_v) \prod_{(i,j) \notin E_H} e^{-x_i x_j} dx_1 \dots dx_k < \infty. \quad (6.14)$$

We can show similarly to (6.2) and (6.3) that this integral is finite when all x values are larger than one, or when all are smaller than one. We now show that the contribution where the integral runs from 1 to ∞ for vertices in some nonempty set S , and from 0 to 1 for vertices in $\bar{S} = V \setminus S$. W.l.o.g., assume $S = \{1, \dots, t\}$ for some $t \geq 1$. We use the bound $e^{-x_u x_v} \leq 1$ when u and v are not both in S . This bounds (6.14) as

$$\int_1^\infty \dots \int_1^\infty (x_1 \dots x_t)^{-\tau} \prod_{u,v \in S: (u,v) \notin E_H} e^{-x_u x_v} \prod_{i=t+1}^k \left(\int_0^1 x_i^{-\tau+d_{i(S)}} \prod_{j \in S: (i,j) \in E_H} \min(x_i x_j, 1) dx_i \right) dx_t \dots dx_1. \quad (6.15)$$

First, we consider the case where the vertices of S induce a complete graph in H . Then (6.15) results in

$$\int_1^\infty \dots \int_1^\infty (x_1 \dots x_t)^{-\tau} \prod_{i=t+1}^k \left(\int_0^1 x_i^{-\tau+d_{i(S)}} \prod_{j \in S: (i,j) \in E_H} \min(x_i x_j, 1) dx_i \right) dx_t \dots dx_1, \quad (6.16)$$

which is the exact same integral as in (6.4). Then we can use the same proof as in the previous lemma to show that this integral is finite. This method works because S induces a complete graph

on H . Indeed, the sets that are constructed in the proof corresponding to S_2 in (2.3) are all subsets of S , and therefore satisfy the constraint (2.4). Then, the sets S_2 constructed in the proof also satisfy (6.11).

When S does not induce a complete graph on H , denote by $M = \{i \in S : d_{i(S)} < |S| - 1\}$ the vertices in S that do not have edges to all other vertices in S . W.l.o.g. assume $M = \{l, \dots, t\}$. We bound $\min(x_i x_j, 1) \leq x_i x_j$ for all $i \in \bar{S}, j \in M$, which bounds (6.15) as

$$\begin{aligned} & \int_1^\infty \cdots \int_1^\infty (x_1 \cdots x_t)^{-\tau} \prod_{u,v \in M: (u,v) \notin E_H} e^{-x_u x_v} \prod_{i=t+1}^k \left(\int_0^1 x_i^{-\tau+d_{i(\bar{S})}} \prod_{j \in S: (i,j) \in E_H} \min(x_i x_j, 1) dx_i \right) dx_t \cdots dx_1 \\ & \leq \int_1^\infty \cdots \int_1^\infty (x_1 \cdots x_l)^{-\tau} \prod_{i=t+1}^k \left(\int_0^1 x_i^{-\tau+d_{i(\bar{S})}+d_{i(M)}} \prod_{j \in S \setminus M: (i,j) \in E_H} \min(x_i x_j, 1) dx_i \right) dx_{l-1} \cdots dx_1 \\ & \quad \times \int_1^\infty \cdots \int_1^\infty \prod_{i=l+1}^t x_i^{-\tau+d_{i(S)}} \prod_{u,v \in M: (u,v) \notin E_H} e^{-x_u x_v} dx_l \cdots dx_t. \end{aligned} \tag{6.17}$$

The last integral is finite, since every vertex in M has at least one missing edge to another vertex in M . Thus, we only need to show that

$$\int_1^\infty \cdots \int_1^\infty (x_1 \cdots x_l)^{-\tau} \prod_{i=t+1}^k \left(\int_0^1 x_i^{-\tau+d_{i(\bar{S})}+d_{i(M)}} \prod_{j \in S \setminus M: (i,j) \in E_H} \min(x_i x_j, 1) dx_i \right) dx_l \cdots dx_1 < \infty. \tag{6.18}$$

We define u_j and f_j as in (6.6) and (6.7) for all $j \in S \setminus M$. If we again let $W_i = \{i \in S \setminus M : u_j = i\}$, then similarly to (6.9) the largest contribution to (6.18) equals

$$C \int_1^\infty \int_{x_1}^\infty \cdots \int_{x_{l-1}}^\infty \prod_{i=1}^t x_i^{-\tau+d_{i,S,\bar{S}}+(\tau-1)|W_i|-2E_{W_i}-E_{W_i,W_i}} dx_l \cdots dx_1. \tag{6.19}$$

We then construct the sets S_1 , S_2 , and S_3 as in the previous lemma. Since $S \setminus M$ induces a complete graph on H , all sets S_2 constructed in the proof satisfy the constraint in (2.4), and thus we can use the proof of the previous lemma to show that this integral is finite. \square

Acknowledgements. This work was supported by NWO TOP grant 613.001.451 and by the NWO Gravitation Networks grant 024.002.003. The work of RvdH is further supported by the NWO VICI grant 639.033.806. The work of JvL is further supported by an NWO TOP-GO grant and by an ERC Starting Grant.

