

OPTIMAL SUBGRAPH STRUCTURES IN SCALE-FREE CONFIGURATION MODELS

BY REMCO VAN DER HOFSTAD¹, JOHAN S. H. VAN LEEUWAARDEN² AND CLARA STEGEHUIS³

¹*Department of Mathematics and Computer Science, Eindhoven University of Technology, r.v.d.hofstad@tue.nl*

²*Department of Econometrics and Operations Research, Tilburg University, j.s.h.vanleeuwen@tilburguniversity.edu*

³*Department of Electrical Engineering, Mathematics and Computer Science, University of Twente, c.stegehuis@utwente.nl*

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, 22, 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, that is, vertices of extremely high degrees.

Another property of real-world scale-free networks is that the clustering coefficient (the probability that two uniformly chosen neighbors of a vertex are neighbors themselves) decreases with the vertex degree [4, 10, 24, 31, 34], 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 [29]. 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 [25, 26, 35] 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 [12, 25, 27]. 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, 17]. For $\tau \in (2, 3)$, however, the configuration model fails to create simple graphs with high probability [21]. We therefore consider the erased configuration model ([8] and [17], Chapter 7), which constructs a configuration model and

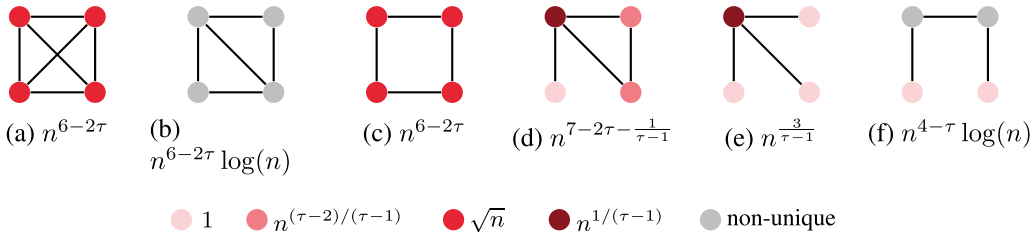


FIG. 1. *Optimal structures and asymptotic counts of induced subgraphs on 4 vertices. The vertex colors indicate the typical degrees and the scaling of the number of subgraphs is given below the pictures. Our results do not apply to the grey vertices.*

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 , typically occurs 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.

1.1. *The erased configuration model.* Let $[n] = \{1, 2, \dots, n\}$. Given a *degree sequence*, that is, a sequence of n positive integers $\mathbf{D} = (D_1, D_2, \dots, D_n)$, the *configuration model* is a (multi)graph with vertex set $[n]$, where vertex $v \in [n]$ has degree D_v . It is defined as follows (see, e.g., [6] or [17], Chapter 7): given a degree sequence with $\sum_{v \in [n]} D_v$ even, we start with D_v free half-edges adjacent to vertex v , for $v \in [n]$. 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 [17], 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 [20] (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

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

where $\tau \in (2, 3)$ so that $\mathbb{E}[D^2] = \infty$ and

$$(1.2) \quad \mathbb{E}[D] = \mu < \infty.$$

When this sample constructs a degree sequence such that the sum of the degrees is odd, we add an extra half-edge to the last vertex. This does not affect our computations. In this setting, D_{\max} is of order $n^{1/(\tau-1)}$, where $D_{\max} = \max_{v \in [n]} D_v$ denotes the maximal degree of the degree sequence. Denote the erased configuration model on n vertices by $\text{ECM}^{(n)}$ when the degrees are an independent and identically distributed (i.i.d.) sample of (1.1), and $\text{ECM}^{(n)}(\mathbf{D})$ when the degree sequence equals \mathbf{D} .

1.2. *Quenched and annealed.* Note that the erased configuration model as defined above has two sources of randomness: the independent and identically distributed (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 \mathcal{E} in the quenched setting by

$$(1.3) \quad \mathbb{P}_n(\mathcal{E}) = \mathbb{P}(\mathcal{E} \mid (D_v)_{v \in [n]}),$$

and we define \mathbb{E}_n and Var_n accordingly.

1.3. *Subgraph counts.* Let $H = (V_H, \mathcal{E}_H)$ be a small, connected graph. We denote the induced subgraph count of H , the number of subgraphs of $\text{ECM}^{(n)}$ 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 $\text{ECM}^{(n)}$, by $N^{(\text{sub})}(H)$.

Throughout this paper, we denote the sampled degree of a vertex $v \in [n]$ in the erased configuration model by D_v . 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 .

1.4. *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 7. We finally show how the proofs for subgraphs can be adjusted to prove the theorems on induced subgraphs in Section 8.

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

2. Main results. The key insight obtained in this paper is that the creation of subgraphs is crucially affected by the following trade-off, inherently present in power-law networks: On the one hand, hubs contribute substantially to the 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 $\text{ECM}^{(n)}|_v$ denote the induced subgraph of the erased configuration model on vertices v . Write the probability that a subgraph $H = (V_H, \mathcal{E}_H)$ with $|V_H| = k$ is created on k uniformly chosen vertices $v = (v_1, \dots, v_k)$ in $\text{ECM}^{(n)}$ as

$$(2.1) \quad \mathbb{P}(\text{ECM}^{(n)}|_v \supseteq H) = \sum_{D'} \mathbb{P}(\text{ECM}^{(n)}|_v \supseteq H \mid D_v = D') \mathbb{P}(D_v = D'),$$

where the sum is over all possible degrees on k vertices $D' = (D'_i)_{i \in [k]}$, and $D_v = (D_{v_i})_{i \in [k]}$ 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, \dots, 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 $n^{(\tau-2)/(\tau-1)}$, to \sqrt{n} or to $n^{1/(\tau-1)}$.

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 $H = (V_H, \mathcal{E}_H)$ be a small, connected graph on $k \geq 3$ 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 $n^{(\tau-2)/(\tau-1)}$, S_2 the ones with degrees proportional to $n^{1/(\tau-1)}$, and S_3 the vertices with degrees proportional to \sqrt{n} . The optimization problem finds the partition of the vertices into these three orders of magnitude 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, that is, it does not grow with n .

Given a partition $\mathcal{P} = (S_1, S_2, S_3)$ of $V_H \setminus V_1$, let \mathcal{E}_{S_i} denote the set of edges in H between vertices in S_i and $E_{S_i} = |\mathcal{E}_{S_i}|$ its size, \mathcal{E}_{S_i, S_j} the set of edges between vertices in S_i and S_j and $E_{S_i, S_j} = |\mathcal{E}_{S_i, S_j}|$ its size, and finally \mathcal{E}_{S_i, V_1} the set of edges between vertices in V_1 and S_i and $E_{S_i, V_1} = |\mathcal{E}_{S_i, V_1}|$ its size. We now define the optimization problem for subgraphs that optimizes the summand in (2.1) as

$$(2.2) \quad B^{(\text{sub})}(H) = \max_{\mathcal{P}} \left[|S_1| - |S_2| - \frac{2E_{S_1} + E_{S_1, S_3} + E_{S_1, V_1} - E_{S_2, V_1}}{\tau - 1} \right].$$

The first two terms in the optimization problem give a positive contribution for all vertices in S_1 , which have relatively low degree, and a negative contribution for vertices in S_2 , which have 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 the similar optimization problem

$$(2.3) \quad B^{(\text{ind})}(H) = \max_{\mathcal{P}^{(\text{ind})}} \left[|S_1| - |S_2| - \frac{2E_{S_1} + E_{S_1, S_3} + E_{S_1, V_1} - E_{S_2, V_1}}{\tau - 1} \right],$$

s.t. $\{i, j\} \in \mathcal{E}_H \forall i \in S_2, j \in S_2 \cup S_3,$

where again $\mathcal{P}^{(\text{ind})} = (S_1, S_2, S_3)$ 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 $S_3 = V_H$, that is, (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 $\alpha = (\alpha_1, \dots, \alpha_k)$ such that $\alpha_i \in [0, 1/(\tau - 1)]$, define

$$(2.4) \quad M_n^{(\alpha)}(\varepsilon) = \{(v_1, \dots, v_k) : D_{v_i} \in [\varepsilon, 1/\varepsilon](\mu n)^{\alpha_i} \forall i \in [k]\}.$$

These are the sets of vertices (v_1, \dots, v_k) such that D_{v_1} is proportional to n^{α_1} and D_{v_2} proportional to n^{α_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

$$(2.5) \quad \alpha_i^{(\text{sub})} = \begin{cases} (\tau - 2)/(\tau - 1) & i \in S_1^{(\text{sub})}, \\ 1/(\tau - 1) & i \in S_2^{(\text{sub})}, \\ \frac{1}{2} & i \in S_3^{(\text{sub})}, \\ 0 & i \in V_1. \end{cases}$$

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 ε_n such that $\lim_{n \rightarrow \infty} \varepsilon_n = 0$,*

$$(2.6) \quad \frac{N^{(\text{sub})}(H, M_n^{\alpha^{(\text{sub})}}(\varepsilon_n))}{N^{(\text{sub})}(H)} \xrightarrow{\mathbb{P}} 1.$$

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

$$(2.7) \quad \frac{N^{(\text{sub})}(H, M_n^{\alpha^{(\text{sub})}}(\varepsilon))}{n^{\frac{3-\tau}{2}(k_{2+} + B^{(\text{sub})}(H)) + k_1/2}} \leq f(\varepsilon) + o_{\mathbb{P}}(1),$$

and

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

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) and the optimization problem in (2.2) by that in (2.3).

Theorem 2.1(ii) only provides the scaling in n and some functions $f(\varepsilon), \tilde{f}(\varepsilon)$, which could tend to ∞ when $\varepsilon \searrow 0$. 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 $\text{ECM}^{(n)}$. 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 nonnegative 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,*

$$(2.9) \quad \frac{N^{(\text{sub})}(H)}{n^{\frac{k}{2}(3-\tau)}} \xrightarrow{\mathbb{P}} A^{(\text{sub})}(H) < \infty,$$

with

$$(2.10) \quad A^{(\text{sub})}(H) = c^k \mu^{-\frac{k}{2}(\tau-1)} \int_0^\infty \cdots \int_0^\infty (x_1 \cdots x_k)^{-\tau} \prod_{\{i,j\} \in \mathcal{E}_H} (1 - e^{-x_i x_j}) dx_1 \cdots dx_k.$$

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

$$(2.11) \quad \begin{aligned} A^{(\text{ind})}(H) &= c^k \mu^{-\frac{k}{2}(\tau-1)} \int_0^\infty \cdots \int_0^\infty (x_1 \cdots x_k)^{-\tau} \prod_{\{i,j\} \in \mathcal{E}_H} (1 - e^{-x_i x_j}) \\ &\times \prod_{\{i,j\} \notin \mathcal{E}_H} e^{-x_i x_j} dx_1 \cdots dx_k. \end{aligned}$$

In the erased configuration model, the probability that a vertex with degree D_{v_i} connects to a vertex with degree D_{v_j} can be approximated by $1 - e^{-D_{v_i} D_{v_j} / L_n}$, where $L_n = \sum_{v \in [n]} D_v$ denotes the sum of all degrees. When rescaling, and taking $D_{v_i} \approx x_i \sqrt{n/\mu}$ and $D_{v_j} \approx x_j \sqrt{n/\mu}$, this results in the factors $1 - e^{-x_i x_j}$ in (2.10) for all edges $\{i, j\} \in \mathcal{E}_H$ 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 factors $e^{-x_i x_j}$ for $\{i, j\} \notin \mathcal{E}_H$ in (2.11).

2.4. *Subgraphs on 4 and 5 vertices.* We 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 1(d) show that

$$(2.12) \quad B^{(\text{ind})}(H) = 2 - 1 + \frac{2 + 0 + 0 - 1}{\tau - 1} = 1 + \frac{1}{\tau - 1}.$$

By Theorem 2.1, the scaling of the induced subgraph in Figure 1(d) then equals

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

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) is unique. However, the gray vertices in

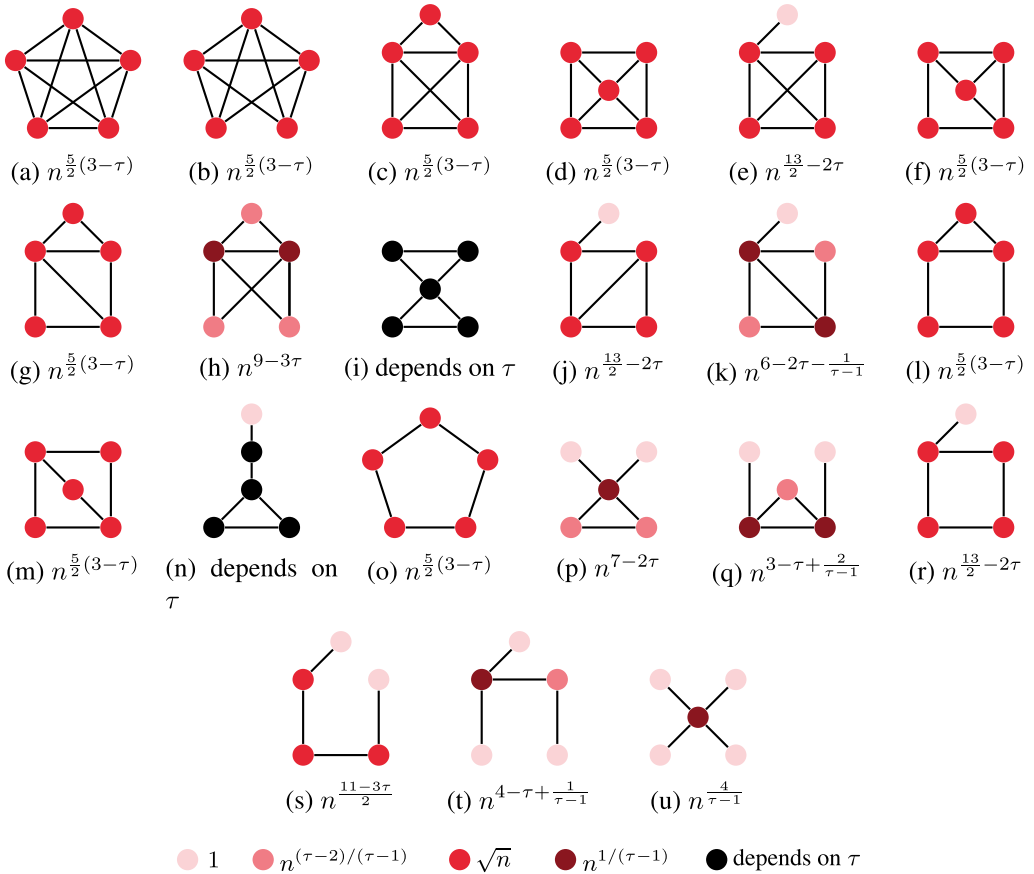


FIG. 2. The scaling of the number of subgraphs in n for induced subgraphs on 5 vertices. The vertex colors correspond to the typical degrees.

Figure 1 do not have unique optimizers, so that our theorems do not apply. Still, a 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 1(b) 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^α and the other two vertices have degrees $n^{1-\alpha}$ give an equal contribution for other values of α . Using that $D_i D_j$ follows a power-law distribution with exponent τ with an extra logarithmic factor ([19], equation (2.16)), then gives the extra factor $\log(n)$ in Figure 1(b).

The bow tie in Figure 2(i) has a unique optimal solution to (2.2), but it depends on τ . 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

$$(2.14) \quad 4 - 1 - \frac{2 \cdot 2}{\tau - 1} = \frac{3\tau - 7}{\tau - 1},$$

which is larger than zero when $\tau > 7/3$. Thus, for τ larger than $7/3$, the major contribution is when the middle vertex has degree $n^{1/(\tau-1)}$, and the other vertices have degrees $n^{(\tau-2)/(\tau-1)}$.

