OPTIMAL SUBGRAPH STRUCTURES IN SCALE-FREE CONFIGURATION MODELS

BY REMCO VAN DER HOFSTAD*,§, JOHAN S. H. VAN LEEUWAARDEN*¶ AND CLARA STEGEHUIS*∥

EINDHOVEN UNIVERSITY OF TECHNOLOGY§, TILBURG UNIVERSITY¶ AND TWENTE UNIVERSITY∥

Subgraphs reveal information about the geometry and functionalities of complex networks. For scale-free networks with unbounded degree fluctuations, we obtain the asymptotics of the number of times a small connected graph occurs as a subgraph or as an induced subgraph. 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 vertices that together span the subgraph. We find that subgraphs typically occur 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 exponent \( \tau \in (2, 3) \) [1, 11, 16, 34]. Many networks have been reported to satisfy these conditions, including metabolic networks, the internet and social networks. Scale-free networks come with the presence of hubs, i.e., vertices of extremely high degrees.

Another property of real-world scale-free networks is that the clustering coefficient (the probability that two neighbors of a vertex are neighbors themselves) decreases with the vertex degree [34, 18, 26, 4, 10], again following a power law. Thus, two neighbors of a hub are less likely to connect. The triangle is the most studied network subgraph, because it not only describes the clustering coefficient, but also signals hierarchy and community structure [23]. However, other subgraphs such as larger cliques are equally important for understanding network organization [2, 33]. Indeed, subgraph
counts might vary considerably across different networks [20, 35, 19] and any given network may have a set of statistically significant subgraphs (also called motifs). Statistical relevance can be expressed by comparing a real-world network to some mathematically tractable model. This comparison filters out the effect of the degree sequence and the network size on the subgraph count. A popular statistic takes the subgraph count, subtracts the expected number of subgraphs in a model, and divides by the standard deviation in the model [19, 21, 12]. Such a standardized test statistic sheds light on whether a subgraph is overrepresented in comparison to the model. This raises the question of what 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 a 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, 27]. For $\tau \in (2, 3)$, however, the configuration model fails to create simple graphs with high probability [15]. We therefore consider the erased configuration model [8], which constructs a configuration model and then removes all self-loops and merges multiple edges. For an erased configuration model with degree exponent $\tau \in (2, 3)$, we count how often a small connected graph $H$ occurs as a subgraph or as an induced subgraph, where edges not present in $H$ are also not allowed to be present in the subgraph.

We find that every (induced) subgraph $H$, occurs typically between vertices in the erased configuration model with degrees in highly specific ranges that depend on the precise subgraph $H$. An example of these typical degree ranges for subgraphs on 4 vertices is shown in Figure 1 (which will be discussed in more detail in Section 2.4). In this paper we show that many subgraphs consist exclusively of $\sqrt{n}$-degree vertices, including cliques of all sizes. Hence, in such subgraphs, hubs (of degree close to the maximal value $n^{1/(\tau-1)}$) are unlikely to participate in a typical subgraph. Hubs can be part, however, of other subgraphs. We define optimization problems that find these optimal degree ranges for every subgraph.

We next define the model.

The erased configuration model. Let $[n] = \{1, 2, \ldots, n\}$. Given a degree sequence, i.e., a sequence of $n$ positive integers $D = (D_1, D_2, \ldots, D_n)$, the configuration model is a (multi)graph where vertex $i$ has degree $D_i$. It is defined as follows, see e.g., [6] or [27, Chapter 7]: given a degree sequence with $\sum_{i \in [n]} D_i$ even, we start with $D_j$ free half-edges adjacent to vertex $j$, for $j = 1, \ldots, n$. The configuration model is constructed by successively pairing, uniformly at random, free half-edges into edges and removing them from the set of free half-edges, until no free half-edges remain. Conditionally
on obtaining a simple graph, the resulting graph is a uniform sample from the ensemble of simple graphs with the prescribed degree sequence [27, Chapter 7]. This is why the configuration model is often used as a 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 erased edges is large [32] (yet small compared to the total number of edges). In particular, we take the degrees to be an i.i.d. copies of a random variable $D$ such that

$$\Pr(D = k) = ck^{-\tau}(1 + o(1)), \quad \text{as } k \to \infty,$$

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, $D_{\text{max}}$ is of order $n^{1/\tau - 1}$, where $D_{\text{max}}$ denotes the maximal degree of the degree sequence. Denote the erased configuration model on $n$ vertices by ECM$^{(\alpha)}(n)$ when the degrees are an i.i.d. sample of (1.1), and ECM$^{(\alpha)}(D)$ when the degree sequence equals $D$.

*Quenched and annealed.* Note that the erased configuration model as defined above has two sources of randomness: the i.i.d. degrees and the random pairing of the half-edges in constructing the graph. Studying the behavior of subgraphs in the erased configuration model once the degree sequence has been fixed corresponds to the *quenched* setting, whereas the erased configuration model with random degrees corresponds to the *annealed* setting. Our
main result on the number of subgraphs in the erased configuration model is in the annealed setting. However, in the proof of our results we often study subgraph counts in the quenched setting. Throughout this paper, we denote the probability of an event $E$ in the quenched setting by

$$P_n(E) = \mathbb{P}(E \mid (D_i)_{i \in [n]}),$$

and we define $E_n$ and $\text{Var}_n$ accordingly.

**Subgraph counts.** Let $H = (V_H, E_H)$ be a small, connected graph. We denote the induced subgraph count of $H$, the number of subgraphs of ECM$^{(n)}$ that are isomorphic to $H$, by $N^{(\text{ind})}(H)$. We denote the subgraph count, the number of occurrences of $H$ as a subgraph of ECM$^{(n)}$, by $N^{(\text{sub})}(H)$.

Throughout this paper, we denote the sampled degree of a vertex $i \in [n]$ in the erased configuration model by $D_i$. Note that this may not be the same as the actual degree of a vertex in the erased configuration model, since self-loops are removed and multiple edges are merged. Since we study subgraphs $H$, we sometimes also need to use the degree of a vertex in $H$ inside the subgraph. We denote the degree of a vertex $i$ of a subgraph $H$ by $d_i$.

**Paper outline.** We present our main results in Section 2, including the theorems that characterize all optimal 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 provide an overview of the proof structures in Section 3. We then prove the first part of the main theorems for subgraphs in Section 4 and for $\sqrt{n}$-optimal subgraphs in Section 5. The proofs of some lemmas introduced along the way are deferred to Section 6. The proof of the second part of the main theorem can be found in Section 5. We finally show how the proofs for subgraphs can be adjusted to prove the theorems on induced subgraphs in Section 8.

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

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
OPTIMAL SUBGRAPH STRUCTURES IN CONFIGURATION MODELS

present in power-law networks: on the one hand, hubs contribute substantially to the subgraph count, because they are well connected, and therefore potentially contribute to many subgraphs. 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 vertices with specific degrees that ‘optimizes’ this trade-off and hence maximizes the probability that the subgraph occurs.

Let $\text{ECM}^{(n)}|\mathbf{v}$ denote the induced subgraph of the erased configuration model on vertices $\mathbf{v}$. Write the probability that a subgraph $H = (V_H,E_H)$ with $|V_H| = k$ is created on $k$ uniformly chosen vertices $\mathbf{v}$ in $\text{ECM}^{(n)}$ as

$$P\left(\text{ECM}^{(n)}|\mathbf{v} \supseteq H\right) = \sum_{D'} P\left(\text{ECM}^{(n)}|\mathbf{v} \supseteq H \mid D_{\mathbf{v}} = D'\right) P\left(D_{\mathbf{v}} = D'\right),$$

where the sum is over all possible degrees on $k$ vertices $D' = (D'_v)_{v \in [k]}$, and $D_{\mathbf{v}} = (D_{v_i})_{i \in [k]}$ denotes the degrees of the randomly chosen set of $k$ vertices.

We show that for every (induced) subgraph, there is a specific range of $D'_1, \ldots, D'_k$ that gives the maximal contribution to (2.1), large enough even to completely ignore all other degree ranges.

We show that when (2.1) is maximized by a unique range of degrees, there are only four possible ranges of degrees that maximize the term inside the sum in (2.1). These ranges are constant degrees, or degrees proportional to $n^{(r-2)/(r-1)}$, to $\sqrt{n}$ or to $n^{1/(r-1)}$.

2.1. An optimization problem We now present the optimization problems that maximizes the summand in (2.1), first for subgraphs and later for induced subgraphs. 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 inside $H$ 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 the following typical orders of magnitude of the degrees of the vertices of $H$ embedded in $\text{ECM}^{(n)}$: $S_1$ denotes the vertices with degree proportional to $n^{(r-2)/(r-1)}$, $S_2$ the ones with degrees proportional to $n^{1/(r-1)}$, 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 that maximizes the contribution to the number of (induced) subgraphs. When a vertex in $H$ has degree 1, its degree in $\text{ECM}^{(n)}$ is typically small, i.e., it does not grow in $n$.

Given a partition $\mathcal{P}$, let $E_{S_1}$ denote the number of edges in $H$ between vertices in $S_1$, $E_{S_i,S_j}$ the number of edges between vertices in $S_i$ and $S_j$ and $E_{S_i,V_1}$ the number of edges between vertices in $V_1$ and $S_i$. We now define the
optimization problem for subgraphs that optimizes the summand in (2.1) as
\[
B^{(\text{sub})}(H) = \max_{\mathcal{P}} \left| S_1 \right| - \left| S_2 \right| - \frac{2E_{S_1} + E_{S_1,S_3} + E_{S_1,V_1} - E_{S_2,V_1}}{\tau - 1}.
\]

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 subgraph. This captures the other part of the trade-off: high-degree vertices are more likely to connect to other vertices than low degree vertices. Note that \( B^{(\text{sub})}(H) \geq 0 \), since putting all vertices in \( S_3 \) yields zero.

For induced subgraphs, we define a similar optimization problem
\[
B^{(\text{ind})}(H) = \max_{\mathcal{P}^{(\text{ind})}} \left| S_1 \right| - \left| S_2 \right| - \frac{2E_{S_1} + E_{S_1,S_3} + E_{S_1,V_1} - E_{S_2,V_1}}{\tau - 1},
\]
\[
\text{s.t. } (u,v) \in E_H \quad \forall u \in S_2, v \in S_2 \cup S_3,
\]
where again \( \mathcal{P}^{(\text{ind})} \) is a partition of \( V_H \setminus V_1 \) into three sets. The constraint in (2.3) ensures that edges that are not present in \( H \) are not present in the subgraph. Again, \( B^{(\text{ind})}(H) \geq 0 \) because \( S_3 = V_H \setminus V_1 \) is a valid solution.

Our main result shows that indeed the optimization problems (2.2) and (2.3) find the typical vertex degrees for any (induced) subgraph and determine the scaling of the number of subgraphs. We then investigate a special class of subgraphs, where the optimal contribution to (2.2) or (2.3) is \( S_3 = V_H \), i.e., (induced) subgraphs where all typical vertex degrees are 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 \( \alpha = (\alpha_1, \cdots, \alpha_k) \) such that \( \alpha_i \in \left[ 0, 1/(\tau - 1) \right] \), define
\[
M_n^{(\alpha)}(\varepsilon) = \{(u_1, \cdots, u_k) : D_{u_i} \in [\varepsilon, 1/\varepsilon](\mu n)^{\alpha_i} \forall i \in [k]\}.
\]

These are the sets of degrees such that \( D_1 \) is proportional to \( n^{\alpha_1} \) and \( D_2 \) proportional to \( n^{\alpha_2} \) and so on. Denote the number of subgraphs with vertices in \( M_n^{(\alpha)}(\varepsilon) \) by \( N^{(\text{sub})}(H, M_n^{(\alpha)}(\varepsilon)) \). Define the vector \( \alpha^{(\text{sub})} \) as
\[
\alpha_i^{(\text{sub})} = \begin{cases} 
(\tau - 2)/(\tau - 1) & i \in S_1^{(\text{sub})}, \\
1/(\tau - 1) & i \in S_2^{(\text{sub})}, \\
\frac{1}{2} & i \in S_3^{(\text{sub})}, \\
0 & i \in V_1.
\end{cases}
\]
For induced subgraphs, let $S_{1}^{(\text{ind})}, S_{2}^{(\text{ind})}, S_{3}^{(\text{ind})}$ be a maximizer of (2.3), and define $\alpha^{(\text{ind})}$ as in (2.5), replacing $S_{i}^{(\text{sub})}$ by $S_{i}^{(\text{ind})}$. The next theorem shows that sets of vertices in $M_{n}^{\alpha^{(\text{sub})}}(\varepsilon)$ or $M_{n}^{\alpha^{(\text{ind})}}(\varepsilon)$ contain a large number of subgraphs, and computes the scaling of the number of (induced) subgraphs:

**Theorem 2.1 (General (induced) subgraphs).** Let $H$ be a subgraph on $k$ vertices such that the solution to (2.2) is unique.

(i) For any $\varepsilon_{n}$ such that $\lim_{n \to \infty} \varepsilon_{n} = 0$,

$$\frac{N^{(\text{sub})}(H, M_{n}^{\alpha^{(\text{sub})}})(\varepsilon_{n})}{N^{(\text{sub})}(H)} \xrightarrow{p} 1.$$  

(ii) Furthermore, for any fixed $0 < \varepsilon < 1$,

$$\frac{N^{(\text{sub})}(H, M_{n}^{\alpha^{(\text{sub})}})(\varepsilon))}{n^{\frac{3}{2} + \frac{2}{k_{2}^{+} + \beta^{(\text{sub})}} + k_{1}/2}} \leq f(\varepsilon) + o_{p}(1),$$

and

$$\frac{N^{(\text{sub})}(H, M_{n}^{\alpha^{(\text{sub})}})(\varepsilon))}{n^{\frac{3}{2} + \frac{2}{k_{2}^{+} + \beta^{(\text{sub})}} + k_{1}/2}} \geq \tilde{f}(\varepsilon) + o_{p}(1),$$

for some functions $f(\varepsilon), \tilde{f}(\varepsilon) < \infty$ 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 the subgraph $H$.

For induced subgraphs the same statements hold, replacing (sub) by (ind).

Theorem 2.1(ii) only provides the scaling in $n$ and some function $f(\varepsilon)$, which could tend to $\infty$ when $\varepsilon \downarrow 0$. For subgraphs with $S_{3} = V_{H}$, we obtain more precise asymptotics in the next section.