References

- [1] R. Albert, H. Jeong, and A.-L. Barabási. Internet: Diameter of the world-wide web. *Nature*, 401(6749):130–131, 1999.
- [2] A. R. Benson, D. F. Gleich, and J. Leskovec. Higher-order organization of complex networks. *Science*, 353(6295):163–166, 2016.
- [3] T. Bläsius, T. Friedrich, and A. Krohmer. Cliques in hyperbolic random graphs. *Algorithmica*, pages 1–22, 2017.
- [4] M. Boguñá and R. Pastor-Satorras. Class of correlated random networks with hidden variables. *Phys. Rev. E*, 68:036112, 2003.
- [5] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.*, 1(4):311–316, 1980.

- [6] B. Bollobás. *Random Graphs*, volume 74 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, 2 edition, 2001.
- [7] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.
- [8] T. Britton, M. Deijfen, and A. Martin-Löf. Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.*, 124(6):1377–1397, 2006.
- [9] F. Chung and L. Lu. The average distances in random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA*, **99**(25):15879–15882 (electronic), 2002.
- [10] P. Colomer-de Simon and M. Boguñá. Clustering of random scale-free networks. *Phys. Rev. E*, 86:026120, 2012.
- [11] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the internet topology. In *ACM SIGCOMM Computer Communication Review*, volume 29, pages 251–262. ACM, 1999.
- [12] C. Gao and J. Lafferty. Testing network structure using relations between small subgraph probabilities. arXiv:1704.06742, 2017.
- [13] R. van der Hofstad. *Random Graphs and Complex Networks Vol. 1*. Cambridge University Press, 2017.
- [14] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem. Distances in random graphs with finite variance degrees. *Random Structures & Algorithms*, 27(1):76–123, 2005.
- [15] R. van der Hofstad, A. J. E. M. Janssen, J. S. H. van Leeuwen, and C. Stegehuis. Local clustering in scale-free networks with hidden variables. *Phys. Rev. E*, 95:022307, 2017.
- [16] P. van der Hoorn and N. Litvak. *Upper Bounds for Number of Removed Edges in the Erased Configuration Model*, pages 54–65. Springer International Publishing, Cham, 2015.
- [17] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási. The large-scale organization of metabolic networks. *Nature*, 407(6804):651–654, 2000.
- [18] N. Kashtan, S. Itzkovitz, R. Milo, and U. Alon. Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics*, 20(11):1746–1758, 2004.
- [19] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá. Hyperbolic geometry of complex networks. *Phys. Rev. E*, 82(3):036106, 2010.
- [20] D. Marcus and Y. Shavitt. Rage – a rapid graphlet enumerator for large networks. *Comput. Networks*, 56(2):810–819, 2012.
- [21] S. Maslov, K. Sneppen, and A. Zaliznyak. Detection of topological patterns in complex networks: correlation profile of the internet. *Phys. A*, 333:529 – 540, 2004.
- [22] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon. Superfamilies of evolved and designed networks. *Science*, 303(5663):1538–1542, 2004.
- [23] 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.
- [24] J.-P. Onnela, J. Saramäki, J. Kertész, and K. Kaski. Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E*, 71:065103, 2005.

- [25] M. Ostilli. Fluctuation analysis in complex networks modeled by hidden-variable models: Necessity of a large cutoff in hidden-variable models. *Phys. Rev. E*, 89:022807, 2014.
- [26] E. Ravasz and A.-L. Barabási. Hierarchical organization in complex networks. *Phys. Rev. E*, 67:026112, 2003.
- [27] C. Stegehuis. Degree correlations in scale-free null models. arXiv:1709.01085, 2017.
- [28] C. Stegehuis, R. van der Hofstad, J. S. H. van Leeuwen, and A. J. E. M. Janssen. Clustering spectrum of hierarchical scale-free networks. arXiv:1706.01727, 2017.
- [29] C. E. Tsourakakis, J. Pachocki, and M. Mitzenmacher. Scalable motif-aware graph clustering. In *Proceedings of the 26th International Conference on World Wide Web, WWW '17*, pages 1451–1460, Republic and Canton of Geneva, Switzerland, 2017. International World Wide Web Conferences Steering Committee.
- [30] R. van der Hofstad and N. Litvak. Degree-degree dependencies in random graphs with heavy-tailed degrees. *Internet Mathematics*, 10(3-4):287–334, 2014.
- [31] A. Vázquez, R. Pastor-Satorras, and A. Vespignani. Large-scale topological and dynamical properties of the internet. *Phys. Rev. E*, 65:066130, 2002.
- [32] S. Wernicke and F. Rasche. Fanmod: a tool for fast network motif detection. *Bioinformatics*, 22(9):1152–1153, 2006.
- [33] S. Wuchty, Z. N. Oltvai, and A.-L. Barabási. Evolutionary conservation of motif constituents in the yeast protein interaction network. *Nat. Genet.*, 35(2):176–179, 2003.