When 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 1(e) is similar to counting the number of sets of three neighbors for every vertex. The only sets of neighbors that we do not count, are neighbors that are connected. This is a small fraction of the pairs of neighbors ([15], equations (5)–(7)), thus the number of claws is approximately equal to

$$(2.15) \quad \sum_{v \in [n]} \frac{1}{6} D_v (D_v - 1)(D_v - 2) \approx \frac{1}{6} \sum_{v \in [n]} D_v^3.$$

Since the degrees are an i.i.d. sample from a power-law distribution, $\sum_{v \in [n]} D_v^3$ 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 to the deterministic leading order for \sqrt{n} degrees of Theorem 2.2.

The scaling of the number of (noninduced) subgraphs can be deduced from Figure 1. For example, we count the number of square subgraphs (the subgraph of Figure 1(c)) by adding the contributions from the induced subgraphs in Figures 1(a), 1(b) and 1(c), 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 1(b), which indeed contains a square, and occurs more frequently than the subgraphs of Figures 1(a) and 1(c). 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.

2.5.1. *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 vertices v_i, v_j with degrees D_{v_i}, D_{v_j} such that $D_{v_i} D_{v_j} = \Theta(n)$ across some edges $\{i, j\} \in \mathcal{E}_H$. One such example is the diamond of Figure 1(b) 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.

2.5.2. *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 .

2.5.3. *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 [28], so that

$$(2.16) \quad \limsup_{n \rightarrow \infty} \frac{\text{Var}(N^{(\text{sub})}(H))}{\mathbb{E}[N^{(\text{sub})}(H)]^2} \neq 0.$$

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 [28], the number of triangles is non-self-averaging in the annealed sense when τ is close to 3. This indicates that most realizations of $\text{ECM}^{(n)}(\mathbf{D})$ 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.

2.5.4. 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 analyze for triangles using similar techniques as in this paper in [13].

Another random graph model that generates simple power-law random graphs is the rank-1 inhomogeneous random graph [4, 9]. 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 [32].

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 [23]. 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 $\Theta(n^{\frac{k}{2}(3-\tau)})$ times as a subgraph in $\text{ECM}^{(n)}$. 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 power-law random graphs dynamically, are those that 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 $\text{ECM}^{(n)}$, with random degree sequence. In the proofs of Theorems 2.1 and 2.2, we often first study the quenched version of $\text{ECM}^{(n)}(\mathbf{D})$ instead, where the degree sequence \mathbf{D} is fixed.

We relate $L_n = \sum_{v \in [n]} D_v$, the total number of half-edges before erasure, to its expected value μn by defining the event

$$(3.1) \quad J_n = \{|L_n - \mu n| \leq n^{2/\tau}\}.$$

By [16], Lemma 2.3, $\mathbb{P}(J_n) \rightarrow 1$ as $n \rightarrow \infty$. 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 Var_n , we condition on the event J_n . We do not include J_n into the notation of \mathbb{P}_n , since given \mathbf{D} , J_n either happens with probability one, or with probability zero. This could be treated more formally by denoting

$$(3.2) \quad \mathbb{P}_n(\mathcal{E}) = \mathbb{1}_{J_n} \mathbb{P}(\mathcal{E} \mid \mathbf{D}),$$

but keep notation light, when using \mathbb{P}_n , we always assume the event J_n to hold.

Denote the indicator that an edge is present between vertices u and v by $X_{u,v}$. 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 $\text{ECM}^{(n)}(\mathbf{D})$ (see the proof of Lemma 4.1 for a more precise statement),

$$(3.3) \quad \mathbb{P}_n(X_{u,v} = 1) \approx 1 - e^{-D_u D_v / L_n}.$$

However, subgraphs often contain more than just one edge, and edges in $\text{ECM}^{(n)}(\mathbf{D})$ 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

$$(3.4) \quad N^{(\text{sub})}(H, M_n^{(\alpha)}(\varepsilon)) = \Theta_{\mathbb{P}}\left(n^{k+(1-\tau)\sum_i \alpha_i} \prod_{\{i,j\} \in \mathcal{E}_H: \alpha_i + \alpha_j < 1} n^{\alpha_i + \alpha_j - 1}\right).$$

To prove Theorem 2.1(ii) we optimize this over $\alpha = (\alpha_1, \dots, \alpha_k)$. Here ε 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 $N^{(\text{sub})}(H, M_n^{(\alpha)}(\varepsilon))$ 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 $\text{ECM}^{(n)}(\mathbf{D})$ 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 $\text{ECM}^{(n)}(\mathbf{D})$. 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 $\text{ECM}^{(n)}$. We explain how this changes the proof of Theorems 2.1 and 2.2 in more detail in Section 8.

4. Maximum contribution: Proof of Theorem 2.1(ii).

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*

$$(4.1) \quad \mathcal{E} = \{X_{u_i, v_i} = 0 \ \forall i \in [m]\}.$$

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

$$(4.2) \quad \mathbb{P}_n(X_{u_{m+1}, v_{m+1}} = 0 \mid \mathcal{E}) = O(e^{-D_{u_{m+1}} D_{v_{m+1}} / 4L_n}).$$

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

$$(4.3) \quad \mathbb{P}_n(X_{u_{m+1}, v_{m+1}} = 0 \mid \mathcal{E}) = e^{-\frac{D_{u_{m+1}} D_{v_{m+1}}}{L_n}} \left(1 + O\left(\frac{D_{u_{m+1}} D_{v_{m+1}}}{L_n} n^{-\gamma}\right)\right).$$

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 for the probability that a subgraph is present.

PROOF. For $m = 0$ the claim is proven in [18], eq (4.6) and (4.9), which states that for two vertices u and v with $D_u > D_v$,

$$(4.4) \quad \mathbb{P}_n(X_{u,v} = 0) = e^{-D_u D_v / L_n} + O(D_u^2 D_v / L_n^2),$$

and that, by using [18], equation (4.5), as well as $L_n - 2i + 1 \leq L_n$ and $1 - x \leq e^{-x}$,

$$(4.5) \quad \mathbb{P}_n(X_{u,v} = 0) \leq \prod_{i=1}^{D_u/2} \left(1 - \frac{D_v}{L_n - 2i + 1} \right) \leq e^{-D_u D_v / 2L_n}.$$

Thus we assume that $m \geq 1$. Note that $\Omega := \{\{u_i, v_i\}\}_{i \in [m]}$ may contain the same vertices multiple times. Denote the number of distinct vertices in $\{\{u_i, v_i\}\}_{i \in [m]}$ by r , and denote these distinct vertices by w_1, \dots, w_r . Let u_{m+1}, v_{m+1} correspond to w_r and w_{r-1} (if they are present in w_1, \dots, w_r at all). The ordering of the other vertices may be arbitrary.

We now construct $\text{ECM}^{(n)}(\mathbf{D})$ conditionally on the edges Ω not being present. We pair the half-edges of the erased configuration model attached to w_1, \dots, w_r . First we pair all half-edges adjacent to w_1 . Since we condition on the edges Ω not being present, no half-edge from w_1 is allowed to pair to any of its neighbors in Ω . 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 Ω , and so on. We continue until all of the forbidden edges Ω 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 Ω are present.

Let B denote the number of vertices we have to pair before all of the forbidden edges Ω 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 Ω have already been paired before arriving at u_{m+1} or v_{m+1} . Let $\hat{X}_{u,v}$ denote the number of half-edges between u and v in the configuration model, so that the edge indicator of the erased configuration model can be written as $X_{u,v} = \mathbb{1}\{\hat{X}_{u,v} > 0\}$. Furthermore, let $\mathcal{F}_{\leq s} = \sigma(\{\hat{X}_{w_i, j}\}_{i \leq s, j \in [n]})$ be the information about the pairings that have been constructed up to time s .

After pairing the half-edges incident to vertices in $[B]$, denote

$$(4.6) \quad \tilde{L}_n = L_n - 2 \sum_{i \in [B]} (D_{w_i} - \hat{X}_{w_i, w_i}),$$

which equals the remaining half-edges after pairing the half-edges incident to $(w_i)_{i \in [B]}$. 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, w_i} . Define $\tilde{D}_{u_{m+1}} = D_{u_{m+1}} - \sum_{i \in [B]} \hat{X}_{i, u_{m+1}}$, and define $\tilde{D}_{v_{m+1}}$ similarly. These quantities are all measurable with respect to $\mathcal{F}_{\leq B}$. The probability that u_{m+1} does not pair to v_{m+1} is the probability that u_{m+1} of degree $\tilde{D}_{u_{m+1}}$ does not connect to v_{m+1} of degree $\tilde{D}_{v_{m+1}}$ in a configuration model with \tilde{L}_n half-edges. Thus, using (4.4),

$$(4.7) \quad \mathbb{P}_n(X_{u_{m+1}, v_{m+1}} = 0 \mid \mathcal{F}_{\leq B}) = e^{-\tilde{D}_{u_{m+1}} \tilde{D}_{v_{m+1}} / \tilde{L}_n} + O(\tilde{D}_{u_{m+1}}^2 \tilde{D}_{v_{m+1}} / \tilde{L}_n^2),$$

where we have assumed w.l.o.g. that $\tilde{D}_{u_{m+1}} \geq \tilde{D}_{v_{m+1}}$.

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

$$(4.8) \quad \begin{aligned} \mathbb{P}_n(\text{jth half-edge pairs to } u_{m+1}) &\leq \frac{D_{u_{m+1}}}{L_n - 2j - 3 - 2 \sum_{s \in [i-1]} D_{w_s}} \\ &\leq K D_{u_{m+1}} / L_n, \end{aligned}$$

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 $\hat{X}_{w_i, u_{m+1}}$ by Y_{w_i} , where $Y_{w_i} \sim \text{Bin}(n^{1/(\tau-1)}/\varepsilon, K D_{u_{m+1}}/L_n)$, so that $\mathbb{E}[Y_{w_i}] = K_1 n^{-\beta} D_{u_{m+1}}$ 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 $\tilde{K} > 0$,

$$(4.9) \quad \mathbb{P}(Y_{w_i} > K_1 n^{-\beta} D_{u_{m+1}} (1 + n^\gamma)) \leq e^{-\tilde{K} n^{2\gamma-\beta} D_{u_{m+1}}}.$$

Define the events

$$(4.10) \quad \mathcal{B}_{n,u} = \{\exists i \in [B] : \hat{X}_{w_i, u_{m+1}} > K_1 n^{-\beta} D_{u_{m+1}} (1 + n^\gamma)\},$$

$$(4.11) \quad \mathcal{B}_{n,v} = \{\exists i \in [B] : \hat{X}_{w_i, v_{m+1}} > K_1 n^{-\beta} D_{v_{m+1}} (1 + n^\gamma)\},$$

and let $\mathcal{B}_{n,u}^c$ and $\mathcal{B}_{n,v}^c$ denote their respective complements, so that, by a union bound,

$$(4.12) \quad \mathbb{P}(\mathcal{B}_{n,u}^c) \geq 1 - B e^{-\tilde{K} n^{2\gamma-\beta} D_{u_{m+1}}}.$$

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

$$(4.13) \quad \tilde{D}_{u_{m+1}} \geq D_{u_{m+1}} \left(1 - \sum_{i \in [B]} \hat{X}_{w_i, u_{m+1}} \right) = D_{u_{m+1}} (1 + O(n^{\gamma-\beta})).$$

Similarly, $\tilde{D}_{v_{m+1}} = D_{v_{m+1}} (1 + O(n^{\gamma-\beta}))$ on $\mathcal{B}_{n,v}^c$, where $\mathbb{P}(\mathcal{B}_{n,v}^c) \geq 1 - B e^{-\tilde{K} n^{2\gamma-\beta} D_{v_{m+1}}}$. Then, when $D_{u_{m+1}} D_{v_{m+1}} = O(n)$ as assumed for (4.3), (4.7) becomes

$$(4.14) \quad \begin{aligned} &\mathbb{P}_n(X_{u_{m+1}, v_{m+1}} = 0 \mid \mathcal{F}_{B+1}, \mathcal{B}_{n,u}^c, \mathcal{B}_{n,v}^c) \\ &= e^{-\frac{D_{u_{m+1}} D_{v_{m+1}}}{L_n} (1 + O(n^{-\gamma}))} + O\left(\frac{D_{u_{m+1}}^2 D_{v_{m+1}}}{L_n^2}\right) \\ &= e^{-\frac{D_{u_{m+1}} D_{v_{m+1}}}{L_n}} \left(1 + O\left(\frac{D_{u_{m+1}} D_{v_{m+1}}}{L_n} n^{-\gamma}\right) \right), \end{aligned}$$

where we have used that $D_{u_{m+1}} = O(n^{1/(\tau-1)})$. Furthermore, $2\gamma - \beta > 0$, whereas by assumption $D_{u_{m+1}} D_{v_{m+1}} / L_n = O(1)$, so that (4.12) together with (4.14) proves (4.3).

To prove (4.2), we use (4.5) and the fact that on the event $\mathcal{B}_{n,u}^c \cap \mathcal{B}_{n,v}^c$, $D_{v_{m+1}} D_{u_{m+1}} \geq \tilde{D}_{v_{m+1}} \tilde{D}_{u_{m+1}} / 2$ for n sufficiently large to obtain

$$(4.15) \quad \begin{aligned} \mathbb{P}_n(X_{u_{m+1}, v_{m+1}} = 0 \mid \mathcal{F}_{B+1} \mathcal{B}_{n,u}^c, \mathcal{B}_{n,v}^c) &\leq e^{-\tilde{D}_{u_{m+1}} \tilde{D}_{v_{m+1}} / 2 \tilde{L}_n} \\ &\leq e^{-\tilde{D}_{u_{m+1}} \tilde{D}_{v_{m+1}} / 2 L_n} \\ &\leq e^{-D_{u_{m+1}} D_{v_{m+1}} / 4 L_n}. \end{aligned}$$

Combining this with (4.12) and the fact that $2\gamma - \beta > 0$ completes the proof of (4.2). \square

4.2. *An optimization problem.* We now use Lemma 4.1 to study the probability that a subgraph is present on vertices (v_1, \dots, v_k) of specific degrees. Assume that $D_{v_i} \in [\varepsilon, 1/\varepsilon]n^{\alpha_i}$ with $\alpha_i \in [0, 1/(\tau - 1)]$ for all i , so that $D_{v_i} = \Theta(n^{\alpha_i})$.

Let H be a subgraph on k vertices labeled as $1, \dots, k$, and with m edges labeled a $\mathcal{E}_H = \{\{i_1, j_1\}, \dots, \{i_m, j_m\}\}$. Furthermore, let $\text{ECM}^{(n)}(\mathbf{D})|_v$ be the induced subgraph of $\text{ECM}^{(n)}(\mathbf{D})$ on the vertices $v = (v_1, \dots, v_k)$. We aim to study the probability that this occurs.