2.3. **Sharp asymptotics for $\sqrt{n}$-class of subgraphs** Now we study the special class of subgraphs for which the unique maximum of (2.2) or (2.3) is $S_{3} = V_{H}$. By the above interpretation of $S_{1}, S_{2}$ and $S_{3}$, we study (induced) subgraphs where the maximum contribution to the number of such subgraphs comes from vertices that have degrees proportional to $\sqrt{n}$ in ECM. Examples of subgraphs that fall into this category are all complete graphs. Bipartite graphs on the other hand, do not fall into the $\sqrt{n}$-class subgraphs, since we can use the two parts of the bipartition 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 $\sqrt{n}$-(induced) subgraphs:
Theorem 2.2 ((Induced) subgraphs 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,

\[
N^{(\text{sub})}(H) \frac{n^{3/2}}{\mu} \overset{p}{\to} A^{(\text{sub})}(H) < \infty,
\]

with

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

For induced subgraphs the same statements hold, replacing \( \text{(sub)} \) by \( \text{(ind)} \) and (2.2) by (2.3), where

\[
A^{(\text{ind})}(H) = c^k \mu^{-k/2} (\tau - 1) \int_0^\infty \cdots \int_0^\infty (x_1 \cdots x_k)^{-\tau} \prod_{(u,v) \in E_H} (1 - e^{-x_u x_v})
\]

\[
\times \prod_{(u,v) \notin E_H} e^{-x_u x_v} dx_1 \cdots dx_k.
\]

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, and taking \( D_i \approx x_u \sqrt{n/\mu} \) and \( D_j \approx x_v \sqrt{n/\mu} \), this results in the factors \( 1 - e^{-x_u x_v} \) in (2.10) for all edges in subgraph \( H \). For induced subgraphs, the fact that no other edges than the edges in \( H \) are allowed to be present gives the extra terms \( e^{-x_u x_v} \) in (2.11).

2.4. Subgraphs on 4 and 5 vertices We now apply Theorem 2.1 to characterize the optimal subgraph configurations on 4 or 5 vertices. We 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 Theorem 2.1 to calculate the scaling of the number of such (induced) subgraphs. Figures 1 and 2 show the order of magnitude of the number of induced subgraphs on 4 and 5 vertices obtained in this way, together with the optimizing sets of (2.3). For example, the optimal values of \( S_1, S_2 \) and \( S_3 \) for the subgraph in Figure 1d show that

\[
B^{(\text{ind})}(H) = 2 - 1 + \frac{2 + 0 + 0 - 1}{\tau - 1} = 1 + \frac{1}{\tau - 1}.
\]
By Theorem 2.1, the scaling of the induced subgraph in Figure 1d then equals

\[(2.13) \quad n^{(3-\tau)(4-1/(\tau-1))/2+1/2} = n^{7-2\tau-\frac{1}{\tau-1}}.\]

The scaling of the other induced subgraphs are computed similarly.

Most induced subgraphs in Figures 1 and 2 satisfy the constraint in Theorem 2.1 that the solution to the optimization problem (2.2) or (2.3) should be unique. However, the gray vertices in Figure 1 do not have unique optimizers, so that our theorems do not apply. Still, similar analysis as in Section 4.2 shows that there exist ranges of degrees that give the major
contribution to the rescaled number of such (induced) subgraphs. The only difference is that these ranges are wider than for the vertices with unique maximizers. For example, for the diamond subgraph in Figure 1b the major contribution is from vertices where the degrees of vertices at each side of an edge \( \{i, j\} \) in the square around the diamond satisfy \( D_i D_j = \Theta(n) \). Note that having all degrees proportional to \( \sqrt{n} \) therefore is one of the main contributors. However, contributions where the bottom left vertex and the top right vertex have degrees proportional to \( n^\alpha \) and the other two vertices have degrees \( n^{1-\alpha} \) give an equal contribution for other values of \( \alpha \). Using that \( D_i D_j \) follows a power-law distribution with exponent \( \tau \) with an extra factor \( \log(n) \) \([30, Eq. (2.16)]\) then gives the extra factor \( \log(n) \) in Figure 1b.

The bow tie in Figure 2i has a unique optimal solution to (2.2), but it depends on \( \tau \). For \( \tau < 7/3 \), the maximum of (2.3) is uniquely attained at 0, so that the optimal composition is with all vertices of degree \( \Theta(\sqrt{n}) \). On the other hand, when \( \tau > 7/3 \), \( S_1 \) contains the degree 2 vertices while the middle vertex is in \( S_2 \). This partition gives a contribution to (2.3) of

\[
4 - 1 - \frac{2 \cdot 2}{\tau - 1} = \frac{3\tau - 7}{\tau - 1},
\]

which is larger than zero if \( \tau > 7/3 \). Thus, for \( \tau \) 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 the maximal contribution to an induced subgraph comes from vertices with degrees proportional to \( \sqrt{n} \), then by Theorem 2.2, the number of such induced subgraphs converges to a constant when properly rescaled. When the maximal contribution contains vertices in \( S_2 \) and \( S_1 \), this may not hold. For example, counting the number of induced claws of Figure 1e is similar to counting the number of sets of three neighbors for every vertex. The only sets of neighbors we do not count, are neighbors that are connected. This is a small fraction of the pairs of neighbors \([29, E1, (5)-(7)]\), thus the number of claws is approximately equal to

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

Since the degrees are an i.i.d. sample from a power-law distribution, \( \sum_i D_i^3 \) converges to a stable law when normalized properly. Thus, when vertices of degrees proportional to \( n^{1/(\tau-1)} \) contribute, the leading order of the number of (induced) subgraphs may contain stable random variables, in contrast with the deterministic leading order for \( \sqrt{n} \) degrees of Theorem 2.2.
The scaling of the number of (non-induced) subgraphs can be deduced from Figure 1. For example, we count the number of square subgraphs (the subgraph of Figure 1c) by adding the contributions from the induced subgraphs in Figures 1a, 1b and 1c, that all contain a square, which shows that a square occurs $\Theta(n^{6-2\tau}\log(n))$ times as a subgraph. The major contribution to the number of square subgraphs is from the induced subgraphs in Figure 1b, which indeed contains a square, and occurs more frequently than the subgraphs of Figures 1a and 1c. In this manner we can infer the order of magnitude of the number of subgraphs from the number of induced subgraphs.

2.5. Discussion and outlook

Uniqueness of the optimum. Theorem 2.1 only holds when the optimum of (2.2), respectively (2.3), is unique. Figures 1 and 2 show that for most subgraphs on 4 or 5 vertices, this is indeed the case. In Section 4, we show that (2.2) and (2.3) can both be interpreted as piecewise linear optimization problems over the optimal degrees of the vertices that together form the subgraph. Thus, if the optimum is not unique, then it is attained by an entire range of degrees. In Section 4 we show that in this situation the optimum is attained for degrees such that $D_i D_j = \Theta(n)$ across some edges $\{i,j\}$. One such example is the diamond of Figure 1b discussed in Section 2.4. We believe that the number of subgraphs where the optimum is not unique scales as in Theorem 2.1 with some additional multiplicative factors of $\log(n)$. Proving this remains open for further research.

Automorphisms of $H$. An automorphism of a graph $H$ is a map $V_H \mapsto V_H$ such that the resulting graph is isomorphic to $H$. In Theorems 2.1 and 2.2 we count automorphisms of $H$ as separate copies of $H$, so that we may count multiple copies of $H$ on one set of vertices. Since $|V_H|$ is fixed, and Theorem 2.1 only considers the scaling of the number of subgraphs, this does not influence Theorem 2.1. Because Theorem 2.2 studies the exact scaling of the number of subgraphs, to count the number of subgraphs without automorphisms, one should divide the results of Theorem 2.2 by the number of automorphisms of $H$.

Self-averaging. A random variable is called self-averaging if its coefficient of variation tends to zero, otherwise it is called non-self-averaging. When the degree distribution follows a power-law with exponent $\tau \in (2,3)$, the number of subgraphs may be non-self-averaging [22], so that

$$\limsup_{n \to \infty} \frac{\text{Var}(N^{(\text{sub})}(H))}{\mathbb{E}[N^{(\text{sub})}(H)]^2} \neq 0.$$
One such example is the triangle. While the triangle subgraph satisfies the conditions of Theorem 2.2, so that the rescaled number of triangles converges in probability to a constant, it was shown in [22], the number of triangles is non-self-averaging in the annealed sense when $\tau$ is close to 3. This indicates that most realizations of $\text{ECM}^{(\alpha)}(D)$ will have a number of triangles that is close to the value predicted by Theorem 2.2. However, since the number of triangles is non-self-averaging making its standard deviation quite large, some realizations will have a number of triangles that is much larger or smaller than the value predicted in Theorem 2.2.

Other random graph models. An interesting question is whether Theorems 2.1 and 2.2 also apply to other models that create simple power-law random graphs. A very natural model for simple power-law random graphs is the uniform random graph, which samples a uniform graph from the ensemble of all simple graphs on a given degree sequence, which we plan to analyze using similar techniques as in this paper [13].

Another random graph model that generates simple power-law random graphs is the rank-1 inhomogeneous random graph [9, 4]. In this model, vertices have weights $h_i$, where the weights are an i.i.d. sample of a power-law random variable with exponent $\tau \in (2, 3)$. Then, two vertices are connected with probability $f_n(h_i, h_j)$. Two standard connection probability functions are $f_n(h_i, h_j) = \min(h_i h_j / (\mu n), 1)$ [9], and $f_n(h_i, h_j) = 1 - e^{-h_i h_j / (\mu n)}$ [7]. Conditionally on the weight sequence, the edge statuses are independent, which is different from the erased configuration model, where the edge statuses are not independent, even when conditioning on the degree sequence. We prove Theorems 2.1 and 2.2 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, Theorems 2.1 and 2.2 hold also for the rank-1 inhomogeneous random graph with these connection probabilities instead [25].

A third model that creates simple power-law random graphs, is the hyperbolic random graph where vertices are sampled in a disk, and connected if their hyperbolic distance is sufficiently small [17]. The geometry in the hyperbolic random graph makes the presence of triangles and other subgraphs containing cycles likely. By Theorem 2.2, a complete graph on $k$ vertices occurs $\Theta(n^{\frac{1}{2}(3-\tau)})$ times as a subgraph in $\text{ECM}^{(\alpha)}$. Interestingly, this is also true for hyperbolic random graphs for $k$ sufficiently large [3]. It would be interesting to investigate the presence of other subgraphs in hyperbolic random graphs.

Another class of popular models, which create simple power-law random graphs dynamically are those which incorporate preferential attachment. In these models, subgraph counts scale significantly differently from the erased
configuration model and uniform random graphs [14].

3. Overview of the proofs We now provide an overview of the proof structure of Theorems 2.1 and 2.2. Our main results study the annealed version ECM\((n)\), with random degree sequence. In the proofs of Theorems 2.1 and 2.2, we often first study the quenched version of ECM\((n)\)(\(D\)) instead, where the degree sequence \(D\) is fixed.

We relate \(L_n = \sum_{i \in [n]} D_i\), the total number of half-edges before erasure, to its expected value \(\mu n\) by defining the event

\[
J_n = \{ |L_n - \mu n| \leq n^{2/\tau} \}.
\]

By [31, Lemma 2.3], \(P(J_n) \to 1\) as \(n \to \infty\). When we condition on the degree sequence, we will work on the event \(J_n\), so that we can write \(L_n = \mu n(1 + o(1))\). Similarly, when we work with \(\mathbb{E}_n\) and \(\text{Var}_n\), we condition on the event \(J_n\). We do not include \(J_n\) into the notation of \(P_n\), since given \(D\), \(J_n\) either happens with probability one, or with probability zero. This could be treated more formally by denoting

\[
P_n(E) = 1_{J_n}P(E | D),
\]

but keep notation light, when using \(P_n\), we always assume the event \(J_n\) to hold.

Denote by \(X_{u,v}\) the indicator that an edge is present between vertices \(u\) and \(v\). To obtain the probability that a specific subgraph is present on a given set of vertices, we investigate the probability of a set of edges being present in the erased configuration model. In ECM\((n)\)(\(D\)) (see the proof of Lemma 4.1 for a more precise statement),

\[
P_n(X_{u,v} = 1) \approx 1 - e^{-D_u D_v / L_n}.
\]

However, subgraphs often contain more than just one edge, and edges in ECM\((n)\)(\(D\)) are not present independently. In Section 4.1, we show that these dependencies are weak, so that we can use the approximation (3.3) for all edges in a subgraph as if they were present independently.

We then compute the probability that a subgraph is present on a specific set of vertices as a function of their degrees, which shows that

\[
\Theta_{\varepsilon}\left(n^{k+(1-\tau)\sum_i \alpha_i} \prod_{\{i,j\} \in E_H: \alpha_i + \alpha_j < 1} n^{\alpha_i + \alpha_j - 1}\right).
\]

To prove Theorem 2.1(ii) we optimize this over \(\alpha = (\alpha_1, \ldots, \alpha_k)\). Here \(\varepsilon\) does not appear in the scaling, since it is independent of \(n\). To prove
Theorem 2.1(i) for $\varepsilon_n \downarrow 0$, we analyze $N^{(\mathrm{sub})}(H, M_n^{(\alpha)}(\varepsilon))$ in more detail in Section 7.

To prove the sharp asymptotics of Theorem 2.2, we compute the contribution to the expectation and the variance of the number of subgraphs in $\text{ECM}^{(\alpha)}(D)$ from vertices with degrees proportional to $\sqrt{n}$ in Section 6. We use a second moment method to show that the number of subgraphs concentrates around its expectation in $\text{ECM}^{(\alpha)}(D)$. We then investigate the asymptotic behavior of this expectation. A first moment method which shows that the expected contribution to the number of subgraphs from vertices with other degrees is small completes the proof of Theorem 2.2.

Theorem 2.2 for induced subgraphs can be proven similarly, the only difference being that we have to take into account that to form an induced subgraph, some edges are not allowed to be present in $\text{ECM}^{(\alpha)}(D)$. We explain how this changes the proof of Theorem 2.2 in more detail in Section 8.