When $\alpha_i + \alpha_j < 1$, by (4.4),

$$(4.16) \quad \mathbb{P}_n(X_{v_i, v_j} = 1) = (1 - e^{-\Theta(n^{\alpha_i + \alpha_j - 1})})(1 + o(1)) = \Theta(n^{\alpha_i + \alpha_j - 1}).$$

Furthermore, by (4.4), $\mathbb{P}_n(X_{v_i, v_j} = 1) = \Theta(1)$ and $\mathbb{P}_n(X_{v_i, v_j} = 0) = \Theta(1)$ when $\alpha_i + \alpha_j = 1$. When $\alpha_i + \alpha_j > 1$ instead, by (4.5),

$$(4.17) \quad \mathbb{P}_n(X_{v_i, v_j} = 1) = 1 - O(e^{-n^{\alpha_i + \alpha_j - 1}/(4\mu)}),$$

so that it equals 1 minus a stretched exponentially small term. For vertices $v_i, v_j \in [n]$ with $D_{v_i} \in [\varepsilon, 1/\varepsilon]n^{\alpha_i}$ and $D_{v_j} \in [\varepsilon, 1/\varepsilon]n^{\alpha_j}$, denote

$$(4.18) \quad w_{i,j} = n^{\alpha_i + \alpha_j - 1 - \gamma},$$

with γ as in (4.3). By Lemma 4.1, for any set of m edges in $[n]$, and $D_{u_p} \in [\varepsilon, 1/\varepsilon]n^{\alpha_{ip}}$, $D_{v_p} \in [\varepsilon, 1/\varepsilon]n^{\alpha_{jp}}$,

$$(4.19) \quad \begin{aligned} \mathbb{P}_n(X_{u_1, v_1} = \dots = X_{u_m, v_m} = 0) &= \prod_{p: \alpha_{ip} + \alpha_{jp} < 1} (1 + O(w_{i_p, j_p})) (1 - \Theta(n^{\alpha_{ip} + \alpha_{jp} - 1})) \\ &\times \prod_{p: \alpha_{ip} + \alpha_{jp} = 1} e^{-D_{u_p} D_{v_p} / (\mu n)} (1 + O(n^{-(\tau-2)/(\tau-1)})) \\ &\times \prod_{p: \alpha_{ip} + \alpha_{jp} > 1} O(e^{-n^{\alpha_{ip} + \alpha_{jp} - 1}/(4\mu)}). \end{aligned}$$

For ease of notation, we denote

$$(4.20) \quad q(i, j) = \begin{cases} (1 + O(w_{i,j}))(1 - \Theta(n^{\alpha_i + \alpha_j - 1})) & \text{if } \alpha_i + \alpha_j < 1, \\ e^{-D_{v_i} D_{v_j} / (\mu n)} (1 + O(n^{-\gamma})) & \text{if } \alpha_i + \alpha_j = 1, \\ O(e^{-n^{\alpha_i + \alpha_j - 1}/(4\mu)}) & \text{if } \alpha_i + \alpha_j > 1. \end{cases}$$

We write the probability that H is present on a specified subset of vertices $v = (v_1, \dots, v_k)$ as

$$(4.21) \quad \begin{aligned} &\mathbb{P}_n(\text{ECM}^{(n)}(\mathbf{D})|_v \supseteq \mathcal{E}_H) \\ &= 1 - \sum_{l=1}^m \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = 0) + \sum_{l \neq p} \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = X_{v_{i_p}, v_{j_p}} = 0) \\ &\quad - \sum_{l \neq p \neq r} \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = X_{v_{i_p}, v_{j_p}} = X_{v_{i_r}, v_{j_r}} = 0) + \dots \\ &\quad + (-1)^m \mathbb{P}_n(X_{v_{i_1}, v_{j_1}} = \dots = X_{v_{i_m}, v_{j_m}} = 0) \\ &= 1 - \sum_{l=1}^m q(i_l, j_l) + \sum_{l \neq p} q(i_l, j_l)q(i_p, j_p) \end{aligned}$$

$$\begin{aligned}
 & - \sum_{l \neq p \neq r} q(i_l, j_l)q(i_p, j_p)q(i_r, j_r) + \dots + (-1)^m \prod_{l \in [m]} q(i_l, j_l) \\
 & = \prod_{l \in [m]} (1 - q(i_l, j_l)) = \Theta \left(\prod_{\{i, j\} \in \mathcal{E}_H : \alpha_i + \alpha_j < 1} n^{\alpha_i + \alpha_j - 1} \right).
 \end{aligned}$$

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

$$(4.22) \quad 1 - q(i, j) = 1 - (1 - \Theta(n^{\alpha_i + \alpha_j - 1}))(1 + O(w_{i,j})) = \Theta(n^{\alpha_i + \alpha_j - 1}),$$

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

$$(4.23) \quad 1 - q(i, j) = 1 - O(e^{-n^{\alpha_i + \alpha_j - 1}/(4\mu)}) = 1 + o(1).$$

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

$$(4.24) \quad 1 - q(i, j) = (1 + O(n^{-\gamma}))(1 - e^{-D_i D_j / (\mu n)}) = \Theta(1),$$

so that edges $\{i, j\}$ 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,

$$\begin{aligned}
 \mathbb{P}(D_1 \in [\varepsilon, 1/\varepsilon](\mu n)^\alpha) & = \sum_{x = \varepsilon(\mu n)^\alpha}^{(\mu n)^\alpha / \varepsilon} cx^{-\tau}(1 + o(1)) \\
 (4.25) \quad & = O(1) \int_{\varepsilon(\mu n)^\alpha}^{1/\varepsilon(\mu n)^\alpha} cx^{-\tau} dx = K(\varepsilon) O((\mu n)^{\alpha(1-\tau)}),
 \end{aligned}$$

for some $K(\varepsilon)$ not depending on n . The number of vertices with degrees in $[\varepsilon, 1/\varepsilon](\mu n)^\alpha$ is Binomial($n, \mathbb{P}(D_1 \in [\varepsilon, 1/\varepsilon](\mu n)^\alpha)$), so that the number of vertices with degrees in $[\varepsilon, 1/\varepsilon](\mu n)^\alpha$ is $\Theta_{\mathbb{P}}(n^{(1-\tau)\alpha+1})$ for $\alpha \leq \frac{1}{\tau-1}$. Then, for $M_n^{(\alpha)}$ as in (2.4),

$$(4.26) \quad \# \text{ sets of vertices with degrees in } M_n^{(\alpha)} = \Theta_{\mathbb{P}}(n^{k+(1-\tau)\sum_i \alpha_i}).$$

Combining (4.21) and (4.26) yields

$$(4.27) \quad N^{(\text{sub})}(H, M_n^{(\alpha)}(\varepsilon)) = \Theta_{\mathbb{P}} \left(n^{k+(1-\tau)\sum_i \alpha_i} \prod_{\{i, j\} \in \mathcal{E}_H : \alpha_i + \alpha_j < 1} n^{\alpha_i + \alpha_j - 1} \right).$$

The maximum contribution is obtained for α that maximizes

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

The following lemma shows that this optimization problem attains its maximum for specific values of the exponents α_i :

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 .*

PROOF. Defining $\beta_i = \alpha_i - \frac{1}{2}$ yields that (4.28) equals

$$(4.29) \quad \max \frac{1-\tau}{2}k + (1-\tau) \sum_i \beta_i + \sum_{\{i,j\} \in \mathcal{E}_H: \beta_i + \beta_j < 0} (\beta_i + \beta_j),$$

over all possible values of $\beta_i \in [-\frac{1}{2}, \frac{3-\tau}{2(\tau-1)}]$. Then, we have to prove that $\beta_i \in \{-\frac{1}{2}, \frac{\tau-3}{2(\tau-1)}, 0, \frac{3-\tau}{2(\tau-1)}\}$ for all i in the optimal solution. Note that (4.29) is a piecewise linear function in β_1, \dots, β_k . Therefore, if (4.29) has a unique maximum, then it must be attained at the boundary for β_i or at a border of one of the linear sections. Thus, any unique optimal value of β_i satisfies $\beta_i = -\frac{1}{2}$, $\beta_i = \frac{\tau-3}{2(\tau-1)}$ or $\beta_i + \beta_j = 0$ for some j . We ignore the constant factor of $(1-\tau)\frac{k}{2}$ in (4.29), since it does not influence the optimal β values. Rewriting (4.29) without the constant factor yields

$$(4.30) \quad \max \sum_i \beta_i (1-\tau + |\{s \in [k] : \{s, i\} \in \mathcal{E}_H \text{ and } \beta_s < -\beta_i\}|).$$

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 $|\beta_i| \in (0, \frac{3-\tau}{2(\tau-1)})$.

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

Step 1. Let i be a vertex of degree 1 in H , and j be the neighbor of i . Let N_j denote the number of edges in H from j to other vertices v not equal to i with $\beta_v < -\beta_j$. The contribution from vertices i and j to (4.30) is

$$(4.31) \quad \beta_j(1-\tau + N_j) + \beta_i(1-\tau + \mathbb{1}_{\{\beta_i > -\beta_j\}}) + \beta_j \mathbb{1}_{\{\beta_i < -\beta_j\}}.$$

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

Let i be a vertex in V_H , and recall that d_i denotes the degree of i in H . Let i be such that $d_i \geq 2$ in H , and suppose that $\beta_i < \frac{\tau-3}{2(\tau-1)}$. Because the maximal value of β_j for $j \neq i$ is $\frac{3-\tau}{2(\tau-1)}$, the contribution to the i th term of (4.30) is

$$(4.32) \quad -\frac{1}{2}(1-\tau + d_i) < 0,$$

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

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

$$(4.33) \quad \tilde{\beta} = \min_{i: |\beta_i| > 0} |\beta_i|.$$

Let $N_{\tilde{\beta}^-}$ denote the number of vertices with their β 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

$$(4.34) \quad \tilde{\beta}((1-\tau)(N_{\tilde{\beta}^+} - N_{\tilde{\beta}^-}) + E_{\tilde{\beta}^+} - E_{\tilde{\beta}^-}).$$

Because we assume β 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 $\beta_i \in \{\frac{\tau-3}{2(\tau-1)}, 0, \frac{3-\tau}{2(\tau-1)}\}$ for all i such that $d_i \geq 2$.

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

$$(4.35) \quad \beta_* [(1 - \tau)(N_{\beta_*^+} - N_{\beta_*^-}) + E_{\beta_*^+} - E_{\beta_*^-}].$$

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

$$(4.36) \quad \hat{\beta}_j ((1 - \tau)(N_{\hat{\beta}_j^+} - N_{\hat{\beta}_j^-}) + E_{\hat{\beta}_j^+} - E_{\hat{\beta}_j^-}) = 0 \quad \forall j \in [s].$$

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 $i \in V_H$. \square

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 $\alpha_i^{(\text{sub})} = \frac{\tau-2}{\tau-1}$, vertices in S_2 have $\alpha_i^{(\text{sub})} = \frac{1}{\tau-1}$, vertices in S_3 have $\alpha_i^{(\text{sub})} = \frac{1}{2}$ 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. Recall that the number of edges inside S_1 is denoted 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

$$(4.37) \quad \max_{\mathcal{P}} \left[(1 - \tau) \left(\frac{\tau - 2}{\tau - 1} |S_1| + \frac{1}{\tau - 1} |S_2| + \frac{1}{2} |S_3| \right) + \frac{\tau - 3}{\tau - 1} E_{S_1} + \frac{\tau - 3}{2(\tau - 1)} E_{S_1, S_3} - \frac{E_{S_1, V_1}}{\tau - 1} - \frac{\tau - 2}{\tau - 1} E_{S_2, V_1} - \frac{1}{2} E_{S_3, V_1} \right],$$

over all partitions $\mathcal{P} = (S_1, S_2, S_3)$ of $V_H \setminus V_1$. Using that $|S_3| = k - |S_1| - |S_2| - k_1$, $E_{S_3, V_1} = k_1 - E_{S_1, V_1} - E_{S_2, V_1}$, where $k_1 = |V_1|$ and extracting a factor $(3 - \tau)/2$ shows that this is equivalent to

$$(4.38) \quad \frac{1 - \tau}{2} k + \max_{\mathcal{P}} \frac{(3 - \tau)}{2} \left(|S_1| - |S_2| + \frac{\tau - 2}{3 - \tau} k_1 - \frac{2E_{S_1} + E_{S_1, S_3}}{\tau - 1} - \frac{E_{S_1, V_1} - E_{S_2, V_1}}{\tau - 1} \right).$$

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

$$(4.39) \quad B^{(\text{sub})}(H) = \max_{\mathcal{P}} \left[|S_1| - |S_2| - \frac{2E_{S_1} + E_{S_1, S_3} + E_{S_1, V_1} - E_{S_2, V_1}}{\tau - 1} \right],$$

which equals (2.2). By (4.27), the contribution of the maximum is then given by

$$(4.40) \quad n^{\frac{3-\tau}{2}(k+B^{(\text{sub})}(H))+\frac{\tau-2}{2}k_1} = n^{\frac{3-\tau}{2}(k_2+B^{(\text{sub})}(H))+k_1/2},$$

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 $\alpha_i = \frac{1}{2}$ for all $i \in V_H = [k]$ as

$$(5.1) \quad W_n^k(\varepsilon) = \{(v_1, \dots, v_k) : D_{v_s} \in [\varepsilon, 1/\varepsilon]\sqrt{\mu n} \forall s \in [k]\},$$

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 $k \geq 3$ vertices such that (2.2) is uniquely optimized at $S_3 = [k]$, so that $B^{(\text{sub})}(H) = 0$. Then:*

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

$$(5.2) \quad \frac{N^{(\text{sub})}(H, W_n^k(\varepsilon))}{n^{\frac{k}{2}(3-\tau)}} = (1 + o_{\mathbb{P}}(1))c^k \mu^{-\frac{k}{2}(\tau-1)} \int_{\varepsilon}^{1/\varepsilon} \dots \int_{\varepsilon}^{1/\varepsilon} (x_1 \dots x_k)^{-\tau} \\ \times \prod_{\{i,j\} \in \mathcal{E}_H} (1 - e^{-x_i x_j}) dx_1 \dots dx_k + f_n(\varepsilon),$$

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

$$(5.3) \quad \lim_{\varepsilon \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(f_n(\varepsilon) > \delta \mid J_n) = 0;$$

(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 start by studying the expected number of subgraphs with vertices outside $W_n^k(\varepsilon)$. First, we investigate the expected number of subgraphs in the case where vertex 1 of the subgraph has degree smaller than $\varepsilon\sqrt{\mu n}$. Similar to (4.21), we can use Lemma 4.1 to show that the probability that H is present on a specified subset of vertices $\mathbf{v} = (v_1, \dots, v_k)$ can be written as

$$(5.4) \quad \mathbb{P}_n(\text{ECM}^{(n)}(\mathbf{D})|_{\mathbf{v}} \supseteq \mathcal{E}_H) = \Theta\left(\prod_{\{i,j\} \in \mathcal{E}_H : D_{v_i} D_{v_j} < L_n} (1 - e^{-D_{v_i} D_{v_j} / L_n})\right) \\ = \Theta\left(\prod_{\{i,j\} \in \mathcal{E}_H} (1 - e^{-D_{v_i} D_{v_j} / L_n})\right).$$

Furthermore, by (1.1), there exists C_0 such that $\mathbb{P}(D = k) \leq C_0 k^{-\tau}$ for all k . Let $I^{(\text{sub})}(H, \mathbf{v}) = \mathbb{1}_{\{\text{ECM}^{(n)}(\mathbf{D})|_{\mathbf{v}} \supseteq \mathcal{E}_H\}}$, so that $N^{(\text{sub})}(H) = \sum_{\mathbf{v}} I^{(\text{sub})}(H, \mathbf{v})$. Then, 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

$$(5.5) \quad \sum_{\mathbf{v}} \mathbb{E}[I^{(\text{sub})}(H, \mathbf{v}) \mathbb{1}_{\{D_{v_1} < \varepsilon\sqrt{\mu n}\}} \mid J_n] \\ \leq \Theta(1)n^k \int_1^{\varepsilon\sqrt{\mu n}} \int_1^{\infty} \dots \int_1^{\infty} (x_1 \dots x_k)^{-\tau} \prod_{\{i,j\} \in \mathcal{E}_H} (1 - e^{-x_i x_j / (\mu n)}) dx_1 \dots dx_k \\ = \Theta(1)n^k (\mu n)^{\frac{k}{2}(1-\tau)} \int_0^{\varepsilon} \int_0^{\infty} \dots \int_0^{\infty} (t_1 \dots t_k)^{-\tau} \prod_{\{i,j\} \in \mathcal{E}_H} (1 - e^{-t_i t_j}) dt_1 \dots dt_k \\ = O(n^{\frac{k}{2}(3-\tau)})h_1(\varepsilon),$$

where $h_1(\varepsilon)$ is a function of ε . By Lemma 5.1(ii), $h_1(\varepsilon) \rightarrow 0$ as $\varepsilon \searrow 0$. 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}/ε , similarly. This yields

$$(5.6) \quad \mathbb{E}[N^{(\text{sub})}(H, \bar{W}_n^k(\varepsilon)) \mid J_n] = O(n^{\frac{k}{2}(3-\tau)})h(\varepsilon),$$

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

$$(5.7) \quad N^{(\text{sub})}(H, \bar{W}_n^k(\varepsilon)) = h(\varepsilon)O_{\mathbb{P}}(n^{\frac{k}{2}(3-\tau)}).$$

Therefore, for any $\delta > 0$,

$$(5.8) \quad \limsup_{\varepsilon \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\frac{N^{(\text{sub})}(H, \bar{W}_n^k(\varepsilon))}{n^{k(3-\tau)/2}} > \delta \mid J_n\right) = 0.$$

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

$$(5.9) \quad \frac{N^{(\text{sub})}(H)}{n^{\frac{k}{2}(3-\tau)}} \xrightarrow{\mathbb{P}} c^k \mu^{-\frac{k}{2}(\tau-1)} \int_0^\infty \dots \int_0^\infty (x_1, \dots, x_k)^{-\tau} \times \prod_{\{i,j\} \in \mathcal{E}_H} (1 - e^{-x_i x_j}) dx_1 \dots dx_k. \quad \square$$

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 $\text{ECM}^{(n)}(\mathbf{D})$. Let H be a subgraph on k vertices, labeled as $[k]$, and m edges, denoted by $e_1 = \{i_1, j_1\}, \dots, e_m = \{i_m, j_m\}$.

LEMMA 6.1 (Conditional expectation of subgraphs). *Let H be a subgraph such that (2.2) has a unique maximum, attained at $S_3^{(\text{sub})} = [k]$ so that $B^{(\text{sub})}(H) = 0$. Then, on the event J_n defined in (3.1),*

$$(6.1) \quad \mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))] = \sum_{(v_1, \dots, v_k) \in W_n^k(\varepsilon)} \prod_{\{i,j\} \in \mathcal{E}_H} (1 - e^{-D_{v_i} D_{v_j}/L_n})(1 + o(1)).$$

PROOF. Let $\mathbf{v} = (v_1, \dots, v_k)$ and $\text{ECM}^{(n)}(\mathbf{D})|_{\mathbf{v}}$ again be the induced subgraph of $\text{ECM}^{(n)}(\mathbf{D})$ on \mathbf{v} . We first derive a more detailed expression for the probability that a subgraph is present on \mathbf{v} than (4.21) which holds when $\mathbf{v} \in W_n^k(\varepsilon)$. Because $\mathbf{v} \in W_n^k(\varepsilon)$, we may use (4.3) for all edge probabilities to obtain

$$(6.2) \quad \mathbb{P}_n(X_{v_{i_1}, v_{j_1}} = \dots = X_{v_{i_m}, v_{j_m}} = 0) = \prod_{l=1}^m \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = 0)(1 + O(n^{(\tau-2)/(\tau-1)})).$$

When $D_{v_i}, D_{v_j} \in [\varepsilon\sqrt{n}, \sqrt{n}/\varepsilon]$, $\mathbb{P}_n(X_{v_i, v_j} = 0) = \Theta(1)$ and $\mathbb{P}_n(X_{v_i, v_j} = 1) = \Theta(1)$. Therefore, similarly to (4.21), for $\mathbf{v} \in W_n^k(\varepsilon)$,

$$\begin{aligned}
 & \mathbb{P}_n(\text{ECM}^{(n)}(\mathbf{D})|_{\mathbf{v}} \supseteq \mathcal{E}_H) \\
 &= 1 - \sum_{l=1}^m \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = 0) + \sum_{l \neq p} \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = X_{v_{i_p}, v_{j_p}} = 0) \\
 (6.3) \quad & - \sum_{l \neq p \neq r} \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = X_{v_{j_p}, v_{i_p}} = X_{v_{i_r}, v_{j_r}} = 0) + \dots \\
 & + (-1)^m \mathbb{P}_n(X_{v_{i_1}, v_{j_1}} = \dots = X_{v_{i_m}, v_{j_m}} = 0) \\
 &= (1 + o(1)) \prod_{l=1}^m (1 - \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = 0)).
 \end{aligned}$$

Thus, the conditional expected value satisfies

$$\begin{aligned}
 \mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))] &= \sum_{\mathbf{v} \in W_n^k(\varepsilon)} \mathbb{P}_n(\text{ECM}^{(n)}(\mathbf{D})|_{\mathbf{v}} \supseteq \mathcal{E}_H) \\
 (6.4) \quad &= (1 + o(1)) \sum_{\mathbf{v} \in W_n^k(\varepsilon)} \prod_{l=1}^m (1 - \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = 0)).
 \end{aligned}$$

Because $D_{v_i} D_{v_j} = O(n)$ and $L_n = \mu n(1 + o(1))$ under J_n , by (4.4),

$$(6.5) \quad \mathbb{P}_n(X_{v_i, v_j} = 1) = 1 - e^{-D_{v_i} D_{v_j} / L_n} + O\left(\frac{D_{v_i}^2 D_{v_j}}{L_n^2}\right) = (1 + o(1))(1 - e^{-D_{v_i} D_{v_j} / L_n}).$$

This results in

$$(6.6) \quad \mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))] = (1 + o(1)) \sum_{\mathbf{v} \in W_n^k(\varepsilon)} \prod_{\{i, j\} \in \mathcal{E}_H} (1 - e^{-D_{v_i} D_{v_j} / L_n}). \quad \square$$

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,*

$$\begin{aligned}
 (6.7) \quad \frac{\mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))]}{n^{\frac{k}{2}(3-\tau)}} &= (1 + o_{\mathbb{P}}(1)) c^k \mu^{-\frac{k}{2}(\tau-1)} \int_{\varepsilon}^{1/\varepsilon} \dots \int_{\varepsilon}^{1/\varepsilon} (x_1 \dots x_k)^{-\tau} \\
 &\times \prod_{\{i, j\} \in \mathcal{E}_H} (1 - e^{-x_i x_j}) dx_1 \dots dx_k + f_n(\varepsilon),
 \end{aligned}$$

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

$$(6.8) \quad \lim_{\varepsilon \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(f_n(\varepsilon) > \delta \mid J_n) = 0.$$

PROOF. Let $|\mathcal{E}_H| = m$ and denote the edges of H by $\{i_1, j_1\}, \dots, \{i_m, j_m\}$. Define

$$(6.9) \quad g(t_1, \dots, t_k) := \prod_{\{i, j\} \in \mathcal{E}_H} (1 - e^{-t_i t_j}).$$

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

$$(6.10) \quad 1 - e^{-xy} = \sum_{i=1}^s \frac{(xy)^i}{i!} (-1)^i + O\left(\frac{\varepsilon^{-s}}{(s+1)!}\right).$$

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

$$(6.11) \quad \begin{aligned} g(t_1, \dots, t_k) &= \sum_{p_1=1}^{s_1} \cdots \sum_{p_m=1}^{s_m} \left((-1)^{p_1} \frac{t_{u_1}^{p_1} t_{v_1}^{p_1}}{p_1!} \cdots (-1)^{p_m} \frac{t_{u_m}^{p_m} t_{v_m}^{p_m}}{p_m!} \right) + \eta(t_1, \dots, t_k) \\ &= \sum_{p_1=1}^{s_1} \cdots \sum_{p_m=1}^{s_m} \left(\frac{(-1)^{p_1+\dots+p_m}}{p_1! \cdots p_m!} t_1^{\gamma_1} t_2^{\gamma_2} \cdots t_k^{\gamma_k} \right) + \eta(t_1, \dots, t_k), \end{aligned}$$

where

$$(6.12) \quad \gamma_j := \gamma_j(p_1, \dots, p_m) = \sum_l p_l \mathbb{1}_{\{i_l=j \text{ or } j_l=j\}}.$$

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

$$(6.13) \quad M^{(n)}([a, b]) = (\mu n)^{\frac{1}{2}(\tau-1)} n^{-1} \sum_{v \in [n]} \mathbb{1}_{\{D_v \in \sqrt{\mu n}[a, b]\}}.$$

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

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

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

$$(6.15) \quad M^{(n)}([a, b]) \xrightarrow{\mathbb{P}} \lambda([a, b]).$$

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

$$(6.16) \quad \begin{aligned} &\frac{\mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))]}{n^{\frac{k}{2}(3-\tau)} \mu^{\frac{k}{2}(1-\tau)}} \\ &= \int_F g(t_1, \dots, t_k) dN^{(n)}(t_1, \dots, t_k) \\ &= \int_F \sum_{p_1=1}^{s_1} \cdots \sum_{p_m=1}^{s_m} \left(\left(\frac{(-1)^{p_1+\dots+p_m}}{p_1! \cdots p_m!} t_1^{\gamma_1} t_2^{\gamma_2} \cdots t_k^{\gamma_k} \right) + \eta(t_1, \dots, t_k) \right) dN^{(n)}(t_1, \dots, t_k) \\ &= \sum_{p_1=1}^{s_1} \cdots \sum_{p_m=1}^{s_m} \frac{(-1)^{p_1+\dots+p_m}}{p_1! \cdots p_m!} \int_\varepsilon^{1/\varepsilon} t_1^{\gamma_1} dM^{(n)}(t_1) \cdots \int_\varepsilon^{1/\varepsilon} t_k^{\gamma_k} dM^{(n)}(t_k) + f_n(\varepsilon). \end{aligned}$$

Here

$$\begin{aligned} f_n(\varepsilon) &= \int_F \sum_{p_1=1}^{s_1} \cdots \sum_{p_m=1}^{s_m} \eta(t_1, \dots, t_k) dN^{(n)}(t_1, \dots, t_k) \\ &\leq \int_F \sum_{p_1=1}^{s_1} \cdots \sum_{p_m=1}^{s_m} \varepsilon^{k(\tau-1)+1} dN^{(n)}(t_1, \dots, t_k) \end{aligned}$$

$$\begin{aligned}
 (6.17) \quad &= \varepsilon^{k(\tau-1)+1} s_1 \cdots s_m M^{(n)}([\varepsilon, 1/\varepsilon])^k \\
 &= \varepsilon^{k(\tau-1)+1} O_{\mathbb{P}}(\lambda([\varepsilon, 1/\varepsilon])^k) \\
 &= \varepsilon^{k(\tau-1)+1} (\varepsilon^{1-\tau} - \varepsilon^{\tau-1})^k O_{\mathbb{P}}(1) = O_{\mathbb{P}}(\varepsilon),
 \end{aligned}$$

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

$$(6.18) \quad \lim_{\varepsilon \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(f_n(\varepsilon) > \delta \mid J_n) = 0.$$

As in [30], equation (55), for any γ ,

$$(6.19) \quad \int_{\varepsilon}^{1/\varepsilon} x^{\gamma} dM^{(n)}(x) \xrightarrow{\mathbb{P}} \int_{\varepsilon}^{1/\varepsilon} x^{\gamma} d\lambda(x).$$

Combining this with (6.16) results in

$$\begin{aligned}
 (6.20) \quad & \frac{\mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))]}{n^{\frac{k}{2}(3-\tau)} \mu^{\frac{k}{2}(1-\tau)}} \\
 &= (1 + o_{\mathbb{P}}(1)) \sum_{p_1=1}^{s_1} \cdots \sum_{p_m=1}^{s_m} \frac{(-1)^{p_1+\cdots+p_m}}{p_1! \cdots p_m!} \\
 & \times \int_{\varepsilon}^{1/\varepsilon} t_1^{\alpha_1} d\lambda(t_1) \cdots \int_{\varepsilon}^{1/\varepsilon} t_k^{\alpha_k} d\lambda(t_k) + f_n(\varepsilon) \\
 &= (1 + o_{\mathbb{P}}(1)) \int_F \sum_{p_1=1}^{s_1} \cdots \sum_{p_m=1}^{s_m} \frac{(-1)^{p_1+\cdots+p_m}}{p_1! \cdots p_m!} t_1^{\alpha_1} \cdots t_k^{\alpha_k} d\lambda(t_1) \cdots d\lambda(t_k) + f_n(\varepsilon) \\
 &= (1 + o_{\mathbb{P}}(1)) \int_F g(t_1, \dots, t_k) d\lambda(t_1) \cdots d\lambda(t_k) + f_n(\varepsilon).
 \end{aligned}$$

Then, by (6.15),

$$\begin{aligned}
 (6.21) \quad & \frac{\mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))]}{n^{\frac{k}{2}(3-\tau)}} = (1 + o_{\mathbb{P}}(1)) c^k \mu^{-\frac{k}{2}(\tau-1)} \int_{\varepsilon}^{1/\varepsilon} \cdots \int_{\varepsilon}^{1/\varepsilon} (t_1 \cdots t_k)^{-\tau} \\
 & \times g(t_1, \dots, t_k) dt_1 \cdots dt_k + f_n(\varepsilon),
 \end{aligned}$$

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 defined in (3.1),*

$$(6.22) \quad \frac{\text{Var}_n(N^{(\text{sub})}(H, W_n^k(\varepsilon)))}{\mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))]^2} \xrightarrow{\mathbb{P}} 0.$$

PROOF. By Lemma 6.2,

$$(6.23) \quad \mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))]^2 = \Theta_{\mathbb{P}}(n^{(3-\tau)k}).$$

Thus, we need to prove that the variance is small compared to $n^{(3-\tau)k}$. Denote $\mathbf{v} = (v_1, \dots, v_k)$ and $\mathbf{u} = (u_1, \dots, u_k)$ and, for ease of notation, we denote $G = \text{ECM}^{(n)}(\mathbf{D})$. We write the variance as

$$(6.24) \quad \text{Var}_n(N^{(\text{sub})}(H, W_n^k(\varepsilon))) = \sum_{\mathbf{v} \in W_n^k(\varepsilon)} \sum_{\mathbf{u} \in W_n^k(\varepsilon)} (\mathbb{P}_n(G|\mathbf{v} \supseteq \mathcal{E}_H, G|\mathbf{u} \supseteq \mathcal{E}_H) - \mathbb{P}_n(G|\mathbf{v} \supseteq \mathcal{E}_H)\mathbb{P}_n(G|\mathbf{u} \supseteq \mathcal{E}_H)).$$

This splits into various cases, depending on the overlap of \mathbf{v} and \mathbf{u} . When \mathbf{v} and \mathbf{u} do not overlap, similar to (6.3),

$$(6.25) \quad \begin{aligned} & \sum_{\mathbf{v} \in W_n^k(\varepsilon)} \sum_{\mathbf{u} \in W_n^k(\varepsilon)} (\mathbb{P}_n(G|\mathbf{v} \supseteq \mathcal{E}_H, G|\mathbf{u} \supseteq \mathcal{E}_H) - \mathbb{P}_n(G|\mathbf{v} \supseteq \mathcal{E}_H)\mathbb{P}_n(G|\mathbf{u} \supseteq \mathcal{E}_H)) \\ &= \sum_{\mathbf{v} \in W_n^k(\varepsilon)} \sum_{\mathbf{u} \in W_n^k(\varepsilon)} \left((1 + o(1)) \prod_{l=1}^m (1 - \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = 0))(1 - \mathbb{P}_n(X_{u_{i_l}, u_{j_l}} = 0)) \right. \\ & \quad \left. - (1 + o(1)) \prod_{l=1}^m (1 - \mathbb{P}_n(X_{v_{i_l}, v_{j_l}} = 0))(1 - \mathbb{P}_n(X_{u_{i_l}, u_{j_l}} = 0)) \right) \\ &= \mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))]^2 o(1). \end{aligned}$$

The other contributions are when \mathbf{v} and \mathbf{u} overlap. We then denote by $|\mathbf{v} \cup \mathbf{u}|$ the number of distinct vertices in \mathbf{v} and \mathbf{u} . In this situation, we use the bound $\mathbb{P}_n(X_{u,v} = 1) \leq 1$. When \mathbf{v} and \mathbf{u} overlap on $s \geq 1$ vertices, we bound the contribution to (6.24) as

$$(6.26) \quad \begin{aligned} \sum_{\mathbf{v}, \mathbf{u} \in W_n^k(\varepsilon) : |\mathbf{v} \cup \mathbf{u}| = 2k-s} \mathbb{P}_n(G|\mathbf{v} \supseteq \mathcal{E}_H, G|\mathbf{u} \supseteq \mathcal{E}_H) &\leq |\{i : D_i \in \sqrt{\mu n}[\varepsilon, 1/\varepsilon]\}|^{2k-s} \\ &= O_{\mathbb{P}}(n^{\frac{(3-\tau)(2k-s)}{2}}), \end{aligned}$$

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

$$(6.27) \quad N^{(\text{sub})}(H, W_n^k(\varepsilon)) = \mathbb{E}_n[N^{(\text{sub})}(H, W_n^k(\varepsilon))](1 + o_{\mathbb{P}}(1)).$$

Combining this with Lemma 6.2 proves Lemma 5.1(i). Lemma 5.1(ii) is a direct consequence of Lemma 7.2 in the next section, when we take $|S_3^*| = k$. Here, we remark that the proof of Lemma 7.2 is entirely self-contained, so no circularity in our argument arises. \square

7. Major contribution to general subgraphs: Proof of Theorem 2.1(i). In this section we prove Theorem 2.1(i) for subgraphs. We start by giving an overview of the proof. We restrict this overview to the expected value of $N^{(\text{sub})}(H, M_n^{(\alpha^{(\text{sub})})}(\varepsilon_n))$.