4. Maximum contribution: proof of Theorem 2.1

4.1. The probability of avoiding a subgraph  

The edges of a subgraph are not present independently. The following lemma computes the probability that an edge is not present conditionally on other edges not being present:

**Lemma 4.1.** Fix $m \in \mathbb{N}$ and $\varepsilon > 0$. Let $(u_i, v_i)_{i \in [m+1]}$ be such that $u_i, v_i \in [n]$ for all $i \in [m+1]$ and $(u_{m+1}, v_{m+1}) \neq (u_i, v_i)$ for all $i \in [m]$. Let

$$E = \{X_{u_i, v_i} = 0, \forall i \in [m]\}.$$  

If $D_{u_i} D_{v_i} \leq n^{1/(\tau-1)}/\varepsilon$ for $i \in [m+1]$, then,

$$P_n \left( X_{u_{m+1}, v_{m+1}} = 0 \mid E \right) = O \left( e^{-\frac{D_{u_{m+1}} D_{v_{m+1}}}{4 L_n}} \right).$$

Furthermore, when $D_{u_{m+1}} D_{v_{m+1}} \leq n/\varepsilon$, for $\gamma \in \left( \frac{\tau-2}{2(\tau-1)}, \frac{\tau-2}{\tau-1} \right)$,

$$P_n \left( X_{u_{m+1}, v_{m+1}} = 0 \mid E \right) = e^{-\frac{D_{u_{m+1}} D_{v_{m+1}}}{L_n}} \left( 1 + O \left( \frac{D_{u_{m+1}} D_{v_{m+1}}}{L_n} n^{-\gamma} \right) \right).$$

Throughout the rest of the paper, we mainly use (4.2) to bound the probability that an edge between two high-degree vertices is absent, whereas we use (4.3) to compute asymptotic identities on the probability that a subgraph is present.
OPTIMAL SUBGRAPH STRUCTURES IN CONFIGURATION MODELS

Proof. For \( m = 0 \) the claim is proven in [28, Eq (4.6) and (4.9)], which states that for two vertices \( i \) and \( j \) with \( D_i > D_j \),

\[
P_n(X_{i,j} = 0) = e^{-D_i D_j / L_n} + O(D_i^2 D_j^2 / L_n^2),
\]

and that, by using [28, Eq. (4.5)] as well as \( L_n - 2i + i \leq L_n \) and \( 1 - x \leq e^{-x} \),

\[
P_n(X_{i,j} = 0) \leq D_i / 2 \prod_{i=1}^{D_i/2} \left( 1 - \frac{D_j}{L_n - 2i - 1} \right) \leq e^{-D_i D_j / 2L_n}.
\]

Thus we assume that \( m > 0 \). Note that \( \Omega := \{u_i, v_i\}_{i \in [m]} \) may contain the same vertices multiple times. Let the number of distinct vertices in \( \{u_i, v_i\}_{i \in [m]} \) be denoted by \( r \), and let these distinct vertices be denoted by \( w_1, \ldots, w_r \). Let \( u_{m+1}, v_{m+1} \) correspond to \( w_r \) and \( w_{r-1} \) (if they are present in \( w_1, \ldots, w_r \) at all). The ordering of the other vertices may be arbitrary. We now construct ECM(\( \omega \))(\( D \)) conditionally on the edges \( \Omega \) not being present. We pair the half-edges of the erased configuration model attached to \( w_1, \ldots, w_r \). First we pair all half-edges adjacent to \( w_1 \). Since we condition on the edges \( \Omega \) not being present, no half-edge from \( w_1 \) is allowed to pair to any of its neighbors in \( \Omega \). 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 \( \Omega \), and so on. We continue until all of the forbidden edges \( \Omega \) have at least one incident vertex whose half-edges have already been paired. Then, if we pair the rest of the half-edges, we know that none of the edges in \( \Omega \) are present. Let \( B \) denote the number of vertices we have to pair before all of the forbidden edges \( \Omega \) have at least one incident vertex whose half-edges have already been paired. We never have to pair half-edges adjacent to \( u_{m+1} \) or to \( v_{m+1} \) (if they are present in \( \{u_i, v_i\}_{i \in [m]} \)), since they are last in the ordering, and \( \{u_{m+1}, v_{m+1}\} \) is not present in \( \{u_i, v_i\}_{i \in [m]} \). Therefore, the half-edges incident to all forbidden neighbors of \( u_{m+1} \) and \( v_{m+1} \) in \( \Omega \) have already been paired before arriving at \( u_{m+1} \) or \( v_{m+1} \). Let \( \hat{X}_{ij} \) denote the number of half-edges between \( i \) and \( j \) in the configuration model, so that the edge indicator of the erased configuration model can be written as \( X_{ij} = 1 \{ \hat{X}_{ij} > 0 \} \). Furthermore, let \( F_{\leq s} = \sigma((\hat{X}_{w_i,j})_{i \leq s, j \in [n]}) \) be the information about the pairings that have been constructed up to time \( s \).

After \( B \) pairings, denote

\[
\tilde{L}_n = L_n - 2 \sum_{i \in [B]} (D_{w_i} - \hat{X}_{w_i,w_i}),
\]

which equals the remaining half-edges after pairing the half-edges incident to \( (w_i)_{i \in [r]} \). Here we subtract \( D_{w_i} \) twice, since the pairing of every half-edge
removes one half-edge incident to $w_i$, and one other half-edge, unless it is paired to another half-edge incident to $w_i$, giving rise to the term $\hat{X}_{w_i,u}^m$. Define $\hat{D}_{um+1} = D_{um+1} - \sum_{i \in [B]} \hat{X}_{i,um+1}$, and define $D_{vm+1}$ similarly. These quantities are all measurable with respect to $\mathcal{F}_B$. The probability that $u_{m+1}$ does not pair to $v_{m+1}$ is the probability that $u_{m+1}$ of degree $\hat{D}_{um+1}$ does not connect to $v_{m+1}$ of degree $\hat{D}_{vm+1}$ in a configuration model with $\hat{L}_n$ half-edges. Thus, using (4.4),
\begin{equation}
\mathbb{P} (X_{um+1,vm+1} = 0 \mid \mathcal{F}_B) = e^{\hat{D}_{um+1} \hat{D}_{vm+1} / \hat{L}_n} + O \left( \frac{\hat{D}_{um+1} \hat{D}_{vm+1} / \hat{L}_n^2}{\hat{D}_{um+1}} \right),
\end{equation}
where we have assumed w.l.o.g. that $\hat{D}_{um+1} \geq \hat{D}_{vm+1}$. Since $D_{u_i}, D_{v_i} \leq n^{1/(\tau - 1)} \varepsilon$, $\hat{L}_n \geq \hat{L}_n/2$.

By (4.5),
\begin{align}
\mathbb{P} (X_{um+1,vm+1} = 0 \mid \mathcal{F}_B) & \leq e^{-D_{um+1} \hat{D}_{vm+1} / 2 \hat{L}_n} \leq e^{-D_{um+1} \hat{D}_{vm+1} / 4 \hat{L}_n} \\
& = O \left( e^{-D_{um+1} \hat{D}_{vm+1} / 4 \hat{L}_n} \right).
\end{align}
which proves (4.2).

We now proceed to prove (4.3). The probability that the $j$th half-edge incident to $w_i$ pairs to $u_{m+1}$ can be bounded as
\begin{equation}
\mathbb{P} (j \text{th half-edge pairs to } u_{m+1}) \leq \frac{D_{um+1}}{\hat{L}_n - 2j - 3 - 2 \sum_{s \in [1]} D_{w_s}} \leq KD_{um+1} / \hat{L}_n,
\end{equation}
for some $K > 0$. We have to pair at most $D_{w_i} \leq n^{1/(\tau - 1)} \varepsilon$ half-edges, since some of the half-edges incident to $w_i$ may have been used already in previous pairings. Therefore, we can stochastically dominate $\hat{X}_{w_i,u}^m$ by $Y_{w_i}$, where $Y_{w_i} \sim \text{Bin}(n^{1/(\tau - 1)} \varepsilon, KD_{um+1}/\hat{L}_n)$, so that $\mathbb{E} [Y_i] = K_1 n^{-\beta} D_{um+1}$ for some $K_1$, where $\beta = (\tau - 2)/(\tau - 1)$.

Choose $\gamma \in (\frac{\tau - 2}{2(\tau - 1)}, \frac{\tau - 2}{\tau - 1})$. By the Chernoff bound, for some $\hat{K} > 0$,
\begin{equation}
\mathbb{P} \left( Y_{w_i} > K_1 n^{-\beta} D_{um+1} (1 + n^\gamma) \right) \leq e^{-\hat{K} n^{2\gamma - \beta} D_{um+1}}.
\end{equation}
Define the events
\begin{align}
\mathcal{B}_n &= \{ \exists i \in [B] : \hat{X}_{w_i,u}^m > K_1 n^{-\beta} D_{um+1} (1 + n^\gamma) \}, \\
\mathcal{B}_n &= \{ \exists i \in [B] : \hat{X}_{w_i,v}^m > K_1 n^{-\beta} D_{vm+1} (1 + n^\gamma) \},
\end{align}
and let $\mathcal{B}_{n,u}^c$ and $\mathcal{B}_{n,v}^c$ denote their respective complements, so that, by a union bound,

$$\Pr\left(\mathcal{B}_{n,u}^c\right) \geq 1 - Be^{-Kn^{2\gamma-\beta}D_{um+1}}.$$  \hfill (4.13)

On the event $\mathcal{B}_{n,u}^c$,

$$\tilde{D}_{um+1} \geq D_{um+1} (1 - \sum_{i \in [B]} \tilde{X}_{ui,um+1}) = D_{um+1} (1 + O(n^{\gamma-\beta})).$$  \hfill (4.14)

Similarly, $\tilde{D}_{vm+1} = D_{vm+1} (1 + O(n^{\gamma-\beta}))$ on $\mathcal{B}_{n,v}^c$, where $\Pr\left(\mathcal{B}_{n,v}^c\right) \geq 1 - Be^{-Kn^{2\gamma-\beta}D_{vm+1}}$. Then, when $D_{um+1}D_{vm+1} = O(n)$ as assumed for (4.3), (4.7) becomes

$$\Pr_n\left( X_{um+1,vm+1} = 0 \mid \mathcal{F}_{B+1}, \mathcal{B}_{n,u}^c, \mathcal{B}_{n,v}^c \right) = e^{-\frac{D_{um+1}D_{vm+1}}{L_n} (1+O(n^{-\gamma}))} + O\left( \frac{D_{um+1}^2D_{vm+1}}{L_n^2} \right)$$

$$= e^{-\frac{D_{um+1}D_{vm+1}}{L_n} \left( 1 + O\left( \frac{D_{um+1}D_{vm+1}}{L_n} n^{-\gamma} \right) \right)},$$  \hfill (4.15)

where we have used that $D_{um+1} = O(n^{1/(\tau-1)})$. Furthermore, $2\gamma - \beta > 0$, whereas by assumption $D_{um+1}D_{vm+1}/L_n = O(1)$, so that (4.13) together with (4.15) proves (4.3).

4.2. An optimization problem  We now use Lemma 4.1 to study the probability that a subgraph is present on vertices of specific degrees. Assume that $D_i \in [\varepsilon, 1/\varepsilon] n^{\alpha_i}$ with $\alpha_i \in [0, 1/(\tau - 1)]$ for all $i$, so that $D_i = \Theta(n^{\alpha_i})$. When $\alpha_i + \alpha_j < 1$, by (4.4)

$$\Pr_n\left( X_{ij} = 1 \right) = \left( 1 - e^{-\Theta(n^{\alpha_i+\alpha_j-1})} \right) (1 + o(1)) = \Theta\left( n^{\alpha_i+\alpha_j-1} \right).$$  \hfill (4.16)

Furthermore, by (4.4), $\Pr_n\left( X_{ij} = 1 \right) = \Theta(1)$ and $\Pr_n\left( X_{ij} = 0 \right) = \Theta(1)$ when $\alpha_i + \alpha_j = 1$. When $\alpha_i + \alpha_j > 1$ instead, by (4.5)

$$\Pr_n\left( X_{ij} = 1 \right) = 1 - O(e^{-n^{\alpha_i+\alpha_j-1}/(4\mu)}),$$  \hfill (4.17)

so that it equals 1 minus a stretched exponentially small term. For vertices $u, v$, denote

$$w_{uv} = n^{\alpha_u+\alpha_v-1-\gamma},$$  \hfill (4.18)
with $\gamma$ as in (4.3). By Lemma 4.1, for any set of $m$ edges,

$$
\mathbb{P}_n (X_{u_1,v_1} = \cdots = X_{u_m,v_m} = 0) = \prod_{\alpha_{u_i} + \alpha_{v_i} < 1} (1 + O(w_{u_i,v_i}))(1 - \Theta(n^{\alpha_{u_i} + \alpha_{v_i} - 1}))
\times \prod_{\alpha_{u_i} + \alpha_{v_i} = 1} e^{-D_{u_i}D_{v_i}/(\mu n)}(1 + O(n^{-(\tau - 2)/(\tau - 1)}))
\times \prod_{\alpha_{u_i} + \alpha_{v_i} > 1} \Theta(e^{-n^{(\alpha_{v_i} + \alpha_{u_i} - 1)/(4\mu)}}).
$$

(4.19)

For ease of notation, we denote

$$
q(i,j) = \begin{cases}
(1 + O(w_{ij}))(1 - \Theta(n^{\alpha_i + \alpha_j - 1})) & \text{if } \alpha_i + \alpha_j < 1, \\
e^{-D_{i}D_{j}/(\mu n)}(1 + O(n^{-\gamma})) & \text{if } \alpha_i + \alpha_j = 1, \\
\Theta(e^{-n^{(\alpha_i + \alpha_j - 1)/(4\mu)}}) & \text{if } \alpha_i + \alpha_j > 1.
\end{cases}
$$

(4.20)

Let $H$ be a subgraph on $k$ vertices labeled as $1, \ldots, k$ with edges $E_H = \{u_1, v_1\}, \ldots, \{u_m, v_m\}$. Furthermore, let $\text{ECM}^{(n)}(D)|_i$ be the induced subgraph of $\text{ECM}^{(n)}(D)$ on vertices $i = (i_1, \ldots, i_k)$. We write the probability that $H$ is present on a specified subset of vertices $i = (i_1, \ldots, i_k)$ as

$$
\mathbb{P}_n (\text{ECM}^{(n)}(D)|_i \supseteq E_H)
= 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}} = X_{i_{u_{u_l},i_{v_{u_l}}}} = 0) + \cdots \\
+ (-1)^m \mathbb{P}_n (X_{i_{u_1},i_{v_1}} = \cdots = X_{i_{u_m},i_{v_m}} = 0)

= 1 - \sum_{l=1}^{m} q(i_{u_l},i_{v_l}) + \sum_{l \neq j} q(i_{u_l},i_{v_l})q(i_{u_j},i_{v_j})
\cdots + (-1)^m \prod_{l \in [m]} q(i_{u_l},i_{v_l})
$$

(4.21)

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

$$
1 - q(i,j) = 1 - (1 - \Theta(n^{\alpha_i + \alpha_j - 1}))(1 + O(w_{ij})) = \Theta(n^{\alpha_i + \alpha_j - 1}),
$$

(4.22)
and that for $\alpha_i + \alpha_j > 1$

\begin{equation}
1 - q(i, j) = 1 - O\left(e^{-n^{(\alpha_i + \alpha_j - 1)/(4\mu)}}\right) = 1 + o(1).
\end{equation}

Furthermore, for $D_i \in [\varepsilon, 1/\varepsilon] n^{\alpha_i}$, $D_j \in [\varepsilon, 1/\varepsilon] n^{\alpha_j}$ and $\alpha_i + \alpha_j = 1$,

\begin{equation}
1 - q(i, j) = (1 + O(n^{-\gamma}))(1 - e^{-D_i D_j / (\mu n)}) = \Theta(1),
\end{equation}

so that edges with $\alpha_i + \alpha_j \geq 1$ do not contribute to the order of magnitude of the last term in (4.21). The degrees are an i.i.d. sample from a power-law distribution. Therefore,\textbf{Lemma 4.2 (Maximum contribution to subgraphs).} Let $H$ be a connected graph on $k$ vertices. If the solution to (4.28) 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 $H$.\textbf{Proof.}
Proof. Defining $\beta_i = \alpha_i - \frac{1}{2}$ yields for (4.28)

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

over all possible values of $\beta_i \in \left[ -\frac{1}{2}, \frac{3 - \tau}{2(\tau - 1)} \right]$. Then, we have to prove that $\beta_i \in \left\{ -\frac{1}{2}, \frac{\tau - 3}{2(\tau - 1)}, 0, \frac{3 - \tau}{2(\tau - 1)} \right\}$ for all $i$ in the optimal solution. Note that (4.29) is a piecewise linear function in $\beta_1, \ldots, \beta_k$. Therefore, if (4.29) has a unique maximum, it must be attained at the boundary for $\beta_i$ or at a border of one of the linear sections. Thus, any unique optimal value of $\beta_i$ 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 (4.29), since it does not influence the optimal $\beta$ values. Rewriting (4.29) without the constant factor yields

$$\max \sum_i \beta_i \left( 1 - \tau + |\{s \in [k] : (s, i) \in E_H \text{ and } \beta_s < -\beta_i\}| \right).$$

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 contain $i$ with $|\beta_i| \in \left( 0, \frac{3 - \tau}{2(\tau - 1)} \right)$.

**Step 3.** Show that any optimal solution that is not unique can be transformed into two different optimal solutions with $\beta_i \in \left\{ -\frac{1}{2}, \frac{\tau - 3}{2(\tau - 1)}, 0, \frac{3 - \tau}{2(\tau - 1)} \right\}$ 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 (4.30) is

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

For any value of $\beta_j \in \left[ -\frac{1}{2}, \frac{3 - \tau}{2(\tau - 1)} \right]$, this contribution is maximized when choosing $\beta_i = -1/2$. Thus, $\beta_i = -1/2$ 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 $\beta_i < \frac{3 - \tau}{2(\tau - 1)}$. Because the maximal value of $\beta_j$ for $j \neq i$ is $\frac{3 - \tau}{2(\tau - 1)}$, the contribution to the $i$th term of (4.30) is

$$-\frac{1}{2} (1 - \tau + d_i) < 0,$$

irrespective of the values of the $\beta_j, j \neq i$. Increasing $\beta_i$ to $\frac{3 - \tau}{2(\tau - 1)}$ then gives a higher contribution. Thus, $\beta_i \geq \frac{3 - \tau}{2(\tau - 1)}$ when $d_i \geq 2$. 

20 R. VAN DER HOFSTAD, J. VAN LEEUWAARDEN AND C. STEGEHUIS
Step 2. Now we show that when the solution to (4.30) is unique, it is never optimal to have $|\beta| \in (0, \frac{3-\tau}{2(\tau-1)})$. Let

\begin{equation}
\tilde{\beta} = \min_{i:|\beta_i|>0} |\beta_i|.
\end{equation}

Let $N_{\tilde{\beta}^-}$ denote the number of vertices with their $\beta$ value equal to $-\tilde{\beta}$, and $N_{\tilde{\beta}^+}$ the number of 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 (4.30) is

\begin{equation}
\tilde{\beta}\big((1-\tau)(N_{\tilde{\beta}^+} - N_{\tilde{\beta}^-}) + E_{\tilde{\beta}^+} - E_{\tilde{\beta}^-}\big).
\end{equation}

Because we assume $\beta$ to be optimal, and the optimum to be unique, the value inside the brackets cannot equal zero. The contribution is linear in $\tilde{\beta}$ and it is the optimal contribution, and therefore $\tilde{\beta} \in \{0, \frac{3-\tau}{2(\tau-1)}\}$. This shows that $\beta_i \in \{\tau - \frac{3}{2(\tau-1)}, 0, \frac{3-\tau}{2(\tau-1)}\}$ for all $i$ such that $d_i \geq 2$.

Step 3. Suppose the solution to (4.30) is not unique. Suppose $\beta_\ast$ appears in one of the optimizers of (4.30). In the same notation as in (4.34), the contribution from vertices with $\beta$-values $\beta_\ast$ and $-\beta_\ast$ equals

\begin{equation}
\beta_\ast\big[(1-\tau)(N_{\beta_\ast^+} - N_{\beta_\ast^-}) + E_{\beta_\ast^+} - E_{\beta_\ast^-}\big].
\end{equation}

Since this contribution is linear in $\beta_\ast$, the contribution of these vertices can only be non-unique if the term within the square brackets equals zero. Thus, for the solution to (4.30) to be non-unique, there must exist $\hat{\beta}_1, \ldots, \hat{\beta}_s > 0$ for some $s \geq 1$ such that

\begin{equation}
\hat{\beta}_j\big((1-\tau)(N_{\beta_j^+} - N_{\beta_j^-}) + E_{\beta_j^+} - E_{\beta_j^-}\big) = 0 \quad \forall j \in [s].
\end{equation}

Setting all $\hat{\beta}_j = 0$ and setting all $\hat{\beta}_j = \frac{3-\tau}{2(\tau-1)}$ are both optimal solutions. Thus, if the solution to (4.30) 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 $i$. \hfill $\Box$

Proof of Theorem 2.1(ii) for subgraphs. Let $\alpha_\text{sub}$ be the unique optimizer of (4.28). By Lemma 4.2, the maximal value of (4.28) 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 $\alpha_\text{sub}^{(1)} = \frac{\tau-2}{\tau-1}$, vertices in $S_2$ have $\alpha_\text{sub}^{(2)} = \frac{1}{\tau-1}$, vertices in $S_3$ have $\alpha_\text{sub}^{(3)} = \frac{1}{2}$.
and vertices in \( V_1 \) have \( \alpha_i^{(\text{sub})} = 0 \). Then, the edges with \( \alpha_i^{(\text{sub})} + \alpha_j^{(\text{sub})} < 1 \) are edges inside \( S_1 \), edges between \( S_1 \) and \( S_3 \) and edges from degree 1 vertices. 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 \( E_{S_i,V_1} \). Then we can rewrite (4.28) as

\[
\max_P \left[ (1 - \tau) \left( \frac{\tau - 2}{\tau - 1} |S_1| + \frac{1}{\tau - 1} |S_2| + \frac{1}{2} |S_3| \right) + \frac{\tau - 3}{\tau - 1} E_{S_1} 
+ \frac{\tau - 3}{2(\tau - 1)} E_{S_1,S_3} - \frac{E_{S_1,V_1}}{\tau - 1} - \frac{\tau - 2}{\tau - 1} E_{S_2,V_1} - \frac{1}{2} E_{S_3,V_1} \right],
\]

over all partitions \( P \) 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 \((3 - \tau)/2 \) shows that this is equivalent to

\[
\frac{1 - \tau}{2} k + \max_P \left( \frac{3 - \tau}{2} \left( |S_1| - |S_2| + \frac{\tau - 2}{3 - \tau} k_1 - \frac{2 E_{S_1} + E_{S_1,S_3}}{\tau - 1} \right) - \frac{E_{S_1,V_1}}{\tau - 1} \right),
\]

since \( k \) and \( k_1 \) are fixed and \( 3 - \tau > 0 \), we need to maximize

\[
B^{(\text{sub})}(H) = \max_P \left[ |S_1| - |S_2| - \frac{2 E_{S_1} + E_{S_1,S_3} + E_{S_1,V_1} - E_{S_2,V_1}}{\tau - 1} \right].
\]

By (4.27), the contribution of the maximum is then given by

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

which proves Theorem 2.1(ii) for subgraphs. \( \square \)

5. Proof of Theorem 2.2  
Define the special case of \( M_n^{(\alpha)}(\varepsilon) \) of (2.4) where \( \alpha_i = 1/2 \ \forall i \) as

\[
W_n^k(\varepsilon) = \{(i_1, \ldots, i_k) : D_{i_s} \in [\varepsilon, 1/\varepsilon]\sqrt{\mu n} \ \forall s \in [k]\},
\]

and let \( \bar{W}_n^k(\varepsilon) \) denote the complement of \( W_n^k(\varepsilon) \). Denote the number of subgraphs \( H \) with all vertices in \( W_n^k(\varepsilon) \) by \( N^{(\text{sub})}(H, W_n^k(\varepsilon)) \).

Lemma 5.1 (Major contribution to subgraphs). Let \( H \) be a connected graph on \( k > 2 \) vertices such that (2.2) is uniquely optimized at 0. Then,
(i) The number of subgraphs with vertices in $W^k_n(\varepsilon)$ satisfies
\[
\frac{N^{(\text{sub})}(H, W^k_n(\varepsilon))}{n^{k(3-\tau)}} = (1 + o_1(1))c^k \mu^{-\frac{k}{2}(\tau-1)} \int_\varepsilon^{1/\varepsilon} \cdots \int_\varepsilon^{1/\varepsilon} (x_1 \cdots x_k)^{-\tau} \times \prod_{(u,v) \in E_H} (1 - e^{-x_u x_v}) dx_1 \cdots dx_k + f_n(\varepsilon),
\]
for some function $f_n(\varepsilon)$ such that for any $\delta > 0$,
\[
\lim_{\varepsilon \to 0} \lim_{n \to \infty} \sup \mathbb{P}(f_n(\varepsilon) > \delta | J_n) = 0.
\]

(ii) $A^{(\text{sub})}(H)$ defined in (2.10) satisfies $A^{(\text{sub})}(H) < \infty$.

The proof of Lemma 5.1 can be found in Section 6. We now prove Theorem 2.2 using this lemma.

**Proof of Theorem 2.2.** We first study the expected number of subgraphs with vertices outside $W^k_n(\varepsilon)$. First, we investigate the expected number of subgraphs in the case where vertex 1 of the subgraph has degree smaller than $\varepsilon \sqrt{\mu n}$. Similarly to (4.21), we can use Lemma 4.1 to show that the probability that $H$ is present on a specified subset of vertices $i = (i_1, \ldots, i_k)$ can be written as
\[
\mathbb{P}_n(\text{ECM}^{(n)}(D) | i \supseteq E_H) = \Theta \left( \prod_{\{i,j\} \in E_H : D_i, D_j < L_n} (1 - e^{-D_i D_j / L_n}) \right)
\]

Furthermore, by (1.1), there exists $C_0$ such that $\mathbb{P}(D = k) \leq C_0 k^{-\tau}$ for all $k$. Thus, the expected number of subgraphs in the case where vertex 1 of the subgraph has degree smaller than $\varepsilon \sqrt{\mu n}$ is bounded by
\[
\mathbb{E} \left[ N^{(\text{sub})}(H) \mathbb{1}_{\{D_1 < \varepsilon \sqrt{\mu n}\}} | J_n \right] \leq \Theta(1)n^k \int_1^{\varepsilon \sqrt{\mu n}} \int_1^{\infty} \cdots \int_1^{\infty} (x_1 \cdots x_k)^{-\tau} \prod_{\{i,j\} \in E_H} (1 - e^{-x_i x_j / (\mu n)}) dx_1 \cdots dx_k
\]
\[
= \Theta(1)n^k(\mu n)^{\frac{k}{2}(1-\tau)} \int_0^{\varepsilon} \int_0^{\infty} \cdots \int_0^{\infty} (t_1 \cdots t_k)^{-\tau} \prod_{\{i,j\} \in E_H} (1 - e^{-t_i t_j}) dt_1 \cdots dt_k
\]
\[
= O \left( n^{k(3-\tau)} \right) h_1(\varepsilon),
\]
where $h_1(\varepsilon)$ is a function of $\varepsilon$. By Lemma 5.1(ii), $h_1(\varepsilon) \to 0$ as $\varepsilon \downarrow 0$. 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 $\sqrt{n}/\varepsilon$, similarly. This yields

$$E \left[ N^{\text{sub}}(H, \tilde{W}_n^k(\varepsilon)) \mid J_n \right] = O \left( n^{k(3-\tau)} \right) h(\varepsilon),$$

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

$$N^{\text{sub}}(H, \tilde{W}_n^k(\varepsilon)) = h(\varepsilon) O_P \left( n^{k(3-\tau)/2} \right).$$

Therefore, for any $\delta > 0$,

$$\limsup \limsup_{\varepsilon \to 0} n \to \infty \mathbb{P} \left( \frac{N^{\text{sub}}(H, \tilde{W}_n^k(\varepsilon))}{n^{k(3-\tau)/2}} > \delta \mid J_n \right) = 0.$$

Combining this with the fact that $\mathbb{P}(J_n) \to 1$ and Lemma 5.1(i) gives

$$\frac{N^{\text{sub}}(H)}{n^{k(3-\tau)}} \xrightarrow{p} c k^{\mu - \frac{k}{\gamma} \left( \tau - 1 \right)} \int_0^\infty \cdots \int_0^\infty (x_1, \ldots, x_k)^{-\tau} \prod_{(u,v) \in E_H} (1 - e^{-x_u x_v}) dx_1 \cdots dx_k.$$

\[\square\]

6. Major contribution to subgraphs: proof of Lemma 5.1 We first prove Lemma 5.1(i). We compute the expected value of the number of subgraphs in the quenched sense in Lemmas 6.1 and 6.2. Then, we study the variance of the number of subgraphs in the quenched sense in Lemma 6.3. Together, these lemmas prove Lemma 5.1(i).

6.1. Conditional expectation In this section, we study the expected number of subgraphs in ECM$^{(n)}(D)$. Let $H$ be a subgraph on $k$ vertices, labeled as $1, \ldots, k$, and $m$ edges, denoted by $e_1 = \{u_1, v_1\}, \ldots, e_m = \{u_m, v_m\}$.

**Lemma 6.1 (Conditional expectation of subgraphs).** Let $H$ be a subgraph such that (2.2) has a unique maximum, attained at 0. Then, on the event $J_n$ defined in (3.1)

$$E_n \left[ N^{\text{sub}}(H, W_n^k(\varepsilon)) \right] = \sum_{(i_1, \ldots, i_k) \in W_n^k(\varepsilon)} \prod_{(u,v) \in E_H} (1 - e^{-D_{i_u} D_{i_v} / L_n}) (1 + o(1)).$$
Proof. Let $i = (i_1, \ldots, i_k)$ and ECM$_n^{(n)}(D)|_i$ again be the induced subgraph of ECM$_n^{(n)}(D)$ on $i$. We first derive a more detailed expression for the probability that a subgraph is present on $i$ than (4.21) which holds when $i \in W_n^k(\varepsilon)$. Because $i \in W_n^k(\varepsilon)$, we may use (4.3) for all edge probabilities to obtain

$$
\mathbb{P}_n(X_{i_{u_1,i_{v_1}}} = \cdots = X_{i_{u_m,i_{v_m}}}) = \prod_{l=1}^m \mathbb{P}_n(X_{i_{u_l,i_{v_l}}} = 0)(1+O(n^{(\tau-2)/(\tau-1)})�).
$$

When $D_i, D_j \in [\varepsilon \sqrt{n}, \sqrt{n}/\varepsilon]$, $\mathbb{P}_n(X_{ij}) = \Theta(1)$ and $\mathbb{P}_n(X_{ij} = 1) = \Theta(1)$. Therefore, similarly to (4.21), for $i \in W_n^k(\varepsilon)$,

$$
\mathbb{P}_n \left( \text{ECM}_n^{(n)}(D)|_i \supseteq E_H \right)
= 1 - \sum_{l=1}^m \mathbb{P}_n(X_{i_{u_l,i_{v_l}}} = 0) + \sum_{l \neq j} \mathbb{P}_n(X_{i_{u_l,i_{v_l}}} = X_{i_{u_j,i_{v_j}}} = 0) - \sum_{l \neq j \neq w} \mathbb{P}_n(X_{i_{u_l,i_{v_l}}} = X_{i_{u_j,i_{v_j}}} = X_{i_{u_w,i_{v_w}}}) + \cdots
+ (-1)^m \mathbb{P}_n(X_{i_{u_1,i_{v_1}}} = \cdots = X_{i_{u_m,i_{v_m}}}) = (1 + o(1)) \prod_{l=1}^m \left( 1 - \mathbb{P}_n \left( X_{i_{u_l,i_{v_l}}} = 0 \right) \right).
$$

Thus, the expected value satisfies

$$
\mathbb{E}_n \left[ N^{(\text{sub})}(H, W_n^k(\varepsilon)) \right] = \sum_{i \in W_n^k(\varepsilon)} \mathbb{P}_n \left( \text{ECM}_n^{(n)}(D)|_i \supseteq E_H \right)
= (1 + o(1)) \sum_{i \in W_n^k(\varepsilon)} \prod_{l=1}^m \left( 1 - \mathbb{P}_n \left( X_{i_{u_l,i_{v_l}}} = 0 \right) \right),
$$

Because $D_iD_j = O(n)$ and $L_n = \mu n(1 + o(1))$ under $J_n$, by (4.4)

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

This results in

$$
\mathbb{E}_n \left[ N^{(\text{sub})}(H, W_n^k(\varepsilon)) \right] = (1 + o(1)) \sum_{i \in W_n^k(\varepsilon)} \prod_{(u,v) \in E_H} \left( 1 - e^{-D_uD_v/L_n} \right).
$$

$\square$
6.2. Convergence of conditional expectation  We now study the asymptotic behavior of the expected number of subgraphs using Lemma 6.1.