To compute $\mathbb{E}[N^{(\text{sub})}(H, M_n^{(\alpha^{(\text{sub})})}(\varepsilon_n))]$, we need to count the expected number of copies of H on vertices $v_1, \dots, v_k \in [n]$ for which $D_{v_i} \approx n^{\alpha_i^{(\text{sub})}}$. This means that we sum the probabilities that $D_{v_i} = k_i$ for all $i \in [k]$ over all k_i that are of the order $n^{\alpha_i^{(\text{sub})}}$, multiplied by the probabilities that $\{v_i, v_j\}$ is an edge in $\text{CM}_n(\mathbf{D})$ for all $\{i, j\} \in \mathcal{E}_H$, conditionally on the degrees. We rescale the arising sum over k_i , and instead integrate over $x_i = k_i n^{-\alpha_i^{(\text{sub})}}$.

After rescaling, we are left with a k -fold integral over the variables x_i for all $i \in [k]$ of which we aim to show that it satisfies the appropriate bounds. Here, we will crucially rely

on the uniqueness of the optimization problem in (2.2). While we are “merely” left with a k -fold integral over relatively simple functions, due to the somewhat implicit information that (2.2) has a unique solution, proving the finiteness of the integral is quite challenging. Indeed, we will need to resort to comparisons over different partitions, and use that they provide a smaller value of the functional in (2.2) to establish the finiteness of the integral.

Let us give some more details about the nature of the integral, and on the organisation of the proof. To simplify notation, we write $S_i^* = S_i^{(\text{sub})}$ for $i = 1, 2, 3$ for the optimal partition \mathcal{P} in (2.2). Recall that, by Lemma 6.1 and on the event J_n ,

$$(7.1) \quad \mathbb{P}_n(\{v_i, v_j\} \text{ is an edge in } \text{CM}_n(\mathbf{D})) \leq \min(D_{v_i} D_{v_j} / (\mu n), 1),$$

and these events are close to being independent for different edges. We bound the minimum in (7.1) by $D_{v_i} D_{v_j} / (\mu n)$ 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 minimum by 1 for $i, j \in S_2^*$ and $i \in S_2^*, j \in S_3^*$ or vice versa. This means that in the integral over the rescaled variables x_i , a factor $x_i^{\zeta_i}$ appears for a certain $\zeta_i \geq 0$ (for a precise definition of ζ_i , see (7.2)). Further, the nice aspect of this bound is that the integrals over x_i for $i \in [k]$ factorize into integrals over x_i for $i \in S_3^*$ and $i \in S_1^* \cup S_2^*$. This allows us to study these integrals *separately*. It turns out that our proof of the finiteness of these integrals depends sensitively on the optimization problem in (2.2) having a *unique* solution. In turn, this explains why some of these integrals are quite hard to bound, as the only ingredient we have is that the optimization problem in (2.2) has a unique solution.

The remainder of the proof is now organized as follows. In Lemma 7.1, we derive bounds on the additional powers ζ_i of x_i in the rescaled integral, which will prove crucial in bounding the arising integrals. In Lemma 7.2, we derive a bound on the integrals over x_i for $i \in S_3^*$, and in Lemma 7.3, we bound the integrals over x_i for $i \in S_1^* \cup S_2^*$. After stating Lemmas 7.2 and 7.3, we complete the proof of Theorem 2.1(i). Then we give the (rather involved) proofs of Lemmas 7.2 and 7.3.

Before giving the details of the argument, we introduce some further 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 $\mathcal{P} = (S_1^*, S_2^*, S_3^*)$ be the optimal partition. Define

$$(7.2) \quad \zeta_i = \begin{cases} 1 & \text{if } d_i = 1, \\ d_{i,S_1^*} + d_{i,S_3^*} + d_{i,V_1} & \text{if } i \in S_1^*, \\ d_{i,V_1} & \text{if } i \in S_2^*, \\ d_{i,S_1^*} + d_{i,V_1} & \text{if } i \in S_3^*. \end{cases}$$

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

LEMMA 7.1 (Bounds on the additional powers of rescaled variables). *Let H be a connected subgraph, such that the optimum of (2.2) is unique, and let $\mathcal{P} = (S_1^*, S_2^*, S_3^*)$ be the optimal partition. Then:*

- (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 first that $i \in S_1^*$. Now consider the partition $\hat{S}_1 = S_1^* \setminus \{i\}$, $\hat{S}_2 = S_2^*$, $S_3 = S_3^* \cup \{i\}$. Then, $E_{\hat{S}_1} = E_{S_1^*} - d_{i,S_1^*}$ and $E_{\hat{S}_1, \hat{S}_3} = E_{S_1^*, S_3^*} + d_{i,S_1^*} - d_{i,S_3^*}$. Furthermore, $E_{\hat{S}_1, V_1} =$

$E_{S_1^*, V_1} - d_{i, V_1}$ and $E_{\hat{S}_2, V_1} = E_{S_2^*, V_1}$. Because the partition into S_1^* , S_2^* and S_3^* achieves the unique optimum of (2.2),

$$(7.3) \quad \begin{aligned} |S_1^*| - |S_2^*| - \frac{2E_{S_1^*} - E_{S_1^*, S_3^*} + E_{S_2^*, V_1} - E_{S_1^*, V_1}}{\tau - 1} \\ > |S_1^*| - 1 - |S_2^*| - \frac{2E_{S_1^*} - E_{S_1^*, S_3^*} - d_{i, S_1^*} - d_{i, S_3^*} + E_{S_2^*, V_1} - E_{S_1^*, V_1} + d_{i, V_1}}{\tau - 1}, \end{aligned}$$

which reduces to

$$(7.4) \quad d_{i, S_1^*} + d_{i, S_3^*} + d_{i, V_1} = \zeta_i < \tau - 1.$$

Using that $\tau \in (2, 3)$ then yields $d_{i, S_1^*} + d_{i, S_3^*} + d_{i, V_1} \leq 1$.

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 $d_{i, S_3^*} + d_{i, S_1^*} + d_{i, V_1} \geq 2$, and considering the partition where i is moved to S_2^* results in the inequality $d_{i, S_1^*} + d_{i, V_1} \leq 1$, so that $\zeta_i \leq 1$. □

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 (Subgraph integrals over S_3^*). *Suppose that the maximum in (2.2) is uniquely attained by $\mathcal{P} = (S_1^*, S_2^*, S_3^*)$ with $|S_3^*| = s > 0$, and say $S_3^* = [s]$. Then*

$$(7.5) \quad \int_0^\infty \cdots \int_0^\infty \prod_{i \in [s]} x_i^{-\tau + \zeta_i} \prod_{\{i, j\} \in \mathcal{E}_{S_3^*}} \min(x_i x_j, 1) dx_s \cdots dx_1 < \infty.$$

The proof of Lemma 7.2 is deferred to after the proof of Theorem 2.1(i). We continue with the integrals over $S_1^* \cup S_2^*$.

LEMMA 7.3 (Subgraph integrals over $S_1^* \cup S_2^*$). *Suppose the optimal solution to (2.2) is unique, and attained by $\mathcal{P} = (S_1^*, S_2^*, S_3^*)$. Say that $S_2^* = [t_2]$ and $S_1^* = [t_2 + t_1] \setminus [t_2]$. Then, for every $a > 0$,*

$$(7.6) \quad \int_0^a \cdots \int_0^a \int_0^\infty \cdots \int_0^\infty \prod_{j \in [t_1 + t_2]} x_j^{-\tau + \zeta_j} \prod_{\{i, j\} \in \mathcal{E}_{S_1^*, S_2^*}} \min(x_i x_j, 1) dx_{t_1 + t_2} \cdots dx_1 < \infty.$$

The proof of Lemma 7.3 is deferred to after the proof of Theorem 2.1(i). Now we are ready to complete the proof of Theorem 2.1(i) for subgraphs.

PROOF OF THEOREM 2.1(i). Because $D_{\max} = O_{\mathbb{P}}(n^{1/(\tau-1)})$, for any $\eta_n \rightarrow 0$, $D_{\max} \leq n^{1/(\tau-1)}/\eta_n$ with high probability. Define

$$(7.7) \quad \gamma_i^u(n) = \begin{cases} n^{1/(\tau-1)}/\eta_n & \text{if } i \in S_2^*, \\ n^{\alpha_i^{(\text{sub})}}/\varepsilon_n & \text{else,} \end{cases}$$

with $\alpha_i^{(\text{sub})}$ as in (2.5), and denote

$$(7.8) \quad \gamma_i^l(n) = \begin{cases} 1 & \text{if } i \in V_1, \\ \varepsilon_n n^{\alpha_i^{(\text{sub})}} & \text{else.} \end{cases}$$

We then show that the expected number of subgraphs where the degree of at least one vertex i satisfies $D_i \notin [\gamma_i^l(n), \gamma_i^u(n)]$ is small, similar to the proof of Theorem 2.2 in Section 5.

With loss of generality, we assume that vertex $1 \in V_H$ satisfies $1 \in V_1$. We count the expected number of $\mathbf{v} = (v_1, \dots, v_k)$ for which the edge $\{v_i, v_j\}$ is present in $\text{CM}_n(\mathbf{D})$ for every $\{i, j\} \in \mathcal{E}_H$. We first study the expected number of copies of H where vertex V_1 has degree $D_{v_1} \in [1, \gamma_1^l(n))$ and all other vertices satisfy $D_{v_i} \in [\gamma_i^l(n), \gamma_i^u(n)]$, by integrating the probability that subgraph H is formed over the range where vertex v_1 has degree $D_{v_1} \in [1, \gamma_1^l(n))$ and all other vertices satisfy $D_{v_i} \in [\gamma_i^l(n), \gamma_i^u(n)]$. Using that the connection probabilities can be bounded by $M_1 \min(D_{v_i} D_{v_j} / n, 1)$ for some $M_1 > 0$ (recall Lemma 6.1, and in particular (7.1)), and the degree distribution can be bounded as $\mathbb{P}(D = k) \leq M_2 k^{-\tau}$ for some $M_2 > 0$ by (1.1), we bound the expected number of such copies of H by

$$(7.9) \quad \sum_{\mathbf{v}} \mathbb{E}[I^{(\text{sub})}(H, \mathbf{v}) \mathbb{1}_{\{D_{v_1} < \gamma_1^l(n), D_{v_i} \in [\gamma_i^l(n), \gamma_i^u(n)] \forall i > 1\}}]$$

$$\leq K n^k \int_1^{\gamma_1^l(n)} \int_{\gamma_2^l(n)}^{\gamma_2^u(n)} \dots \int_{\gamma_k^l(n)}^{\gamma_k^u(n)} (x_1 \dots x_k)^{-\tau} \prod_{\{i, j\} \in \mathcal{E}_H} \min\left(\frac{x_i x_j}{n}, 1\right) dx_k \dots dx_1,$$

for some $K > 0$, and where we recall that $I^{(\text{sub})}(H, \mathbf{v}) = \mathbb{1}_{\{\text{ECM}^{(n)}(\mathbf{D})|_{\mathbf{v}} \supseteq \mathcal{E}_H\}}$. This integral equals zero when vertex 1 is in V_1 , since then $[1, \gamma_1^l(n)) = \emptyset$. Suppose that vertex 1 is in S_2^* . Without loss of generality, assume that $S_2^* = [t_2]$, $S_1^* = [t_1 + t_2] \setminus [t_2]$ and $S_3^* = [t_1 + t_2 + t_3] \setminus [t_1 + t_2]$. We bound the minimum in (7.9) by:

- (a) $x_i x_j / n$ for $i, j \in S_1^*$;
- (b) $x_i x_j / n$ for i or j in V_1 ;
- (c) $x_i x_j / n$ for $i \in S_1^*, j \in S_3^*$ or vice versa; and
- (d) 1 for $i, j \in S_2^*$ and $i \in S_2^*, j \in S_3^*$ or vice versa.