**Lemma 6.2** (Convergence of conditional expectation of \(\sqrt{n}\) subgraphs). Let \(H\) be a subgraph such that (2.2) has a unique maximizer, and the maximum is attained at 0. Then,

\[
\mathbb{E}_n \left[ N_{\text{sub}}(H, W_n^k(\varepsilon)) \right] = (1 + o_p(1)) c^k \mu^{k-\frac{3}{2}(\tau-1)} \int_{\varepsilon}^{1/\varepsilon} \cdots \int_{\varepsilon}^{1/\varepsilon} (x_1 \cdots x_k)^{-\tau} \\
\times \prod_{(u,v) \in E_H} (1 - e^{-x_u x_v}) \, dx_1 \cdots dx_k + f_n(\varepsilon),
\]

for some function \(f_n(\varepsilon)\) such that for any \(\delta > 0\),

\[
\lim_{\varepsilon \downarrow 0} \lim_{n \to \infty} \mathbb{P} \left( f_n(\varepsilon) > \delta \mid J_n \right) = 0.
\]

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

Define

\[
g(t_1, \ldots, t_k) := \prod_{(u,v) \in E_H} (1 - e^{-t_u t_v}).
\]

Using the Taylor expansion of \(1 - e^{-xy}\) on \([\varepsilon, 1/\varepsilon]^2\) results in

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

Since \(g\) is bounded on \(F = [\varepsilon, 1/\varepsilon]^k\), we can find \(s_1, \ldots, s_m\) and \(\eta(t_1, \ldots, t_k) \leq \varepsilon^{k(\tau-1) + 1}\) such that

\[
g(t_1, \ldots, t_k) = \sum_{i_1=1}^{s_1} \cdots \sum_{i_m=1}^{s_m} \left( (-1)^{i_1} \frac{t_{i_1}^1}{i_1!} \cdots (-1)^{i_m} \frac{t_{i_m}^m}{i_m!} \right) + \eta(t_1, \ldots, t_k)
\]

\[
= \sum_{i_1=1}^{s_1} \cdots \sum_{i_m=1}^{s_m} \left( (-1)^{i_1+\cdots+i_m} \frac{t_{i_1}^1}{i_1!} \cdots \frac{t_{i_m}^m}{i_m!} \gamma_{i_1} \cdots \gamma_{i_m} \right) + \eta(t_1, \ldots, t_k),
\]

where

\[
\gamma_j := \gamma_j(i_1, \ldots, i_m) = \sum_l i_l \mathbb{1}_{\{u_l = j \text{ or } v_l = j\}}.
\]
Let $M^{(n)}$ denote the random measure

$$M^{(n)}([a, b]) = (\mu n)^{1/2(n-1)} \sum_{i=1}^{n} \frac{1}{2^n} \{ D_i \in \sqrt{\mu n} [a, b] \}. \tag{6.13}$$

The number of vertices with degrees in a certain interval $[a, b]$ is binomially distributed. By (1.1), we thus get $(\mu n)^{1/2(n-1)} \Pr (D_1 \in \sqrt{\mu n} [a, b]) \xrightarrow{p} \lambda([a, b])$, where

$$\lambda([a, b]) := \int_a^b x^{-\tau} \, dx. \tag{6.14}$$

Hence, by the weak law of large numbers, as $n \to \infty$,

$$M^{(n)}([a, b]) \xrightarrow{p} \lambda([a, b]). \tag{6.15}$$

Let $N^{(n)}$ denote the product measure $M^{(n)} \times M^{(n)} \times \cdots \times M^{(n)}$ ($k$ times). Then (6.11) together with Lemma 6.1 yields

$$E_n \left[ N^{(n)}(H, W_k^\varepsilon) \right] = \int_F g(t_1, \ldots, t_k) dN^{(n)}(t_1, \ldots, t_k)$$

$$= \int_F \sum_{i_1=1}^{s_1} \cdots \sum_{i_m=1}^{s_m} \left( \left( \frac{-1}{i_1! \cdots i_m!} t_1^{\gamma_1} t_2^{\gamma_2} \cdots t_k^{\gamma_k} + \eta(t_1, \ldots, t_k) \right) dN^{(n)}(t_1, \ldots, t_k) \right)$$

$$= \sum_{i_1=1}^{s_1} \cdots \sum_{i_m=1}^{s_m} \left( \frac{-1}{i_1! \cdots i_m!} \int_\varepsilon^{1/\varepsilon} t_1^{\gamma_1} dM^{(n)}(t_1) \cdots \int_\varepsilon^{1/\varepsilon} t_k^{\gamma_k} dM^{(n)}(t_k) \right) + f_n(\varepsilon). \tag{6.16}$$

Here

$$f_n(\varepsilon) = \int_F \sum_{i_1=1}^{s_1} \cdots \sum_{i_m=1}^{s_m} \eta(t_1, \ldots, t_k) dN^{(n)}(t_1, \ldots, t_k)$$

$$\leq \int_F \sum_{i_1=1}^{s_1} \cdots \sum_{i_m=1}^{s_m} \varepsilon^{k(\tau-1)+1} dN^{(n)}(t_1, \ldots, t_k)$$

$$= \varepsilon^{k(\tau-1)+1} s_1 \cdots s_m M^{(n)}([\varepsilon, 1/\varepsilon])^k$$

$$= \varepsilon^{k(\tau-1)+1} O_p \left( \lambda([\varepsilon, 1/\varepsilon])^k \right) = \varepsilon^{k(\tau-1)+1} (\varepsilon^{1-\tau} - \varepsilon^{\tau-1})^k O_p (1), \tag{6.17}$$

which shows that for any $\delta > 0$,

$$\lim_{\varepsilon \downarrow 0} \limsup_{n \to \infty} \Pr (f_n(\varepsilon) > \delta | J_n) = 0. \tag{6.18}$$
As in [24, Eq. (55)] for any \( \gamma \)

\[
\int_{\varepsilon}^{1/\varepsilon} x^{\gamma} dM^{(n)}(x) \xrightarrow{p} \int_{\varepsilon}^{1/\varepsilon} x^{\gamma} d\lambda(x).
\]

Combining this with (6.16) results in

\[
E_{n} \left[ N^{(sub)}(H, W_{n}^{k}(\varepsilon)) \right]_{\gamma} = \Theta_{\varepsilon} n^{(3-\tau)k},
\]

Then, by (6.15)

\[
\frac{E_{n} \left[ N^{(sub)}(H, W_{n}^{k}(\varepsilon)) \right]}{n^{\frac{k}{2}(3-\tau)\mu^{2}(1-\tau)}} = (1 + o_{\varepsilon}(1)) \sum_{i_{1}=1}^{s_{1}} \cdots \sum_{i_{m}=1}^{s_{m}} \frac{(-1)^{i_{1}+\cdots+i_{m}}}{i_{1}! \cdots i_{m}!} \int_{\varepsilon}^{1/\varepsilon} t_{1}^{\alpha_{i_{1}}} d\lambda(t_{1}) \cdots \int_{\varepsilon}^{1/\varepsilon} t_{k}^{\alpha_{i_{m}}} d\lambda(t_{k}) + f_{n}(\varepsilon)
\]

\[
\int_{\varepsilon}^{1/\varepsilon} \cdots \int_{\varepsilon}^{1/\varepsilon} g(t_{1}, \cdots, t_{k}) dt_{1} \cdots dt_{k} + f_{n}(\varepsilon).
\]

which proves the claim. 

\[\square\]

6.3. Conditional variance

We now study the conditional variance of the number of subgraphs in the quenched setting for the degrees. The following lemma shows that the conditional variance of the number of subgraphs is small compared to its expectation.

**Lemma 6.3 (Conditional variance for subgraphs).** Let \( H \) be a subgraph such that (2.2) has a unique maximum attained at 0. Then, on the event \( J_{n} \) as defined in (3.1)

\[
\frac{\text{Var}_{n} \left( N^{(sub)}(H, W_{n}^{k}(\varepsilon)) \right)}{E_{n} \left[ N^{(sub)}(H, W_{n}^{k}(\varepsilon)) \right]^{2}} \xrightarrow{p} 0.
\]

**Proof.** By Lemma 6.2,

\[
E_{n} \left[ N^{(sub)}(H, W_{n}^{k}(\varepsilon)) \right]^{2} = \Theta_{\varepsilon} (n^{(3-\tau)k}),
\]
Thus, we need to prove that the variance is small compared to \(n^{(3-\tau)k}\). Denote \(i = (i_1, \ldots, i_k)\) and \(j = (j_1, \ldots, j_k)\) and for ease of notation we denote \(G = ECM^{(n)}(D)\). We write the variance as

\[
\text{Var}_n \left( N^{(\text{sub})}(H, W_n^k(\varepsilon)) \right) = \sum_{i \in W_n^k(\varepsilon)} \sum_{j \in W_n^k(\varepsilon)} \left( P_n(G|i \supseteq E_H, G|j \supseteq E_H) - P_n(G|i \supseteq E_H) P_n(G|j \supseteq E_H) \right).
\]

(6.24)

This splits into various cases, depending on the overlap of \(i\) and \(j\). When \(i\) and \(j\) do not overlap, similarly to (6.3),

\[
\sum_{i \in W_n^k(\varepsilon)} \sum_{j \in W_n^k(\varepsilon)} \left( (1 + o(1)) \prod_{l=1}^m (1 - P_n(X_{i_{ul}, i_{vl}} = 0)) (1 - P_n(X_{j_{ul}, j_{vl}} = 0)) \right)
\]

\[- (1 + o(1)) \prod_{l=1}^m (1 - P_n(X_{i_{ul}, i_{vl}} = 0)) (1 - P_n(X_{j_{ul}, j_{vl}} = 0)) \right)
\]

\[= E_n \left[ N^{(\text{sub})}(H, W_n^k(\varepsilon)) \right]^2 o(1).
\]

The other contributions are when \(i\) and \(j\) overlap. In this situation, we use the bound \(P_n(X_{ij} = 1) \leq 1\). When \(i\) and \(j\) overlap on \(s \geq 1\) vertices, we bound the contribution to (6.24) as

\[
\sum_{i,j \in W_n^k(\varepsilon) : |i \cup j| = 2k-s} P_n(G|i \supseteq E_H, G|j \supseteq E_H) \leq |\{i : D_i \in \sqrt{mn}[\varepsilon, 1/\varepsilon]\}|^{2k-s}
\]

(6.26)

\[= O_\varepsilon \left( n^{(3-\tau)(2k-s)} \right),
\]

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

**Proof of Lemma 5.1.** We start by proving part (i). By Lemma 6.3 and Chebyshev’s inequality, conditionally on the degrees

\[
N^{(\text{sub})}(H, W_n^k(\varepsilon)) = E_n \left[ N^{(\text{sub})}(H, W_n^k(\varepsilon)) \right] (1 + o_\varepsilon(1)).
\]

(6.27)

Combining this with Lemma 6.2 proves Lemma 5.1(i). Lemma 5.1(ii) is a direct consequence of Lemma 7.2 where \(|S_{3i}^*| = k\). \(\square\)
7. **Major contribution to general subgraphs** In this section we prove Theorem 2.1(i) for subgraphs. We start by introducing some notation. For any \( W \subseteq V_H \), we denote by \( d_{i,W} \) the number of edges from vertex \( i \) to vertices in \( W \). Let \( H \) be a connected subgraph, such that the optimum of (2.2) is unique, and let \( S_1^*, S_2^* \) and \( S_3^* \) be the optimal partition. Define

\[
\zeta_i = \begin{cases} 
1 & \text{if } d_i = 1, \\
 d_i, & \text{if } i \in S_1^*, \\
 d_i, & \text{if } i \in S_2^*, \\
 d_i, & \text{if } i \in S_3^*
\end{cases}
\]  

(7.1)

The following lemma states several properties of the number of edges between vertices in the different optimizing sets:

**Lemma 7.1.** Let \( H \) be a connected subgraph, such that the optimum of (2.2) is unique, and let \( S_1^*, S_2^* \) and \( S_3^* \) be the optimal partition. Then the following holds:

(i) \( \zeta_i \leq 1 \) for \( i \in S_1^* \).
(ii) \( d_i, S_1^* + \zeta_i \geq 2 \) for \( i \in S_2^* \).
(iii) \( \zeta_i \leq 1 \) and \( d_i, S_3^* + \zeta_i \geq 2 \) for \( i \in S_3^* \).

**Proof.** Suppose \( i \in S_1^* \). Now consider the partition \( \hat{S}_1 = S_1^* \setminus \{i\}, \hat{S}_2 = S_2^*, \hat{S}_3 = S_3^* \cup \{i\} \). Then, \( E_{\hat{S}_1} = E_{S_1^*} - d_i, S_1^* \) and \( E_{\hat{S}_1, \hat{S}_2} = E_{S_1^*, S_2^*} + d_i, S_1^* - d_i, S_3^* \). Furthermore, \( E_{\hat{S}_2, 1} = E_{S_2^*, 1} - d_i, 1 \) and \( E_{\hat{S}_2, 1} = E_{S_2^*, 1} \). Because the partition into \( S_1^*, S_2^* \) and \( S_3^* \) achieves the unique optimum of (2.2)

\[
|S_1^*| - |S_2^*| = \frac{2E_{S_1^*} - E_{S_1^*, S_3^*} + E_{S_2^*, 1} - E_{S_1^*, 1}}{\tau - 1} 
\]

(7.2)

\[
> |S_1^*| - 1 - |S_2^*| = \frac{2E_{S_1^*} - E_{S_1^*, S_3^*} - d_i, S_1^* - d_i, S_3^* + E_{S_2^*, 1} - E_{S_1^*, 1} + d_i, V_i}{\tau - 1},
\]

which reduces to

\[
d_i, S_3^* + d_i, S_1^* + d_i, V_i = \zeta_i < \tau - 1.
\]

(7.3)

Using that \( \tau \in (2, 3) \) then yields \( d_i, S_3^* + d_i, S_1^* + d_i, V_i \leq 1 \).