Applying the change of variables $y_i = x_i / n^{\alpha_i^{(\text{sub})}}$ results, for some $\tilde{K} > 0$, in the bound

$$\begin{aligned} & \sum_{\mathbf{v}} \mathbb{E}[I^{(\text{sub})}(H, \mathbf{v}) \mathbb{1}_{\{D_{v_1} < \gamma_1^l(n), D_{v_i} \in [\gamma_i^l(n), \gamma_i^u(n)] \forall i > 1\}}] \\ & \leq \tilde{K} n^{|S_1^*|(2-\tau) + |S_3^*|(1-\tau)/2 - |S_2^*|} \\ & \quad \times n^k n^{\frac{\tau-3}{\tau-1} E_{S_1^*, V_1} + \frac{\tau-3}{2(\tau-1)} E_{S_1^*, S_3^*} - \frac{1}{\tau-1} E_{S_1^*, V_1} - \frac{1}{2} E_{S_3^*, V_1} - \frac{\tau-2}{\tau-1} E_{S_2^*, V_1}} \\ & \quad \times \int_0^{\varepsilon_n} \int_0^{1/\eta_n} \dots \int_0^{1/\eta_n} \int_0^\infty \dots \int_0^\infty \prod_{i \in V_H \setminus V_1} y_i^{-\tau + \zeta_i} \\ & \quad \times \prod_{\{i, j\} \in \mathcal{E}_{S_3^*} \cup E_{S_1^*, S_2^*}} \min(y_i y_j, 1) dy_{t_1+t_2+t_3} \dots dy_1 \prod_{j \in V_1} \int_1^\infty y_j^{1-\tau} dy_j, \end{aligned}$$

where the integrals from 0 to $1/\eta_n$ correspond to vertices in S_2^* and the integrals from 0 to ∞ to vertices in S_1^* and S_3^* . Since $\tau \in (2, 3)$, the integrals corresponding to vertices in V_1 are finite. By the analysis from (4.37) to (4.40),

$$(7.10) \quad \begin{aligned} & |S_1^*|(2-\tau) + |S_3^*|(1-\tau)/2 - |S_2^*| + k + \frac{\tau-3}{\tau-1} E_{S_1^*} + \frac{\tau-3}{2(\tau-1)} E_{S_1^*, S_3^*} \\ & - \frac{1}{\tau-1} E_{S_1^*, V_1} - \frac{1}{2} E_{S_3^*, V_1} - \frac{\tau-2}{\tau-1} E_{S_2^*, V_1} \\ & = \frac{3-\tau}{2} (k_{2+} + B^{(\text{sub})}(H)) + k_1/2. \end{aligned}$$

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

$$\begin{aligned}
 & \int_0^{\varepsilon_n} \int_0^{1/\eta_n} \cdots \int_0^{1/\eta_n} \int_0^\infty \cdots \int_0^\infty \prod_{i \in S_1^* \cup S_2^*} y_i^{-\tau + \zeta_i} \prod_{\{i,j\} \in \mathcal{E}_{S_1^*, S_2^*}} \min(y_i y_j, 1) \, dy_{t_1+t_2} \cdots dy_1 \\
 (7.11) \quad & \times \int_0^\infty \cdots \int_0^\infty \prod_{i \in S_3^*} y_i^{-\tau + \zeta_i} \prod_{\{i,j\} \in \mathcal{E}_{S_3^*}} \min(y_i y_j, 1) \, dy_{t_1+t_2+t_3} \cdots dy_{t_1+t_2+1}.
 \end{aligned}$$

By Lemma 7.2 the set of integrals on the second line of (7.11) is finite. Lemma 7.3 shows that the set of integrals on the first line of (7.11) tends to zero for η_n fixed and $\varepsilon_n \rightarrow 0$. Thus, choosing $\eta_n \rightarrow 0$ sufficiently slowly compared to ε_n yields

$$\begin{aligned}
 (7.12) \quad & \int_0^{\varepsilon_n} \int_0^{1/\eta_n} \cdots \int_0^{1/\eta_n} \int_0^\infty \cdots \int_0^\infty \prod_{i \in S_1^* \cup S_2^*} y_i^{-\tau + \zeta_i} \prod_{\{i,j\} \in \mathcal{E}_{S_1^*, S_2^*}} \min(y_i y_j, 1) \, dy_{t_1+t_2} \cdots dy_1 \\
 & = o(1).
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 (7.13) \quad & \sum_v \mathbb{E}[I^{(\text{sub})}(H, \mathbf{v}) \mathbb{1}_{\{D_{v_1} < \gamma_1^l(n), D_{v_i} \in [\gamma_i^l(n), \gamma_i^u(n)] \forall i > 1\}}] \\
 & = o(n^{\frac{3-\tau}{2}(k_{2+} + B^{(\text{sub})}(H)) + k_1/2}),
 \end{aligned}$$

when vertex 1 satisfies $1 \in S_2^*$. Similarly, we can show that the expected contribution from $D_{v_1} < \gamma_1^l(n)$ satisfies the same bound when vertex 1 is in S_1^* or S_3^* . The expected number of subgraphs where $D_{v_1} > \gamma_1^u(n)$ 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 $D_{v_i} \notin [\gamma_i^l(n), \gamma_i^u(n)]$.

Denote

$$(7.14) \quad \Gamma_n(\varepsilon_n, \eta_n) = \{(v_1, \dots, v_k) : D_{v_i} \in [\gamma_{v_i}^l(n), \gamma_{v_i}^u(n)]\},$$

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 $N^{(\text{sub})}(H, \bar{\Gamma}_n(\varepsilon_n, \eta_n))$. Since $D_{\max} \leq n^{1/(\tau-1)}/\eta_n$ with high probability, $\Gamma_n(\varepsilon_n, \eta_n) = M_n^{(\alpha^{(\text{sub})})}$ with high probability. Therefore, with high probability,

$$(7.15) \quad N^{(\text{sub})}(H, \bar{M}_n^{(\alpha^{(\text{sub})})}(\varepsilon_n)) = N^{(\text{sub})}(H, \bar{\Gamma}_n(\varepsilon_n, \eta_n)),$$

where $N^{(\text{sub})}(H, \bar{M}_n^{(\alpha^{(\text{sub})})}(\varepsilon_n))$ denotes the number of copies of H on vertices not in $M_n^{(\alpha^{(\text{sub})})}(\varepsilon_n)$. By the Markov inequality,

$$(7.16) \quad N^{(\text{sub})}(H, \bar{\Gamma}_n(\varepsilon_n, \eta_n)) = o_{\mathbb{P}}(n^{\frac{3-\tau}{2}(k_{2+} + B^{(\text{sub})}(H)) + k_1/2}).$$

Combining this with the fact that by Theorem 2.1(ii) as proved in Section 4, for fixed $\varepsilon > 0$,

$$\begin{aligned}
 (7.17) \quad & N^{(\text{sub})}(H) = N^{(\text{sub})}(H, M_n^{(\alpha^{(\text{sub})})}(\varepsilon)) + N^{(\text{sub})}(H, \bar{M}_n^{(\alpha^{(\text{sub})})}(\varepsilon)) \\
 & = O_{\mathbb{P}}(n^{\frac{3-\tau}{2}(k_{2+} + B^{(\text{sub})}(H)) + k_1/2})
 \end{aligned}$$

shows that

$$(7.18) \quad \frac{N^{(\text{sub})}(H, M_n^{(\alpha^{(\text{sub})})}(\varepsilon_n))}{N^{(\text{sub})}(H)} \xrightarrow{\mathbb{P}} 1,$$

as required. This completes the proof of Theorem 2.1(i). \square

We close this section by proving the integral Lemmas 7.2 and 7.3.

PROOF OF LEMMA 7.2. Recall that $S_3^* = [s]$. Without loss of generality, we may assume that $x_1 < x_2 < \dots < x_s$. Let $U = [t]$ be such that $x_i < 1$ precisely when $i \in [t]$. Here $U = \emptyset$ when $t = 0$.

The integral (7.5) consists of multiple regions that will be characterized by U , and we will deal with all of them in the sequel. The first region is where $U = \emptyset$, so that $x_1, \dots, x_s \geq 1$. Since $-\tau + \zeta_i < -1$ by Lemma 7.1(iii), this integral can be bounded by the full integral over $[1, \infty)$ for all variables, which is bounded by

$$(7.19) \quad \int_1^\infty \dots \int_1^\infty \prod_{j \in [s]} x_j^{-\tau + \zeta_j} dx_1 \dots dx_s < \infty.$$

The second region is where $U = [s]$, so that $x_1, \dots, x_s \in [0, 1]$. Since by Lemma 7.1, any vertex in S_3^* satisfies $\zeta_i + d_{i, S_3^*} \geq 2$, this integral can be bounded as

$$(7.20) \quad \begin{aligned} & \int_0^1 \dots \int_0^1 \prod_{j \in [s]} x_j^{-\tau + \zeta_j} \prod_{\{i, j\} \in \mathcal{E}_{S_3^*}} x_i x_j dx_1 \dots dx_s \\ &= \int_0^1 \dots \int_0^1 \prod_{j \in [s]} x_j^{-\tau + \zeta_j + d_{j, S_3^*}} dx_1 \dots dx_s \\ &\leq \int_0^1 \dots \int_0^1 (x_1 \dots x_s)^{2-\tau} dx_1 \dots dx_s < \infty. \end{aligned}$$

The other regions arise when $U \neq \emptyset$ and $U \neq [t]$. For these cases, the integral runs from 1 to ∞ for $i \in U$, and from 0 to 1 for $i \in \bar{U} = S_3^* \setminus U$. In such a region, $\min(x_i x_j, 1) = x_i x_j$ when $i, j \notin U$, and $\min(x_i x_j, 1) = 1$ when $i, j \in U$. Then, as we assumed that $x_1 < x_2 < \dots < x_s$, the contribution to (7.5) from the region described by U can be bounded by

$$(7.21) \quad \int_1^\infty \int_{x_1}^\infty \dots \int_{x_{t-1}}^\infty \prod_{j \in [t]} x_j^{-\tau + \zeta_j} \prod_{i=t+1}^s h(i, \mathbf{x}) dx_t \dots dx_1,$$

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

$$(7.22) \quad h(i, \mathbf{x}) = \int_0^1 x_i^{-\tau + \zeta_i + d_{i, \bar{U}}} \prod_{j \in U: \{i, j\} \in \mathcal{E}_{S_3^*}} \min(x_i x_j, 1) dx_i,$$

for $i \in U = [s] \setminus [t]$.

The integral in $h(i, \mathbf{x})$ consists of multiple regions, depending on whether $x_i x_j < 1$ or not. Suppose vertex $i \in \bar{U}$ is connected in H to vertices $j_1, j_2, \dots, j_l \in U$, where $j_1 < j_2 < \dots < j_l$ so that also $1 < x_{j_1} < x_{j_2} < \dots < x_{j_l}$ and $l + d_{i, \bar{U}} = d_{i, S_3^*}$. Then, by splitting the integral depending on how many j_i 's are such that $x_i x_{j_i} \leq 1$, we obtain

$$(7.23) \quad \begin{aligned} h(i, \mathbf{x}) &= \int_0^1 x_i^{-\tau + \zeta_i + d_{i, \bar{U}}} \min(x_i x_{j_1}, 1) \min(x_i x_{j_2}, 1) \dots \min(x_i x_{j_l}, 1) dx_i \\ &= \int_{1/x_{j_1}}^1 x_i^{-\tau + \zeta_i + d_{i, \bar{U}}} dx_i + \dots + x_{j_1} \dots x_{j_{l-1}} \int_{1/x_{j_l}}^{1/x_{j_{l-1}}} x_i^{-\tau + \zeta_i + l - 1 + d_{i, \bar{U}}} dx_i \\ &\quad + x_{j_1} \dots x_{j_l} \int_0^{1/x_{j_l}} x_i^{-\tau + \zeta_i + l + d_{i, \bar{U}}} dx_i. \end{aligned}$$

Since $\zeta_i + d_{i, \bar{U}} + l - \tau = \zeta_i + d_{i, S_3^*} - \tau > -1$ by Lemma 7.1(iii), the last integral is finite.

Computing these integrals yields

$$\begin{aligned}
 h(i, \mathbf{x}) &= C_0 + C_1 x_{j_1}^{\tau - \zeta_i - d_{i, \bar{U}} - 1} + \dots + C_{l-1} x_{j_1} \dots x_{j_{l-2}} x_{j_{l-1}}^{\tau - \zeta_i - l - d_{i, \bar{U}} + 1} \\
 (7.24) \quad &+ C_l x_{j_1} x_{j_2} \dots x_{j_{l-1}} x_{j_l}^{\tau - \zeta_i - l - d_{i, \bar{U}}} \\
 &=: C_0 h_0(i, \mathbf{x}) + C_1 h_1(i, \mathbf{x}) + \dots + C_l h_l(i, \mathbf{x}),
 \end{aligned}$$

for some constants C_0, \dots, C_l . These terms (except for the first term) are all products of powers of x_{j_1}, \dots, x_{j_l} , such that the sum of these powers is $\tau - \zeta_i - d_{i, \bar{U}} - 1$. Furthermore, the exponents of x_{j_1}, \dots, x_{j_b} equal 1 for some $b \in [l]$, and the exponents of $x_{j_{b+2}}, \dots, x_{j_l}$ equal zero, and

$$(7.25) \quad \frac{h_p(i, \mathbf{x})}{h_{p-1}(i, \mathbf{x})} = x_{j_{p-1}} \frac{x_{j_p}^{\tau - \zeta_i - p}}{x_{j_{p-1}}^{\tau - \zeta_i - (p-1)}} = \left(\frac{x_{j_p}}{x_{j_{p-1}}} \right)^{\tau - \zeta_i - p},$$

which is at most 1 for $p \leq \tau - \zeta_i$, and smaller than 1 for $p > \tau - \zeta_i$. Thus, $p^* = p_i^* = \operatorname{argmax}_p h_p(i, \mathbf{x}) = \lfloor \tau - \zeta_i \rfloor$. Therefore, there exists a $K > 0$ such that

$$(7.26) \quad h(i, \mathbf{x}) \leq K h_{p_i^*}(i, \mathbf{x}).$$

In particular, $p_i^* = \lfloor \tau - \zeta_i \rfloor \geq 1$ by Lemma 7.1(iii).

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

$$\begin{aligned}
 (7.27) \quad &\int_1^\infty \int_{x_1}^\infty \dots \int_{x_{t-1}}^\infty \prod_{j \in [t]} x_j^{-\tau + \zeta_j} \prod_{i=t+1}^s h(i, \mathbf{x}) dx_t \dots dx_1 \\
 &\leq \tilde{K} \int_1^\infty \int_{x_1}^\infty \dots \int_{x_{t-1}}^\infty \prod_{j \in [t]} x_j^{-\tau + \zeta_j} \prod_{i=t+1}^s h_{p_i^*}(i, \mathbf{x}) dx_t \dots dx_1.
 \end{aligned}$$

The above steps effectively perform the integrals over x_i for $i \in \bar{U}$, and we are left with the integrals over x_j for $j \in U$. It is here that we will rely on the fact that the optimization problem in (2.2) has a *unique* solution. We start by rewriting the integral in (7.27) so that we can effectively use the uniqueness of (2.2), for which we need to make the dependence on the various x_j for $j \in U$ explicit. We start by introducing some notation to simplify this analysis.

Let $T_i = \{j_q : q \in [p_i^*]\} \subseteq U$ denote the set of neighbors of i that appear in $h_{p_i^*}(i, \mathbf{x})$. For all $j \in U$, let

$$(7.28) \quad Q_j = \{i \in \bar{U} : \{i, j\} \in \mathcal{E}_{S_3^*}, j_{p_i^*} \geq j\}$$

denote the set of neighbors $i \in \bar{U}$ of $j \in U$ such that x_j appears in $h_{p_i^*}(i, \mathbf{x})$ (note that $i < j$ for all $i \in \bar{U}, j \in U$). Then,

$$\begin{aligned}
 (7.29) \quad &\prod_{j=1}^t x_j^{-\tau + \zeta_j} \prod_{i=t+1}^s h_{p_i^*}(i, \mathbf{x}) \\
 &= \tilde{K} \prod_{j=1}^t x_j^{-\tau + \zeta_j + |Q_j|} \prod_{i=t+1}^s x_{j_{p_i^*}}^{\tau - 1 - \zeta_i - d_{i, \bar{U}} - p_i^*},
 \end{aligned}$$

for some constant $\tilde{K} > 0$. We now simplify the above integral.

Let $W_j = \{i \in \bar{U} : x_{j_{p_i^*}} = j\}$ for $j \in [t]$, so that W_j denotes the set of neighbors i of j in \bar{U} such that the factor $x_j^{\tau - \zeta_j - p_i^* - d_{j, \bar{U}}}$ appears in $h_{j_{p_i^*}}(i, \mathbf{x})$. Furthermore, let $\hat{W}_j = (V_1 \cup S_1^* \cup$

$[j] \cup \bar{U} \setminus W_j$. Then, by (7.2) and the fact that $p_i^* = d_{i, [j]_{p_i^*}}$,

$$(7.30) \quad \sum_{i \in W_j} \zeta_i + d_{i, \bar{U}} + p_i^* = \sum_{i \in W_j} d_{i, V_1} + d_{i, S_1^*} + d_{i, \bar{U}} + p_i^* = 2E_{W_j} + E_{W_j, \hat{W}_j},$$

where

$$(7.31) \quad E_{W_j} = |\{\{i, j\} \in \mathcal{E}_{S_3^*} : i, j \in W_j\}|$$

denotes the number of edges inside W_j and E_{W_j, \hat{W}_j} denotes the number of edges in S_3^* between W_j and \hat{W}_j . As a result, (7.27) becomes

$$(7.32) \quad \tilde{K} \int_1^\infty \int_{x_1}^\infty \cdots \int_{x_{t-1}}^\infty \prod_{j=1}^t x_j^{-\tau + \zeta_j + |Q_j| + (\tau-1)|W_j| - 2E_{W_j} - E_{W_j, \hat{W}_j}} dx_t \cdots dx_1.$$

We aim to perform the integrals one by one, starting with the integral over x_t , followed by x_{t-1} , etc. For this, we crucially use the uniqueness of (2.2) to show that

$$(7.33) \quad -\tau + \zeta_t + |Q_t| + (\tau-1)|W_t| - 2E_{W_t} - E_{W_t, \hat{W}_t} < -1,$$

so that the integral in (7.32) over x_t is finite. Indeed, note that

$$(7.34) \quad Q_t = \{i \in \bar{U} : \{i, t\} \in \mathcal{E}_{S_3^*}, j_{p_i^*} = t\} = \{i \in W_t : \{i, t\} \in \mathcal{E}_{S_3^*}\},$$

because t is the maximal index in U , so that $|Q_t| = d_{t, W_t}$. Also, $\hat{W}_t = (V_1 \cup S_1^* \cup S_3^*) \setminus W_t$ because $[t] \cup \bar{U} = S_3^*$.

Define $\hat{S}_2 = \hat{S}_2^* \cup \{t\}$, $\hat{S}_1 = \hat{S}_1^* \cup W_t$ and $\hat{S}_3 = S_3^* \setminus (W_t \cup \{t\})$. This gives

$$(7.35) \quad E_{\hat{S}_1} - E_{S_1^*} = E_{W_t} + E_{W_t, S_1^*},$$

$$(7.36) \quad E_{\hat{S}_1, \hat{S}_3} - E_{S_1^*, S_3^*} = E_{W_t, S_3^*} - E_{W_t} - E_{W_t, S_1^*} - |Q_t| - d_{t, S_1^*},$$

$$(7.37) \quad E_{\hat{S}_1, V_1} - E_{S_1^*, V_1} = E_{W_t, V_1},$$

$$(7.38) \quad E_{\hat{S}_2, V_1} - E_{S_2^*, V_1} = d_{t, V_1}.$$

Because (2.2) is uniquely optimized by S_1^* , S_2^* and S_3^* ,

$$(7.39) \quad \begin{aligned} & |\hat{S}_1| - |\hat{S}_2| - \frac{2E_{\hat{S}_1} + E_{\hat{S}_1, \hat{S}_3} + E_{\hat{S}_1, V_1} - E_{\hat{S}_2, V_1}}{\tau - 1} \\ & < |S_1^*| - |S_2^*| - \frac{2E_{S_1^*} + E_{S_1^*, S_3^*} + E_{S_1^*, V_1} - E_{S_2^*, V_1}}{\tau - 1}. \end{aligned}$$

Using (7.35)–(7.38), this reduces to

$$(7.40) \quad |W_t| - 1 - \frac{2E_{W_t} + E_{W_t, \hat{W}_t} - |Q_t| - d_{t, S_1^*} - d_{t, V_1}}{\tau - 1} < 0,$$

which is equivalent to

$$(7.41) \quad -\tau + (\tau-1)|W_t| + |Q_t| + d_{t, S_1^*} + d_{t, V_1} - 2E_{W_t} - E_{W_t, \hat{W}_t} < -1,$$

which is (7.33). Since $\zeta_t = d_{t, V_1} + d_{t, S_1^*}$ by (7.2), the inner integral of (7.32) is finite. As W_t and W_{t-1} are disjoint, we obtain that the integral over x_t can be evaluated as

$$(7.42) \quad \begin{aligned} & \int_{x_{t-1}}^\infty \prod_{j=t-1}^t x_j^{-\tau + \zeta_j + |Q_j| + (\tau-1)|W_j| - 2E_{W_j} - E_{W_j, \hat{W}_j}} dx_t \\ & = K x_{t-1}^{1-2\tau + \zeta_{t-1} + \zeta_t + |Q_{t-1}| + |Q_t| + (\tau-1)|W_t \cup W_{t-1}| - 2E_{W_t \cup W_{t-1}} - E_{W_t \cup W_{t-1}, \widehat{W_t \cup W_{t-1}}}}, \end{aligned}$$

for some $K > 0$, where $\widehat{W_t \cup W_{t-1}} = V_1 \cup S_1^* \cup S_3^* \setminus (W_t \cup W_{t-1})$.

We next repeat the above procedure to evaluate the integral over x_{t-1} . Choosing $\hat{S}_2 = S_2^* \cup \{t, t-1\}$, $\hat{S}_1 = S_1^* \cup W_t \cup W_{t-1}$ and $\hat{S}_3 = S_3^* \setminus (W_t \cup W_{t-1} \cup \{t, t-1\})$, we can again use (7.39) to prove that the power of x_{t-1} in (7.42) is smaller than -1 , so that integral (7.42) over x_{t-1} from x_{t-2} to ∞ as in (7.32) results in a power 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.

In general, fix $b \in [t-1]$. We let $Z_b = W_t \cup W_{t-1} \cup \dots \cup W_{t-b}$ and $\hat{Z} = V_1 \cup S_1^* \cup S_3^* \setminus Z_b$. Choosing $\hat{S}_2^{(b)} = S_2^* \cup ([t] \setminus [t-b-1])$, $\hat{S}_1^{(b)} = S_1^* \cup Z_b$ and $\hat{S}_3^{(b)} = S_3^* \setminus (Z_b \cup ([t] \setminus [t-b-1]))$ gives

$$(7.43) \quad E_{\hat{S}_1^{(b)}} - E_{S_1^*} = E_{Z_b} + E_{Z_b, S_1^*},$$

$$(7.44) \quad E_{\hat{S}_1^{(b)}, \hat{S}_3^{(b)}} - E_{S_1^*, S_3^*} = E_{Z_b, S_3^*} - E_{Z_b} - E_{Z_b, S_1^*} - |Q_t| - \dots - |Q_{t-b}| \\ - d_{t, S_1^*} - \dots - d_{t-b, S_1^*},$$

$$(7.45) \quad E_{\hat{S}_1^{(b)}, V_1} - E_{S_1^*, V_1} = E_{Z_b, V_1},$$

$$(7.46) \quad E_{\hat{S}_2^{(b)}, V_1} - E_{S_2^*, V_1} = d_{t, V_1} + d_{t-1, V_1} + \dots + d_{t-b, V_1}.$$

Then, (7.39) reduces to

$$(7.47) \quad |Z_b| - b - 1 - \frac{2E_{Z_b} + E_{Z_b, \hat{Z}_b} - |Q_t| - \dots - |Q_{t-b}| - \zeta_t - \dots - \zeta_{t-b}}{\tau - 1} < 0,$$

which is equivalent to

$$(7.48) \quad - (b + 1)\tau + b + (\tau - 1)|Z_b| + |Q_t| + \dots + |Q_{t-b}| + \zeta_t \\ + \dots + \zeta_{t-b} - 2E_{Z_b} - E_{Z_b, \hat{Z}_b} < -1.$$

This is precisely the exponent that appears in the variable x_{t-b} when integrating for x_{t-b} from x_{t-b-1} to ∞ (as in (7.42) for $b = 1$). Thus indeed, evaluating the integrals in (7.32) one by one does not result in diverging integrals at ∞ as their exponents are smaller than -1 . Therefore (7.32) is also finite, so that the claim in (7.5) follows. \square

PROOF OF LEMMA 7.3. We first argue that we may assume that $a = 1$. Indeed, the integral with $a < 1$ is upper bounded by that with $a = 1$, while for $a > 1$, we can do a change of variables and use that $\min(bx_i x_j, 1) \leq b \min(x_i x_j, 1)$ for all $b > 1$. Thus, from now on, we will assume that $a = 1$.

The proof that this integral is finite has a similar structure as the proof of Lemma 7.2, and we will be more concise here to avoid repetitions. Recall that $S_2^* = [t_2]$ and $S_1^* = [t_2 + t_1] \setminus [t_2]$. Now, it will be convenient to order the x_i for $i \in [t_2]$ such that $x_1 > x_2 > \dots > x_{t_2}$, which we can do w.l.o.g. We first rewrite the integral as

$$(7.49) \quad \int_0^1 \dots \int_0^1 \prod_{j \in [t_2]} x_j^{-\tau + \zeta_j} \prod_{i=t_2+1}^{t_1+t_2} \tilde{h}(i, \mathbf{x}) dx_{t_2} \dots dx_1,$$

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

$$(7.50) \quad \tilde{h}(i, \mathbf{x}) = \int_0^\infty x_i^{-\tau + \zeta_i} \prod_{j \in [t_2]: \{i, j\} \in \mathcal{E}_{S_1^*, S_2^*}} \min(x_i x_j, 1) dx_i.$$

Similar to (7.23), suppose that vertex $i \in S_1^* = [t_2 + t_1] \setminus [t_2]$ has vertices j_1, j_2, \dots, j_l as neighbors in $S_2^* = [t_2]$, where $j_1 < j_2 < \dots < j_l$, so that now $1 > x_{j_1} > x_{j_2} > \dots > x_{j_l}$. Note

that $l = d_{i, S_2^*}$. Then,

$$\begin{aligned}
 \tilde{h}(i, \mathbf{x}) &= \int_0^\infty x_i^{-\tau+\zeta_i} \min(x_i x_{j_1}, 1) \min(x_i x_{j_2}, 1) \cdots \min(x_i x_{j_l}, 1) dx_i \\
 (7.51) \quad &= \int_{1/x_{j_l}}^\infty x_i^{-\tau+\zeta_i} dx_i + \cdots + \int_{1/x_{j_2}}^{1/x_{j_3}} x_i^{-\tau+\zeta_i+l-1} x_{j_2} \cdots x_{j_l} dx_i \\
 &\quad + \int_0^{1/x_{j_1}} x_i^{-\tau+\zeta_i+l} x_{j_1} \cdots x_{j_l} dx_i.
 \end{aligned}$$

Because $\zeta_i + l = \zeta_i + d_{i, S_2^*} = d_i \geq 2$ by (7.2), and $\zeta_i \leq 1$ by Lemma 7.1(i), the first and the last integrals are finite. Computing the integrals yields that for some C_1, \dots, C_l ,

$$\begin{aligned}
 (7.52) \quad \tilde{h}(i, \mathbf{x}) &= C_l x_{j_l}^{\tau-\zeta_i-1} + \cdots + C_2 x_{j_2}^{\tau-\zeta_i-l+1} x_{j_3} \cdots x_{j_l} + C_1 x_{j_1}^{\tau-\zeta_i-l} x_{j_2} \cdots x_{j_l} \\
 &=: C_l \tilde{h}_l(i, \mathbf{x}) + C_{l-1} \tilde{h}_{l-1}(i, \mathbf{x}) + \cdots + C_1 \tilde{h}_1(i, \mathbf{x}).
 \end{aligned}$$

Similar to the argument leading to (7.26), for all $i \in [t_1 + t_2] \setminus [t_2]$, there exists p_i^* such that, for all $1 > x_{v_1} > x_{v_2} > \cdots > x_{v_l}$,

$$(7.53) \quad \tilde{h}(i, \mathbf{x}) \leq K \tilde{h}_{p_i^*}(i, \mathbf{x}),$$

for some $K > 0$. Thus,

$$\begin{aligned}
 (7.54) \quad &\int_0^1 \int_0^{x_1} \cdots \int_0^{x_{t_2-1}} \prod_{j \in [t_2]} x_j^{-\tau+\zeta_j} \prod_{i=t_2+1}^{t_1+t_2} \tilde{h}(i, \mathbf{x}) dx_{t_2} \cdots dx_1 \\
 &\leq K \int_0^1 \int_0^{x_1} \cdots \int_0^{x_{t_2-1}} \prod_{j \in [t_2]} x_j^{-\tau+\zeta_j} \prod_{i=t_2+1}^{t_1+t_2} \tilde{h}_{p_i^*}(i, \mathbf{x}) dx_{t_2} \cdots dx_1.
 \end{aligned}$$

Let $T_i = \{j_q : q \geq p_i^*\}$, so that $|T_i| \leq l$, denote the set of neighbors of i whose terms appear in $h_{p_i^*}(i, \mathbf{x})$. Since $\zeta_i + l = \zeta_i + d_{i, S_2^*} = d_i \geq 2$ by (7.2), and $\zeta_i \leq 1$ by Lemma 7.1(i), we have that $l \geq 1$, and therefore $|T_i| \geq 1$ for all $i \in [t_2]$. For $j \in S_2^*$, let

$$(7.55) \quad Q_j = \{i \in S_1^* : \{i, j\} \in \mathcal{E}_H, j_{p_i^*} \leq j\}$$

be the set of indices i such that x_j appears in $\tilde{h}_{p_i^*}(i, \mathbf{x})$. Then,

$$\begin{aligned}
 (7.56) \quad &\prod_{j \in [t_2]} x_j^{-\tau+\zeta_j} \prod_{i=t_2+1}^{t_1+t_2} \tilde{h}_{p_i^*}(i, \mathbf{x}) \\
 &\leq \tilde{K} \prod_{j \in [t_2]} x_j^{-\tau+\zeta_j+|Q_j|} \prod_{i=t_2+1}^{t_1+t_2} (1/x_{j_{p_i^*}})^{\tau-1-\zeta_i-(l-p_i^*+1)},
 \end{aligned}$$

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

Define $W_j = \{i \in S_1^* : j_{p_i^*} = j\}$ for $j \in S_2^*$ and let $\bar{W}_j = V_H \setminus (W_j \cup [j-1])$. Using that $\zeta_i = d_{i, v_1} + d_{i, S_1^*} + d_{i, S_3^*}$ for $i \in S_1^*$ by (7.2), and that $l - p_i^* + 1 = d_{i, S_2^* \setminus [j_{p_i^*} - 1]}$, (7.56) reduces to

$$(7.57) \quad \tilde{K} \prod_{j \in [t_2]} x_j^{-\tau+\zeta_j+|Q_j|+(\tau-1)|W_j|-2E_{W_j}-E_{W_j, \bar{W}_j}}.$$

We set $\hat{S}_1 = S_1^* \setminus W_{t_2}$, $\hat{S}_2 = S_2^* \setminus \{t_2\}$ and $\hat{S}_3 = S_3^* \cup W_{t_2} \cup \{t_2\}$. Notice that

$$(7.58) \quad E_{S_1^*} - E_{\hat{S}_1} = E_{W_{t_2}} + E_{W_{t_2}, S_1^* \setminus W_{t_2}},$$

$$(7.59) \quad E_{S_1^*, S_3^*} - E_{\hat{S}_1, \hat{S}_3} = E_{W_{t_2}, S_3^*} - d_{t_2, S_1^* \setminus W_{t_2}} - E_{W_{t_2}, S_1^* \setminus W_{t_2}},$$

$$(7.60) \quad E_{S_1^*, V_1} - E_{\hat{S}_1, V_1} = E_{W_{t_2}, V_1},$$

$$(7.61) \quad E_{S_2^*, V_1} - E_{\hat{S}_2, V_1} = d_{t_2, V_1}.$$

Because the optimal solution to (2.2) is unique, we obtain using (7.39) that

$$(7.62) \quad \begin{aligned} & -\tau + (\tau - 1)|W_{t_2}| - 2E_{W_{t_2}} - E_{W_{t_2}, S_1^* \setminus W_{t_2}} \\ & - E_{W_{t_2}, S_3^*} - E_{W_{t_2}, V_1} + d_{t_2, S_1^* \setminus W_{t_2}} + d_{t_2, V_1} > -1. \end{aligned}$$

Note that $\bar{W}_{t_2} = V_1 \cup S_1^* \cup S_3^* \cup \{t_2\} \setminus W_{t_2}$ since $S_2^* = [t_2]$. Therefore,

$$(7.63) \quad E_{W_{t_2}, \bar{W}_{t_2}} = E_{W_{t_2}, S_1^* \setminus W_{t_2}} + E_{W_{t_2}, S_3^*} + d_{t_2, W_{t_2}} + E_{W_{t_2}, V_1}.$$

Using (7.62) and that by (7.2) $\zeta_{t_2} = d_{t_2, V_1}$ then shows that

$$(7.64) \quad -\tau + (\tau - 1)|W_{t_2}| - 2E_{W_{t_2}} - E_{W_{t_2}, \bar{W}_{t_2}} + E_{W_{t_2}, S_2^*} + d_{t_2, S_1^* \setminus W_{t_2}} + \zeta_{t_2} > -1.$$

We then use that $d_{t_2, S_1^* \setminus W_{t_2}} + d_{t_2, W_{t_2}} = d_{t_2, S_1^*}$ to obtain

$$(7.65) \quad -\tau + (\tau - 1)|W_{t_2}| - 2E_{W_{t_2}} - E_{W_{t_2}, \bar{W}_{t_2}} + d_{t_2, S_1^*} + \zeta_{t_2} > -1.$$

Finally, by (7.55), $|Q_{t_2}| = d_{t_2, S_1^*}$ as t_2 is the largest index in S_1^* .