Similar arguments give the other inequalities. For example, for \( i \in S_3^* \), considering the partition where \( i \) is moved to \( S_1^* \) gives the inequality \( d_i, S_1^* + d_i, S_3^* + d_i, V_i \geq 2 \), and considering the partition where \( i \) is moved to \( S_2^* \) results in the inequality \( d_i, S_1^* + d_i, V_i \leq 1 \), so that \( \zeta_i \leq 1 \). \( \square \)
We now show that two integrals related to the solution of the optimization problem (2.2) are finite, using Lemma 7.1. These integrals are the key ingredient in proving Theorem 2.1(i) for subgraphs.

**Lemma 7.2.** Suppose that the maximum in (2.2) is uniquely attained with $|S_3^*| = s > 0$, and say $S_3^* = \{1, \ldots, s\}$. Then

\[\int_0^\infty \cdots \int_0^\infty \prod_{i \in [s]} x_i^{-\tau + \zeta_i} \prod_{(u,v) \in E_{S_3^*}} \min(x_u x_v, 1) dx_1 \cdots dx_s < \infty.\]  

**Proof.** The integral (7.4) consists of multiple regions. One region is $x_1, \ldots, x_s \geq 1$. Since $-\tau + \zeta_i < -1$ by Lemma 7.1, this integral satisfies

\[\int_1^\infty \cdots \int_1^\infty \prod_{j \in [s]} x_j^{-\tau + \zeta_j} dx_1 \cdots dx_s < \infty.\]  

Another region is $x_1, \ldots, x_s \in [0, 1]$. Since by Lemma 7.1, any vertex in $S_3^*$ has $\zeta_i + d_i, S_3^* \geq 2$, this integral can be bounded as

\[\int_0^1 \cdots \int_0^1 \prod_{j \in [s]} x_j^{-\tau + \zeta_j} \prod_{(u,v) \in E_{S_3^*}} x_u x_v dx_1 \cdots dx_s \leq \int_0^1 \cdots \int_0^1 (x_1 \cdots x_s)^2\tau dx_1 \cdots dx_s < \infty.\]  

The other regions can be described by the union of all sets $U \subset S_3^*$ such that the integral runs from 1 to $\infty$ for $i \in U$, and from 0 to 1 for $i \in U = S_3^* \setminus U$. In such a region, $\min(x_i, x_j, 1) = x_i x_j$ when $i, j \notin U$, and $\min(x_i, x_j) = 1$ when $i, j \in U$. W.l.o.g. assume $U = \{1, \ldots, t\}$ for some $t \geq 1$. Then, the contribution to (7.4) from the region described by $U$ can be written as

\[\int_1^\infty \cdots \int_1^\infty \prod_{j \in [t]} x_j^{-\tau + \zeta_j} \prod_{i=t+1}^s h(i, x) dx_i dx_1 \cdots dx_t,\]  

where $x = (x_i)_{i \in [t]}$ and

\[h(i, x) = \int_0^1 x_j^{-\tau + \zeta_j + d_j, U} \prod_{j \in U : \{i,j\} \in E_H} \min(x_i x_j, 1) dx_i,\]  

for $i \in \{t, t+1, \ldots, s\}$. The integral in $h(i, x)$ consists of multiple regions, depending on whether $x_i x_j < 1$ or not. Suppose vertex $j \in U$ is connected
to vertices $v_1, v_2, \ldots, v_l \in U$, where $1 < x_{v_1} < x_{v_2} < \cdots < x_{v_l}$. Then,

\begin{equation}
(7.9) \quad h(i, \mathbf{x}) = \int_0^1 x_i^{-\tau + \zeta_j + d_j, 0} \min(x_i x_{v_1}, 1) \min(x_i x_{v_2}, 1) \cdots \min(x_i x_{v_l}, 1) dx_i
\end{equation}

\[= x_{v_1} \cdots x_{v_l} \int_0^{1/x_{v_1}} x_i^{-\tau + \zeta_j + 1 + d_j, 0} dx_i + x_{v_2} \cdots x_{v_l-1} \int_0^{1/x_{v_2}} x_i^{-\tau + \zeta_j + 1 + d_j, 0} dx_i + \cdots + \int_0^{1/x_{v_l}} x_i^{-\tau + \zeta_j + d_j, 0} dx_i.\]

Since $\zeta_j + d_j, 0 + l - \tau = \zeta_j + d_j, 0 - \tau > -1$ by Lemma 7.1, the first integral is finite. Computing these integrals yields

\begin{equation}
(7.10) \quad h(i, \mathbf{x}) = C_1 x_{v_1} x_{v_2} \cdots x_{v_{l-1}} x_{v_l}^{-\tau - \zeta_j - l - d_j, 0} + C_2 x_{v_1} x_{v_2} \cdots x_{v_{l-2}} x_{v_{l-1}}^{-\tau - \zeta_j - l - d_j, 0} + \cdots + C_l x_{v_1}^{-\tau - \zeta_j - l - d_j, 0} + C_{l+1}
\end{equation}

for some constants $C_1, \ldots, C_{l+1}$. These terms (except for the last term) are all products of powers of $x_{v_1}, \ldots, x_{v_l}$, such that the sum of these powers is $\tau - \zeta_j - d_j, 0 - 1$. Furthermore, the exponents of $x_{v_1}, \ldots, x_{v_l}$ equal 1 for some $s \in [l]$, and the exponents of $x_{v_{l+1}}, \ldots, x_{v_l}$ equal zero. Therefore, for each $i$ there exists $i^* \in [l+1]$ such that, for all $1 < x_{v_1} < x_{v_2} < \cdots < x_{v_l},$

\begin{equation}
(7.11) \quad h(i, \mathbf{x}) \leq K h_{i^*}(i, \mathbf{x})
\end{equation}

for some $K > 0$. Then, for some $\tilde{K} > 0,$

\begin{equation}
(7.12) \quad \int_1^\infty \cdots \int_1^\infty \prod_{j \in [l]} x_j^{-\tau + \zeta_j} \prod_{i=t+1}^s h(i, \mathbf{x}) dx_t \cdots dx_1 \leq \tilde{K} \int_1^\infty \cdots \int_1^\infty \prod_{j \in [l]} x_j^{-\tau + \zeta_j} \prod_{i=t+1}^s h_{i^*}(i, \mathbf{x}) dx_t \cdots dx_1.
\end{equation}

We first show that the contribution to the upper bound in (7.12) from the area where $x_1 < x_2 < \cdots < x_t$ is finite. Let $T_i \subseteq U$ denote the set of neighbors of $i$ that appear in $h_{i^*}(i, \mathbf{x})$. Let

\begin{equation}
(7.13) \quad u_i = \max\{j \mid j \in T_i\}
\end{equation}

for all $i \in \bar{U}$ such that $|T_i| \geq 1$. Thus, $u_i$ denotes the index of the $x$-term that appears with exponent $\tau - \zeta_j - d_j, 0 - s$ in $h_{i^*}(i, \mathbf{x})$ for some $s$. Furthermore,
let
\begin{equation}
(7.14) \quad f(i) = \begin{cases} 
  x_{v_i} & \text{if } |T_i| \geq 1, \\
  1 & \text{else},
\end{cases}
\end{equation}
for all \( i \in \bar{U} \) and
\begin{equation}
(7.15) \quad Q_j = \{ i \in \bar{U} : u_i \geq j \}
\end{equation}
for all \( j \in U \). Thus, \( Q_j \) denotes the set of neighbors \( v \) of \( j \) in \( \bar{U} \) such that \( x_j \) appears in \( h^*_t(v, x) \). Then,
\begin{equation}
(7.16) \quad \int_1^{\infty} \cdots \int_{x_{t-1}}^{\infty} \prod_{j=1}^{t} x_j^{-\tau+\zeta_j} \prod_{i=t+1}^{s} \hat{h}_i(i, x) dx_t \cdots dx_1
\end{equation}
for some constant \( \hat{K} > 0 \). Let \( W_j = \{ i \in \bar{U} : u_i = j \} \) for \( j \in [t] \), so that \( W_j \) denotes the set of neighbors \( v \) of \( j \) in \( \bar{U} \) such that the term \( x_j^{\tau-\zeta_t-l-d_t, t} \) appears in \( h^*_t(v, x) \) for some \( l \). Furthermore, let \( \hat{W}_j = (V_1 \cup S^*_1 \cup [j]) \setminus \bar{U} \) \( W_j \).
Then, \( \sum_{i \in W_j} \zeta_i + d_i, t + d_i, t = \sum_{i \in W_j} d_i, v_1 + d_i, S^*_1 + d_i, t + d_i, t = 2E_{W_j} + E_{W_j, \hat{W}_j} \), where \( E_{W_j} \) denotes the number of edges inside \( W_j \) and \( E_{W_j, \hat{W}_j} \) denotes the number of edges between \( W_j \) and \( \hat{W}_j \). Thus, \( (7.16) \) results in
\begin{equation}
(7.17) \quad \hat{K} \int_1^{\infty} \cdots \int_{x_{t-1}}^{\infty} \prod_{j=1}^{t} x_j^{-\tau+\zeta_j+d_{i,Q_i}+(\tau-1)|W_i|} -2E_{W_i} - E_{W_i, \hat{W}_i} dx_t \cdots dx_1.
\end{equation}
We now show that
\begin{equation}
(7.18) \quad \tau + \zeta_t + d_t, Q_t + (\tau - 1)|W_t| - 2E_{W_t} - E_{W_t, \hat{W}_t} < -1,
\end{equation}
so that the integral in \( (7.17) \) over \( x_t \) is finite. Note that \( d_t, Q_t = d_t, W_t \) by definition of \( (7.13) \) and \( W_t \), and since \( t \) is the maximal index in \( U \). Also, \( \hat{W}_t = (V_1 \cup S^*_1 \cup S^*_3) \setminus W_t \). Setting \( \hat{S}_2 = \hat{S}_2 \cup \{ t \} \), \( \hat{S}_1 = \hat{S}_1 \cup W_t \) and \( \hat{S}_3 = S^*_3 \setminus (W_t \cup \{ t \}) \), gives
\begin{align}
(7.19) \quad E_{\hat{S}_1} - E_{\hat{S}_1}^* &= E_{W_t} + E_{W_t, S^*_1}, \\
(7.20) \quad E_{\hat{S}_1, \hat{S}_3} - E_{\hat{S}_1, \hat{S}_3}^* &= E_{W_t, S^*_3} - E_{W_t} - E_{W_t, S^*_1} - d_t, Q_t - d_t, S^*_1,
\end{align}
Because (2.2) is uniquely optimized for \( S_1^*, S_2^* \) and \( S_3^* \),
\[
|\hat{S}_1| - |\hat{S}_2| - \frac{2E_{\hat{S}_1} + E_{\hat{S}_1, S_3} + E_{\hat{S}_1, V_1} - E_{\hat{S}_2, V_1}}{\tau - 1} < |S_1^*| - |S_2^*| - \frac{2E_{S_1^*} + E_{S_1^*, S_3^*} + E_{S_1^*, V_1} - E_{S_2^*, V_1}}{\tau - 1}.
\] (7.23)

Using (7.19)-(7.22) this reduces to
\[
|W_t| - 1 - \frac{2E_{W_t} + E_{W_t, \hat{W}_t} - d_{t, Q_t} - d_{t, S_1^*} - d_{t, V_1}}{\tau - 1} < 0,
\] (7.24)
or
\[
-\tau + (\tau - 1) |W_t| + d_{t, Q_t} + d_{t, S_1^*} + d_{t, V_1} - 2E_{W_t} - E_{W_t, \hat{W}_t} < -1,
\] (7.25)
which is (7.18). Using that \( \zeta_t = d_{t, V_1} + d_{t, S_1^*} \) shows that the inner integral of (7.17) is finite, and that
\[
\int_{x_{t-1}}^\infty \prod_{j=t-1}^t \frac{1 - \tau + \zeta_j + d_{j, Q_j} + (\tau - 1) |W_j| - 2E_{W_j} - E_{W_j, \hat{W}_j} dx_t}{x_j}
\] (7.26)
for some \( K > 0 \), where \( W_t \cup W_{t-1} = V_1 \cup S_1^* \cup S_3^* \setminus (W_t \cup W_{t-1}) \). Then, choosing \( S_2 = S_2^* \setminus \{t-2, t-1\} \), \( S_1 = S_1^* \cup W_t \cup W_{t-1} \) and \( S_3 = S_3^* \setminus (W_t \cup W_{t-1} \cup \{t-1, t-2\}) \), we can again prove using (7.23) that integrating (7.26) over \( x_{t-1} \) from \( x_{t-2} \) to \( \infty \) as in (7.17) results in a finite function of \( x_{t-2} \). We continue this process until we arrive at the integral over \( x_1 \) and show that this final integral is finite. Therefore (7.17) and (7.16) are also finite. Since the area \( x_1 < x_2 < \cdots < x_t \) was arbitrary, the integral over any ordering of \( x_1, \ldots, x_t \) is finite, so that (7.12) is finite as well. ∎

**Lemma 7.3.** Suppose the optimal solution to (2.2) is unique, and attained by \( S_1^*, S_2^* \) and \( S_3^* \). Say that \( S_2^* = \{1, \ldots, t_2\} \) and \( S_1^* = \{t_2+1, \ldots, t_2+t_1\} \). Then,
\[
\int_0^1 \cdots \int_0^1 \int_0^\infty \cdots \int_0^\infty \prod_{j \in [t_1+t_2]} x_j^{-\tau + \zeta_j} \prod_{(u, v) \in E_{S_1^* S_2^*}} \min(x_u x_v, 1) dx_{t_1+t_2} \cdots dx_1 < \infty.
\] (7.27)
Proof. This proof has a similar structure as the proof of Lemma 7.2. We first rewrite the integral as

\[
\int_0^1 \cdots \int_0^1 \prod_{j \in [t_2]} x_j^{-\tau + \zeta_j} \prod_{i=t_2+1}^{t_1+t_2} \tilde{h}(i, x) \, dx_t \cdots dx_1,
\]

where \(x = (x_j)_{j \in [t_2]}\) and, for \(t_2 + 1 \leq i \leq t_1 + t_2\),

\[
\tilde{h}(i, x) = \int_0^\infty x_i^{-\tau + \zeta_i} \prod_{j \in [t_2]: \{i, j\} \in E_H} \min(x_i x_j, 1) \, dx_i.
\]

Similarly to (7.9), suppose that vertex \(i\) has vertices \(v_1, v_2, \ldots, v_l\) as neighbors in \(S^*_2\) where \(1 > x_{v_1} > x_{v_2} > \cdots > x_{v_l}\). Then,

\[
\tilde{h}(i, x) = \int_0^\infty x_i^{-\tau + \zeta_i} \min(x_i x_{v_1}, 1) \min(x_i x_{v_2}, 1) \cdots \min(x_i x_{v_l}, 1) \, dx_i
\]

\[
= \int_0^{1/x_{v_1}} x_i^{-\tau + \zeta_i + l} x_{v_1} \cdots x_{v_l} \, dx_i + \int_0^{1/x_{v_2}} x_i^{-\tau + \zeta_i + l-1} x_{v_2} \cdots x_{v_l} \, dx_i
\]

\[
+ \cdots + \int_0^\infty x_i^{-\tau + \zeta_i} \, dx_i.
\]

Because \(\zeta_i + l = d_i \geq 2\) and \(\zeta_i \leq 1\) by Lemma 7.1, the first and the last integrals are finite. Computing the integrals yields that for some \(C_1, \ldots, C_l\),

\[
\tilde{h}(i, x) = C_1 x_{v_1}^{-\zeta_i - l} x_{v_2} \cdots x_{v_l} + C_2 x_{v_2}^{-\zeta_i - l+1} x_{v_3} \cdots x_{v_l} + \cdots + C_l x_{v_l}^{-\zeta_i - 1}
\]

\[
(7.31) =: \tilde{h}_1(i, x) + \tilde{h}_2(i, x) + \cdots + \tilde{h}_l(i, x).
\]

Similarly to the argument leading to (7.11), for all \(i \in t_2 + 1, \ldots, t_1 + t_2\), there exists \(i^*\) such that for all \(x_{v_1} > x_{v_2} > \cdots > x_{v_l}\)

\[
(7.32) \tilde{h}(i, x) \leq K \tilde{h}_{i^*}(i, x)
\]

for some \(K > 0\). Thus,

\[
\int_0^1 \int_0^{x_1} \cdots \int_0^{x_{t_2-1}} \prod_{j \in [t_2]} x_j^{-\tau + \zeta_j} \prod_{i=t_2+1}^{t_1+t_2} \tilde{h}(i, x) \, dx_t \cdots dx_1
\]

\[
\leq K \int_0^1 \int_0^{x_1} \cdots \int_0^{x_{t_2-1}} \prod_{j \in [t_2]} x_j^{-\tau + \zeta_j} \prod_{i=t_2+1}^{t_1+t_2} \tilde{h}_{i^*}(i, x) \, dx_t \cdots dx_1.
\]

\[
(7.33)
\]

OPTIMAL SUBGRAPH STRUCTURES IN CONFIGURATION MODELS 35
Let $T_i$ denote the set of neighbors of $i$ whose terms appear in $h_{i^*}(t,\mathbf{x})$. By Lemma 7.1 $\zeta_i + l = d_i \geq 2$ and $\zeta_i \leq 1$ so that $l \leq 1$ and therefore $|T_i| \geq 1$ for all $i$. For $i \in S^*_1$, define

$$w_i = \min \{ j \in S^*_1 \mid j \in T_i \},$$

so that $w_i$ describes which term in $h_{i^*}(i, \mathbf{x})$ appears with exponent $\tau - \zeta_i - s$ for some $s$. Also, define

$$Q_j = \{ i \mid w_i \leq j \},$$

so that $Q_j$ is the set of indices $i$ such that such that $x_j$ appears in $h_{i^*}(i, \mathbf{x})$ when $i$ and $j$ are connected. Then,

$$\int_0^{x_1} \cdots \int_0^{x_{t_2}-1} \prod_{j \in [t_2]} \frac{x_j^{-\tau + \zeta_j}}{t_1 + t_2} \prod_{i=t_2+1}^t h_{i^*}(i, \mathbf{x}) dx t_2 \cdots dx_1$$

$$\leq \tilde{K} \int_0^{x_1} \cdots \int_0^{x_{t_2}-1} \prod_{j \in [t_2]} \frac{x_j^{-\tau + \zeta_j + d_j, Q_j}}{t_1 + t_2} \prod_{i=t_2+1}^t (1/x_{w_i})^{\tau - 1 - \zeta_i - d_i, T_i} dx t_2 \cdots dx_1,$$

for some $\tilde{K} > 0$. Define $W_j = \{ i \in S^*_1 \mid w_i = j \}$ for $j \in S^*_2$ and let $W_j = V_H \setminus (W_j \cup \{ j-1 \})$. Using that $\zeta_i = d_i, V_i + d_i, S^*_1 + d_i, S^*_2$ for $i \in S^*_1$, (7.36) reduces to

$$\tilde{K} \int_0^{x_1} \cdots \int_0^{x_{t_2}-1} \prod_{j \in [t_2]} \frac{x_j^{-\tau + \zeta_j + d_j, Q_j + (\tau - 1)|W_j| - 2E_{W_j, W_j}}{t_1 + t_2} dx t_2 \cdots dx_1.$$

We set $\tilde{S}_1 = S^*_1 \setminus W_{t_2}$, $\tilde{S}_2 = S^*_2 \setminus \{ t_2 \}$ and $\tilde{S}_3 = S^*_3 \cup W_{t_2} \cup \{ t_2 \}$. Notice that $E_{S^*_1} - E_{\tilde{S}_1} = E_{W_{t_2}} + E_{W_{t_2}, S^*_1 \setminus W_{t_2}}$, $E_{\tilde{S}_1} - E_{\tilde{S}_1, S^*_3} = E_{W_{t_2}, S^*_3} - E_{t_2, S^*_1 \setminus W_{t_2}} - E_{W_{t_2}, S^*_1 \setminus W_{t_2}}$, $E_{\tilde{S}_1, V_i} - E_{\tilde{S}_1, V_1} = E_{W_{t_2}, V_i}$ and $E_{S^*_2} - E_{\tilde{S}_2, V_1} = E_{W_{t_2}, V_i}$. Because the optimal solution to (2.2) is unique, we obtain using (7.23) that

$$-\tau + (\tau - 1)|W_{t_2}| - 2E_{W_{t_2}} - E_{W_{t_2}, S^*_1 \setminus W_{t_2}}$$

$$- E_{W_{t_2}, S^*_3} - E_{W_{t_2}, V_1} + d_{t_2, S^*_1 \setminus W_{t_2}} + d_{t_2, V_1} > -1.$$

Using that $W_{t_2} = V_1 \cup S^*_1 \cup S^*_3 \cup \{ t_2 \} \setminus W_{t_2}$ and $E_{W_{t_2}, W_{t_2}} = E_{W_{t_2}, S^*_1 \setminus W_{t_2}} + E_{W_{t_2}, S^*_3} + d_{t_2, W_{t_2}} + E_{W_{t_2}, V_1}$ and that $\zeta_{t_2} = d_{t_2, V_1}$ then shows that

$$-\tau + (\tau - 1)|W_{t_2}| - 2E_{W_{t_2}} - E_{W_{t_2}, W_{t_2}} + E_{W_{t_2}, S^*_3} + d_{t_2, S^*_1 \setminus W_{t_2}} + \zeta_{t_2} > -1.$$

Using that $d_{t_2, S^*_1 \setminus W_{t_2}} + d_{t_2, W_{t_2}} = d_{t_2, S^*_1}$ results in

$$-\tau + (\tau - 1)|W_{t_2}| - 2E_{W_{t_2}} - E_{W_{t_2}, W_{t_2}} + d_{t_2, S^*_1} + \zeta_{t_2} > -1.$$
Finally, using that \( Q_{t_2} = S_1^* \) shows that the inner integral in (7.37) is finite. A similar argument, setting \( S_1 = S_1^* \setminus (W_{t_2} \cup W_{t_2 - 1}) \) and \( S_2 = S_2^* \setminus \{t_2, t_2 - 1\} \) shows that the second integral is also finite, and we can proceed to show that the outer integral of (7.37) is finite. Because the ordering \( x_1 > x_2 > \cdots > x_{t_2} \) was arbitrary, the integral is finite over any reordering, so that (7.27) is finite.

**Proof of Theorem 2.1(i).** Because \( D_{\max} = O_p(n^{1/(\tau - 1)}) \), for any \( \eta_n \to 0 \), \( D_{\max} \leq n^{1/(\tau - 1)}/\eta_n \) with high probability. Define

\[
\gamma_i^a(n) = \begin{cases} 
   n^{1/(\tau - 1)}/\eta_n & \text{if } i \in S_2^*, \\
   n^{\alpha_i^{(\text{sub})}}/\varepsilon_n & \text{else},
\end{cases}
\]

with \( \alpha_i^{(\text{sub})} \) as in (2.5)

\[
\gamma_i^b(n) = \begin{cases} 
   1 & \text{if } i \in V_1, \\
   \varepsilon_n n^{\alpha_i^{(\text{sub})}} & \text{else}.
\end{cases}
\]

We then show that the expected number of subgraphs where the degree of at least one vertex \( i \) satisfies \( D_i \notin [\gamma_i^a(n), \gamma_i^b(n)] \) is small, similarly to the proof of Theorem 2.2.

We first study the expected number of copies of \( H \) where vertex 1 has degree in \([1, \gamma_1^a(n)]\) and all other vertices satisfy \( D_i \in [\gamma_i^b(n), \gamma_i^b(n)] \), by integrating the probability that subgraph \( H \) is formed over the range where vertex 1 has degree in \([1, \gamma_1^a(n)]\) and all other vertices satisfy \( D_1 \in [\gamma_1^b(n), \gamma_1^b(n)] \).

Using that the connection probabilities can be bounded by \( M_1 \min(D_1, D_j/(\mu n), 1) \) for some \( M_1 > 0 \) and the degree distribution can be bounded as \( \mathbb{P}(D = k) \leq M_2 k^{-\tau} \) for some \( M_2 > 0 \), we bound the expected number of such copies of \( H \) by

\[
\mathbb{E}
\left[
N^{(\text{sub})}(H) | \{D_1 < \gamma_1^a(n), D_i \in [\gamma_i^b(n), \gamma_i^b(n)] \ \forall i \geq 1\}
\right]
\leq Kn^k \int_{\gamma_1^a(n)}^{\gamma_2^a(n)} \int_{\gamma_2^b(n)}^{\gamma_2^b(n)} \cdots \int_{\gamma_k^b(n)}^{\gamma_k^b(n)} (x_1 \cdots x_k)^{-\tau} \prod_{\{i, j\} \in E_H} \min\left(\frac{x_i x_j}{\mu n}, 1\right) \ dx_k \cdots dx_1,
\]

for some \( K > 0 \). This integral equals zero when vertex 1 is in \( V_1 \). Suppose vertex 1 is in \( S_2^* \). W.l.o.g. assume that \( S_2^* = \{1, \ldots, t_2\} \), \( S_1^* = \{t_2 + 1, \ldots, t_1 + t_2\} \) and \( S_3^* = \{t_1 + t_2 + 1, \ldots, t_1 + t_2 + t_3\} \). We bound the minimum in (7.43) by \( x_i x_j/(\mu n) \) for \( i, j \in S_1^* \), for \( i \) or \( j \) in \( V_1 \) and for \( i \in S_1^*, j \in S_3^* \) or vice versa. We bound the minima by 1 for \( i, j \in S_2^* \) and \( i \in S_2^*, j \in S_3^* \) or vice versa.
Applying the change of variables \( y_i = x_i/n^{a_i^{(sub)}} \) results for some \( \tilde{K} > 0 \) in the bound

\[
\mathbb{E} \left[ N^{(sub)}(H) \mathbf{1}_{\{D_1<\gamma'_1(n),D_i\in[\gamma'_i(n),\gamma''_i(n)] \ \forall i \geq 1\}} \right] \leq \tilde{K} n^{|S^*_1|(2-\tau)+|S^*_2|(1-\tau)/2-|S^*_2|} 
\times n^k n^{-(\tau-1)} E_{S^*_1}^* + \frac{1}{\tau-1} E_{S^*_1}^* - \frac{1}{\tau} E_{V_1} s^*_1 - \frac{1}{\tau} E_{V_1} s^*_2 
\times \int_0^{\varepsilon_n} \int_0^{1/\eta_n} \cdots \int_0^{1/\eta_n} \int_0^{\infty} \prod_{i \in V_H \setminus V_1} y_i^{-\tau+\zeta_i} 
\times \prod_{(u,v) \in E_{S_3^1} \cup E_{S_1^*, S_2^*}} \min(y_u y_v, 1) dy_{t_1+1} \cdots dy_{t_2+1} \cdots dy_{t_3+1},
\]

where the integrals from 0 to 1/\( \eta_n \) correspond to vertices in \( S^*_2 \) and the integrals from 0 to \( \infty \) to vertices in \( S^*_1 \) and \( S^*_2 \). Since \( \tau \in (2,3) \), the integrals corresponding to vertices in \( V_1 \) are finite. By the analysis from (4.37) to (4.40),

\[
\eta |S^*_1| (2-\tau)+|S^*_2| (1-\tau)/2-|S^*_2| n^{\frac{1}{\tau-1} E_{S_1^*} + \frac{1}{\tau-1} E_{S_1^*} - \frac{1}{\tau} E_{V_1} s_1 - \frac{1}{\tau} E_{V_1} s_2} 
= n^{\frac{3-\tau}{\tau} (k_2 + B^{(sub)}(H)) + k_1}
\]

The integrals over \( y_i \in V_H \setminus V_1 \) can be split into

\[
\int_0^{\varepsilon_n} \int_0^{1/\eta_n} \cdots \int_0^{1/\eta_n} \int_0^{\infty} \prod_{i \in S_1^* \cup S_2^*} y_i^{-\tau+\zeta_i} \prod_{(u,v) \in E_{S_1^*, S_2^*}} \min(y_u y_v, 1) dy_{t_1+1} \cdots dy_{t_2+1} \cdots dy_{t_3+1},
\]

By Lemma 7.2 the set of integrals on the second line of (7.45) is finite. Lemma 7.3 shows that the set of integrals on the first line of (7.45) tends to zero for \( \eta_n \) fixed and \( \varepsilon_n \to 0 \). Thus, choosing \( \eta_n \to 0 \) sufficiently slowly compared to \( \varepsilon_n \) yields

\[
\int_0^{\varepsilon_n} \int_0^{1/\eta_n} \cdots \int_0^{1/\eta_n} \int_0^{\infty} \prod_{i \in S_1^* \cup S_2^*} y_i^{-\tau+\zeta_i} \prod_{(u,v) \in E_{S_1^*, S_2^*}} \min(y_u y_v, 1) dy_{t_1+1} \cdots dy_{t_2+1} \cdots dy_{t_3+1} 
= o(1).
\]
Therefore, 
(7.47) \[ \mathbb{E} \left[ N^{(sub)}(H) \mathbb{1}\{D_1 < \gamma_1(n), D_i \in [\gamma_i^L(n), \gamma_i^U(n)] \ \forall i > 1\} \right] = o \left( n \frac{3-\tau}{2} (k_2 + B^{(sub)}(H) + k_1) \right) , \]
when vertex 1 \( \not\in S^*_2 \). Similarly, we can show that the expected contribution from \( D_1 < \gamma_1(n) \) satisfies the same bound when vertex 1 is in \( S^*_1 \) or \( S^*_3 \). The expected number of subgraphs where \( D_1 > \gamma_1(n) \) if vertex 1 is in \( S^*_1 \), \( S^*_3 \) or \( V_1 \) can be bounded similarly, as well as the expected contribution where multiple vertices have \( D_i \not\in [\gamma_i^L(n), \gamma_i^U(n)] \).

Denote (7.48) \( \Gamma_n(\varepsilon_n, \eta_n) = \{ (v_1, \ldots, v_k) : D_{v_i} \in [\gamma_i^L, \gamma_i^U] \} \), and define \( \bar{\Gamma}_n(\varepsilon_n, \eta_n) \) as its complement. Denote the number of subgraphs with vertices in \( \bar{\Gamma}_n(\varepsilon_n, \eta_n) \) by \( N^{(sub)}(H, \bar{\Gamma}_n(\varepsilon_n, \eta_n)) \). Since \( D_{\max} \leq n^{1/(\tau-1)}/\eta_n \) with high probability, \( \Gamma_n(\varepsilon_n, \eta_n) = M_n(\alpha^{(sub)}) \) with high probability. Therefore, with high probability
(7.49) \[ N^{(sub)}(H, M_n(\alpha^{(sub)})) (\varepsilon_n) = N^{(sub)}(H, \bar{\Gamma}_n(\varepsilon_n, \eta_n)) , \]
where \( N^{(sub)}(H, M_n(\alpha^{(sub)})) (\varepsilon_n) \) denotes the number of copies of \( H \) on vertices not in \( M_n(\alpha^{(sub)})(\varepsilon_n) \). By the Markov inequality
(7.50) \[ N^{(sub)}(H, \bar{\Gamma}_n(\varepsilon_n, \eta_n)) = o_p \left( n \frac{3-\tau}{2} (k_2 + B^{(sub)}(H) + k_1) \right) . \]

Combining this with the fact that by Theorem 2.1(ii) for fixed \( \varepsilon \)
(7.51) \[ N^{(sub)}(H) = N^{(sub)}(H, M_n(\alpha^{(sub)}))(\varepsilon) + N^{(sub)}(H, \bar{M}_n(\alpha^{(sub)}))(\varepsilon) \\
= o_p \left( n \frac{3-\tau}{2} (k_2 + B^{(sub)}(H) + k_1) \right) \]
shows that
(7.52) \[ \frac{N^{(sub)}(H, M_n(\alpha^{(sub)}))(\varepsilon_n)}{N^{(sub)}(H)} \overset{p}{\rightarrow} 1. \]
\( \square \)
8. Induced subgraphs We now describe how to adapt the analysis of subgraphs to induced subgraphs. For induced subgraphs we can define a similar optimization problem as (4.29). When \( \alpha_i + \alpha_j < 1 \), (4.4) results in

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

whereas for \( \alpha_i + \alpha_j > 1 \), (4.17) yields

\[
\mathbb{P}_n(X_{ij} = 0) = o(1),
\]

and for \( \alpha_i + \alpha_j = 1 \) (4.4) yields

\[
\mathbb{P}_n(X_{ij} = 0) = \Theta(1).
\]

Similarly to (4.21), we can write the probability that \( H \) occurs as an induced subgraph on \( v = (v_1, \ldots, v_k) \) as

\[
\mathbb{P}_n(\text{ECM}^{(\alpha)}(D) | v = E_H) = \Theta^2 \left( \prod_{(v_i, v_j) \in E_H : \alpha_i + \alpha_j < 1} n^{\alpha_i+\alpha_j-1} \prod_{(i,j) \notin E_H : \alpha_i + \alpha_j > 1} e^{-n^{\alpha_i+\alpha_j-1}/2} \right). \tag{8.3}
\]

Similarly to (4.21), edges with \( \alpha_i + \alpha_j = 1 \) do not contribute to the order of magnitude of (8.3). Thus, the probability that \( H \) is an induced subgraph on \( v \) is stretched 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 (4.28) for induced subgraphs becomes

\[
\max (1 - \tau) \sum_i \alpha_i + \sum_{\{i,j\} \in E_H : \alpha_i + \alpha_j < 1} \alpha_i + \alpha_j - 1, \tag{8.4}
\]

s.t. \( \alpha_i + \alpha_j \leq 1 \quad \forall (i,j) \notin E_H \).

The following lemma shows that this optimization problem attains its optimum for very specific values of \( \alpha \) (similarly to Lemma 4.2 for subgraphs):

**Lemma 8.1 (Maximum contribution to induced subgraphs).** Let \( H \) be a connected graph on \( k \) vertices. If the solution to (8.5) 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 similar to the proof of Lemma 4.2. First, we again define \( \beta_i = \alpha_i - \frac{1}{2} \), so that (8.4) becomes

\[
\max \frac{1-\tau}{2} k + (1-\tau) \sum_i \beta_i + \sum_{\{i,j\} \in E_H : \beta_i + \beta_j < 0} \beta_i + \beta_j, \tag{8.5}
\]

s.t. \( \beta_i + \beta_j \leq 0 \quad \forall (i,j) \notin E_H \).
The proof of Step 1 from Lemma 4.2 then also holds for induced subgraphs. Now we prove that if the optimal solution to (8.5) 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 (4.33), and assume that \( \tilde{\beta} < \frac{3 - \tau}{2(\tau - 1)} \). The contribution of the vertices with \( |\beta_i| = \tilde{\beta} \) is as in (4.34). By increasing \( \tilde{\beta} \) 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 4.2 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 4.2 shows that if the solution to (8.5) 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.

Following the same lines as the proof of Theorem 2.1(ii) for subgraphs, Theorem 2.1(ii) for induced subgraphs follows, where we now use Lemma 8.1 instead of 4.2. We now state an equivalent lemma to Lemma 5.1 for induced subgraphs:

**Lemma 8.2** (Convergence of major contribution to induced subgraphs). Let \( H \) be a connected graph on \( k > 2 \) vertices such that (2.3) is uniquely optimized at 0. Then,

(i) The number of induced subgraphs with vertices in \( W_n^k(\varepsilon) \) satisfies

\[
\frac{N^{(\text{ind})}(H, W_n^k(\varepsilon))}{n^{k/2}(3-\gamma)} = (1 + o_\theta(1))e^{\frac{k}{2}(\gamma - 1)} \int_{\varepsilon}^{1/\varepsilon} \cdots \int_{\varepsilon}^{1/\varepsilon} (x_1 \cdots x_k)^{-\gamma} \times \prod_{\{i,j\} \in E_H} (1 - e^{-x_i x_j}) \prod_{(i,j) \notin E_H} e^{-x_i x_j} dx_1 \cdots dx_k + f_n(\varepsilon),
\]

for some function \( f_n(\varepsilon) \) such that for any \( \delta > 0 \),

\[
\lim_{\varepsilon \downarrow 0} \limsup_{n \to \infty} P(f_n(\varepsilon) > \delta | J_n) = 0.
\]

(ii) \( A^{(\text{ind})}(H) \) defined in (2.11) satisfies \( A^{(\text{ind})}(H) < \infty \).

The proof of Theorem 2.2 for induced subgraphs is similar to the proof of Theorem 2.2 for subgraphs, using Lemma 8.2 instead of Lemma 5.1. The proof of Lemma 8.2(i) in turn follows from straightforward extensions of Lemmas 6.1 6.3 and 6.2 for induced subgraphs, now also using that the probability that an edge \( \{i,j\} \notin E_H \) is not present in the subgraph can be approximated by \( \exp(-D_i D_j / L_n) \). Lemma 8.2(ii) is an application of the following equivalent lemma to Lemma 7.2 for \( S^*_3 = V_H \):
Lemma 8.3. Suppose that the maximum in (2.3) is uniquely attained for $|S^*_3| = s > 0$, and say $S^*_3 = \{1, \ldots, s\}$. Then

$$\int_0^\infty \cdots \int_0^\infty \prod_{i \in [s]} x_i^{\tau_+ + \zeta_i} \prod_{(u,v) \in E_{S^*_3}} \min(x_u x_v, 1) \prod_{(u,v) \notin E_{S^*_3}} e^{-x_u x_v} dx_s \ldots dx_1 < \infty. \tag{8.8}$$

Proof. This integral is finite if

$$\int_0^\infty \cdots \int_0^\infty \prod_{i \in [s]} x_i^{\tau_+ + \zeta_i} \prod_{(u,v) \in E_{S^*_3}} \min(x_u x_v, 1) \prod_{(u,v) \notin E_{S^*_3}} \mathbb{1}_{\{x_u x_v < 1\}} dx_s \ldots dx_1 < \infty, \tag{8.9}$$

since if

$$\int_a^b \int_0^{1/x_1} x_1^{x_2} e^{-x_1 x_2} dx_2 dx_1 < \infty, \tag{8.10}$$

then also

$$\int_a^b \int_{1/x_1}^\infty x_1^{x_2} e^{-x_1 x_2} dx_2 dx_1 < \infty. \tag{8.11}$$

We can show similarly to (7.5) and (7.6) that the integral is finite when all integrands are larger than one, or when all are smaller than one. We compute the contribution to (8.9) where the integrand runs from 1 to $\infty$ for vertices in some nonempty set $U$, and from 0 to 1 for vertices in $\bar{U} = S^*_3 \setminus U$. W.l.o.g., assume $U = \{1, \ldots, t\}$ for some $1 \leq t \leq s$. Define, for $i \in \bar{U}$,

$$\hat{h}(i, x) = \int_0^1 x_i^{\tau_+ + \zeta_i + d_i} \prod_{j \in U \setminus \{i,j\} \in E_H} \min(x_i x_j, 1) \prod_{v \in U \setminus \{i,v\} \notin E_H} \mathbb{1}_{\{x_i x_v < 1\}} dx_i. \tag{8.12}$$

Then (8.8) results in

$$\int_1^\infty \cdots \int_1^\infty \prod_{j \in [t]} x_j^{\tau_+ + \zeta_j} \prod_{u,v \in U \setminus \{u,v\} \notin E_H} \mathbb{1}_{\{x_u x_v < 1\}} \prod_{i = t+1}^k \hat{h}(i, x) dx_t \ldots dx_1. \tag{8.13}$$

When the induced subgraph of $H$ formed by the vertices of $U$ is not a complete graph, this integral equals zero. Thus, we assume that the induced subgraph of $H$ formed by the vertices of $U$ is a complete graph so that $\{(u,v) \in U \mid (u,v) \notin E_H\} \neq \emptyset$. We first bound the region of the integral where $1 < x_1 < \ldots < x_t$. When $i$ is connected to all vertices in $U$, $\hat{h}(i, x)$
equals \( h(i, x) \) defined in (7.8), which can be bounded by (7.11). Otherwise, define

\[
(8.14) \quad a_i = \max \{ j \in [t] : \{ i, j \} \notin E_H \}.
\]

Thus, \( i \) is connected to vertices \( t, t-1, \ldots, a_i-1 \) and we can write \( \hat{h}(i, x) \) as

\[
\hat{h}(i, x) = \int_0^1 x_1^{-\tau + \zeta_i + d_i, S_3^*} dx_1 \cdot x_t \cdot x_{t-1} \cdot x_{a_i-1} \prod_{j \in [a_i-1] \setminus \{ i \}} x_j + \cdots
\]

\[
+ \int_{1/xa_i-1}^{1/xa_{i-1}} x_1^{-\tau + \zeta_i + d_i, S_3^*} dx_1 \cdot x_t \cdot x_{t-1} \cdot x_{a_i-1} \prod_{j \in [a_i-1] \setminus \{ i \}} x_j + \cdots
\]

(8.15)

By a similar argument as in Lemma 7.1, \( \zeta_i + d_i, S_3^* \geq 2 \) for \( i \in S_3^* \) so that the first integral is finite. Thus, for some constants \( C_t, \ldots, C_{t-a_i+1} \)

\[
\hat{h}(i, x) = \prod_{j \in [a_i-1] \setminus \{ i \}} x_j \left( C_1 x_t^{-\zeta_i - d_i, S_3^*} x_t \cdot x_{t-1} \cdot x_{a_i-1} + C_2 x_t^{-\zeta_i - d_i, S_3^* + 1} x_t \cdot x_{t-1} \cdot x_{a_i-1} + \cdots + C_{t-a_i+1} x_{a_i} \right)
\]

(8.16)

\( =: \hat{h}_1(i, x) + \cdots + \hat{h}_{t-a_i+1}(i, x) \).

As in (7.11), for every \( i \) we can find an \( i^* \) such that for all \( 1 > x_1 > \cdots > x_t \),

\[
(8.17) \quad \hat{h}(i, x) \leq Kh_{i^*}(i, x)
\]

for some \( K > 0 \). Again, let \( T_i \) denote the set of neighbors of vertex \( i \) appearing in \( h_{i^*}(i, x) \), and set \( u_i, f(i) \) and \( Q_j \) as in (7.13)-(7.15), and let \( W_j = \{ i \in U : u_i = j \} \). Then,

\[
(8.18) \quad \int_1^\infty \cdots \int_1^\infty \prod_{j \in [t]} x_j^{-\tau + \zeta_i} \prod_{i=t+1}^k \hat{h}(i, x) dx_t \cdots dx_1
\]

\[
\leq K \int_1^\infty \cdots \int_1^\infty \prod_{j \in [t]} x_j^{-\tau + \zeta_i} \prod_{i=t+1}^k \hat{h}_{i^*}(i, x) dx_t \cdots dx_1
\]

\[
\leq \int_1^\infty \cdots \int_1^\infty \prod_{j \in [t]} x_j^{-\tau + \zeta_j + d_j, Q_j} \prod_{j=1}^t \hat{h}_{i^*}(i, x) dx_t \cdots dx_1
\]
for some $K > 0$, where $\hat{W}_j = (V_1 \cup S_1^* \cup [j] \cup \bar{U}) \setminus W_j$. We can now show that the integral over $x_t$ is finite in a similar manner as in Lemma 7.2. Define $\hat{S}_1 = S_1^* \cup W_t$, $\hat{S}_2 = S_2^* \cup \{t\}$ and $\hat{S}_3 = S_3^* \setminus (W_t \cup \{t\})$. Because $t \in S_3^*$, by constraint (2.3), $t$ is connected to all other vertices in $\hat{S}_2$, so that the vertices of $\hat{S}_2$ still form a complete graph. Furthermore, $t \in U$, so that $t$ is connected to all other vertices in $U$, since the vertices of $U$ formed a complete graph. Also, when $i \in \bar{U}$ is not connected to $t$, then $i \in W_t$ by (8.15) and the definition of $a_i$ in (8.14). Thus, $t$ is connected to all vertices in $U \cup \bar{U} \setminus (W_t \cup \{t\}) = \hat{S}_3$. Therefore, $\hat{S}_1$, $\hat{S}_2$ and $\hat{S}_3$ still satisfy the constraint in (2.3), and we may proceed as in Lemma 7.2 using (7.23) to show that the integral over $x_t$ finite. Iterating this proves Lemma 8.3.

The following lemma is the counterpart of Lemma 7.3 for induced subgraphs:

**Lemma 8.4.** Suppose that the optimal solution to (2.3) is unique, and attained by $S_1^*$, $S_2^*$ and $S_3^*$. Say that $S_2^* = \{1, \ldots, t_2\}$ and $S_1^* = \{t_2+1, \ldots, t_2+t_1\}$. Then,

$$
\int_0^1 \cdots \int_0^1 \int_0^\infty \cdots \int_0^\infty \prod_{j \in [t_1+t_2]} x_j^{-\tau + \zeta_j} \prod_{(u,v) \in E_{S_1^*, S_2^*}} \min(x_u x_v, 1) \times \prod_{(u,v) \notin E_{S_1^*, S_2^*}} e^{-x_u x_v} \, dx_{t_1+t_2} \cdots dx_1 < \infty.
$$

(8.19)

**Proof.** This lemma can be proven along similar lines as Lemma 7.3, with similar adjustments as the adjustments to prove Lemma 8.3 for induced subgraphs from its counterpart for subgraphs, Lemma 7.2.

From these lemmas, the proof of Theorem 2.1(i) for induced subgraphs follows along the same lines as the proof of Theorem 2.1(i) for subgraphs.

**REFERENCES**