This shows that the integral of (7.57) over $x_{t_2} \in [0, x_{t_2-1}]$ equals a power of x_{t_2-1} . A similar argument, setting $\hat{S}_1 = S_1^* \setminus (W_{t_2} \cup W_{t_2-1})$ and $\hat{S}_2 = S_2^* \setminus \{t_2, t_2 - 1\}$ shows that the integral of (7.57) over $x_{t_2-1} \in [0, x_{t_2-2}]$ equals a power of x_{t_2-2} , and we can proceed to show that the outer integral of (7.57) is finite. We conclude that (7.6) is finite. \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

$$(8.1) \quad \mathbb{P}_n(X_{v_i, v_j} = 0) = e^{-\Theta(n^{\alpha_i + \alpha_j - 1})} (1 + o(1)) = 1 + o(1),$$

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

$$(8.2) \quad \mathbb{P}_n(X_{v_i, v_j} = 0) = o(1),$$

and for $\alpha_i + \alpha_j = 1$ (4.4) yields $\mathbb{P}_n(X_{v_i, v_j} = 0) = \Theta(1)$. Similar to (4.21), we can write the probability that H occurs as an induced subgraph on $\mathbf{v} = (v_1, \dots, v_k)$ as

$$(8.3) \quad \begin{aligned} & \mathbb{P}_n(\text{ECM}^{(n)}(\mathbf{D}) | \mathbf{v} = \mathcal{E}_H) \\ & = \Theta_{\mathbb{P}} \left(\prod_{\{i, j\} \in \mathcal{E}_H : \alpha_i + \alpha_j < 1} n^{\alpha_i + \alpha_j - 1} \prod_{\{i, j\} \notin \mathcal{E}_H : \alpha_i + \alpha_j > 1} e^{-n^{\alpha_i + \alpha_j - 1} / 2} \right). \end{aligned}$$

Similar 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 \mathbf{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

$$(8.4) \quad \begin{aligned} & \max(1 - \tau) \sum_i \alpha_i + \sum_{\{i, j\} \in \mathcal{E}_H : \alpha_i + \alpha_j < 1} \alpha_i + \alpha_j - 1, \\ & \text{s.t. } \alpha_i + \alpha_j \leq 1 \quad \forall \{i, j\} \notin \mathcal{E}_H. \end{aligned}$$

The following lemma shows that this optimization problem attains its optimum for very specific values of α (similar 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

$$(8.5) \quad \max \frac{1-\tau}{2}k + (1-\tau) \sum_i \beta_i + \sum_{\{i,j\} \in \mathcal{E}_H: \beta_i + \beta_j < 0} \beta_i + \beta_j,$$

$$\text{s.t. } \beta_i + \beta_j \leq 0 \quad \forall \{i, j\} \notin \mathcal{E}_H.$$

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. \square

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 by $S_3^* = V_H$ with $B^{(\text{ind})}(H) = 0$. Then:*

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

$$(8.6) \quad \frac{N^{(\text{ind})}(H, W_n^k(\varepsilon))}{n^{\frac{k}{2}(3-\tau)}} = (1 + o_{\mathbb{P}}(1))c^k \mu^{-\frac{k}{2}(\tau-1)} \int_{\varepsilon}^{1/\varepsilon} \dots \int_{\varepsilon}^{1/\varepsilon} (x_1 \dots x_k)^{-\tau}$$

$$\times \prod_{\{i,j\} \in \mathcal{E}_H} (1 - e^{-x_i x_j}) \prod_{\{i,j\} \notin \mathcal{E}_H} e^{-x_i x_j} dx_1 \dots dx_k + f_n(\varepsilon),$$

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

$$(8.7) \quad \lim_{\varepsilon \searrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(f_n(\varepsilon) > \delta \mid J_n) = 0.$$

(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, 6.2 and 6.3 to induced subgraphs, now also using that the probability that an edge $\{i, j\} \notin \mathcal{E}_H$ is not present in the subgraph can be approximated by $\exp(-D_{v_i} D_{v_j} / L_n)$. Lemma 8.2(ii) is an application of the following equivalent lemma to Lemma 7.2 for $S_3^* = V_H$.

LEMMA 8.3 (Induced subgraph integrals over S_3^*). *Suppose that the maximum in (2.3) is uniquely attained by $\mathcal{P} = (S_1^*, S_2^*, S_3^*)$ with $|S_3^*| = s > 0$, and say that $S_3^* = [s]$. Then*

$$(8.8) \quad \int_0^\infty \cdots \int_0^\infty \prod_{i \in [s]} x_i^{-\tau+\zeta_i} \prod_{\{i,j\} \in \mathcal{E}_{S_3^*}} \min(x_i x_j, 1) \prod_{\{i,j\} \notin \mathcal{E}_{S_3^*}} e^{-x_i x_j} dx_s \cdots dx_1 < \infty.$$

PROOF. The proof follows that of Lemma 7.2, where now we need to rely on the uniqueness of (2.3) instead of that of (2.2), and we obtain extra factors $e^{-x_i x_j}$ for all $\{i, j\} \notin \mathcal{E}_H$. We will therefore be more brief, and focus on the differences compared to the proof of Lemma 7.2.

This integral is finite if

$$(8.9) \quad \int_0^\infty \cdots \int_0^\infty \prod_{i \in [s]} x_i^{-\tau+\zeta_i} \prod_{\{i,j\} \in \mathcal{E}_{S_3^*}} \min(x_i x_j, 1) \prod_{\{i,j\} \notin \mathcal{E}_{S_3^*}} \mathbb{1}_{\{x_i x_j < 1\}} dx_s \cdots dx_1 < \infty,$$

since if

$$(8.10) \quad \int_a^b \int_0^{1/x_1} x_1^{\gamma_1} x_2^{\gamma_2} e^{-x_1 x_2} dx_2 dx_1 < \infty,$$

then also

$$(8.11) \quad \int_a^b \int_{1/x_1}^\infty x_1^{\gamma_1} x_2^{\gamma_2} e^{-x_1 x_2} dx_2 dx_1 < \infty.$$

We can show similar to (7.19) and (7.20) 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 ∞ for vertices in some nonempty set U , and from 0 to 1 for vertices in $\bar{U} = S_3^* \setminus U$. Without loss of generality, assume that $U = [t]$ for some $1 \leq t < s$. Define, for $i \in \bar{U}$,

$$(8.12) \quad \hat{h}(i, \mathbf{x}) = \int_0^1 x_i^{-\tau+\zeta_i+d_{i,\bar{U}}} \prod_{j \in U: \{i,j\} \in \mathcal{E}_{S_3^*}} \min(x_i x_j, 1) \prod_{j \in U: \{i,j\} \notin \mathcal{E}_{S_3^*}} \mathbb{1}_{\{x_i x_j < 1\}} dx_i.$$

Then (8.8) results in

$$(8.13) \quad \int_1^\infty \cdots \int_1^\infty \prod_{p \in [t]} x_p^{-\tau+\zeta_j} \prod_{i,j \in U: \{i,j\} \notin \mathcal{E}_{S_3^*}} \mathbb{1}_{\{x_i x_j < 1\}} \prod_{i=t+1}^k \hat{h}(i, \mathbf{x}) dx_t \cdots dx_1.$$

When the induced subgraph of H formed by the vertices of U is not a complete graph, this integral equals zero, as $\mathbb{1}_{\{x_i x_j < 1\}} = 0$ when $i, j \in U$. Thus, we assume that the induced subgraph of H formed by the vertices of U is a complete graph so that $\{\{i, j\} \in U : \{i, j\} \notin \mathcal{E}_H\} = \emptyset$.

We first bound the region of (8.8) where $1 < x_1 < \cdots < x_t$. When $i \in \bar{U}$ is connected to all vertices in U , $\hat{h}(i, \mathbf{x})$ equals $h(i, \mathbf{x})$ defined in (7.22), which can be bounded by (7.26). Otherwise, define

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

Thus, i is connected to vertices $[t] \setminus [a_i]$ and we can write $\hat{h}(i, \mathbf{x})$ as

$$\hat{h}(i, \mathbf{x}) = \int_0^{1/x_t} x_i^{-\tau+\zeta_i+d_{i,S_3^*}} dx_i \cdot x_t \cdots x_{a_i+1} \prod_{j \in [a_i]: \{i,j\} \in \mathcal{E}_H} x_j$$

$$\begin{aligned}
 (8.15) \quad & + \int_{1/x_t}^{1/x_{t-1}} x_i^{-\tau-1+\zeta_i+d_{i,S_3^*}} dx_i \cdot x_{t-1} \cdots x_{a_i+1} \prod_{j \in [a_i]: \{i,j\} \in \mathcal{E}_H} x_j + \cdots \\
 & + \int_{1/x_{a_i+1}}^{1/x_{a_i}} x_i^{-\tau+\zeta_i+d_{i,S_3^*}-t+a_i} dx_i \prod_{j \in [a_i]: \{i,j\} \in \mathcal{E}_H} x_j.
 \end{aligned}$$

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, \dots, C_{a_i} ,

$$\begin{aligned}
 (8.16) \quad \hat{h}(i, \mathbf{x}) &= \prod_{j \in [a_i]: \{i,j\} \in \mathcal{E}_H} x_j (C_t x_t^{\tau-\zeta_i-d_{i,S_3^*}} x_{t-1} \cdots x_{a_i+1} + C_{t-1} x_{t-1}^{\tau-\zeta_i-d_{i,S_3^*}+1} x_{t-2} \cdots x_{a_i+1} \\
 &+ \cdots + C_{a_i} x_{a_i}^{\tau-\zeta_i-d_{i,S_3^*}+t-a_i}) \\
 &=: \hat{h}_t(i, \mathbf{x}) + \cdots + \hat{h}_{a_i}(i, \mathbf{x}).
 \end{aligned}$$

As in (7.26), for every i we can find a p_i^* such that, for all $1 > x_1 > \cdots > x_t$,

$$(8.17) \quad \hat{h}(i, \mathbf{x}) \leq K h_{p_i^*}(i, \mathbf{x})$$

for some $K > 0$. Again, let T_i denote the set of neighbors of vertex i appearing in $h_{p_i^*}(i, \mathbf{x})$, and Q_j as in (7.28), and let $W_j = \{i \in \bar{U} : j_{p_i^*} = j\}$. Then,

$$\begin{aligned}
 (8.18) \quad & \int_1^\infty \cdots \int_{x_{t-1}}^\infty \prod_{j \in [t]} x_j^{-\tau+\zeta_j} \mathbb{1}_{\{x_i x_j < 1\}} \prod_{i=t+1}^k \hat{h}(i, \mathbf{x}) dx_t \cdots dx_1 \\
 & \leq \tilde{K} \int_1^\infty \cdots \int_{x_{t-1}}^\infty \prod_{j \in [t]} x_j^{-\tau+\zeta_j} \prod_{i=t+1}^k \hat{h}_{p_i^*}(i, \mathbf{x}) dx_t \cdots dx_1 \\
 & \leq \int_1^\infty \int_{x_1}^\infty \cdots \int_{x_{t-1}}^\infty \prod_{j \in [t]} x_j^{-\tau+\zeta_j+|Q_j|+(\tau-1)|W_j|-2E_{W_j}-E_{W_j, \hat{W}_j}} dx_t \cdots dx_1
 \end{aligned}$$

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

Indeed, we will now again use the uniqueness of the solution of the optimization problem in (2.3) to again prove that this integral is finite. For this, we 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 $U \cup \bar{U} \setminus (W_t \cup \{t\}) = \hat{S}_3$.

We conclude that \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 (7.39) to show that the integral over x_t is finite. Iterating this proves Lemma 8.3. \square

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

LEMMA 8.4 (Induced subgraph integrals over $S_1^* \cup S_2^*$). *Suppose the optimal solution to (2.3) is unique, and attained by $\mathcal{P} = (S_1^*, S_2^*, S_3^*)$. Say that $S_2^* = [t_2]$ and $S_1^* = [t_2 + t_1] \setminus [t_2]$.*

Then, for every $a > 0$,

$$(8.19) \quad \int_0^a \cdots \int_0^a \int_0^\infty \cdots \int_0^\infty \prod_{j \in [t_1+t_2]} x_j^{-\tau+\zeta_j} \prod_{\{i,j\} \in \mathcal{E}_{S_1^*, S_2^*}} \min(x_i x_j, 1) \\ \times \prod_{\{i,j\} \notin \mathcal{E}_{S_1^*, S_2^*}} e^{-x_i x_j} dx_{t_1+t_2} \cdots dx_1 < \infty.$$

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. \square

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.

Acknowledgments. The first author was supported by NWO TOP grant 613.001.451 and also supported by NWO Gravitation Networks grant 024.002.003 and NWO VICI grant 639.033.806.

The second author was supported by NWO TOP grant 613.001.451 and also supported by NWO Gravitation Networks grant 024.002.003, an NWO TOP-GO grant and an ERC Starting Grant.

The third author was supported by NWO TOP grant 613.001.451.

REFERENCES

- [1] ALBERT, R., JEONG, H. and BARABÁSI, A.-L. (1999). Internet: Diameter of the world-wide web. *Nature* **401** 130–131. <https://doi.org/10.1038/43601>
- [2] BENSON, A. R., GLEICH, D. F. and LESKOVEC, J. (2016). Higher-order organization of complex networks. *Science* **353** 163–166.
- [3] BLÁSIUS, T., FRIEDRICH, T. and KROHMER, A. (2018). Cliques in hyperbolic random graphs. *Algorithmica* **80** 2324–2344. MR3800263 <https://doi.org/10.1007/s00453-017-0323-3>
- [4] BOGUÑÁ, M. and PASTOR-SATORRAS, R. (2003). Class of correlated random networks with hidden variables. *Phys. Rev. E* **68** 036112. <https://doi.org/10.1103/PhysRevE.68.036112>
- [5] BOLLOBÁS, B. (1980). A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European J. Combin.* **1** 311–316. MR0595929 [https://doi.org/10.1016/S0195-6698\(80\)80030-8](https://doi.org/10.1016/S0195-6698(80)80030-8)
- [6] BOLLOBÁS, B. (2001). *Random Graphs*, 2nd ed. *Cambridge Studies in Advanced Mathematics* **73**. Cambridge Univ. Press, Cambridge. MR1864966 <https://doi.org/10.1017/CBO9780511814068>
- [7] BOLLOBÁS, B., JANSON, S. and RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31** 3–122. MR2337396 <https://doi.org/10.1002/rsa.20168>
- [8] BRITTON, T., DEIJFEN, M. and MARTIN-LÖF, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.* **124** 1377–1397. MR2266448 <https://doi.org/10.1007/s10955-006-9168-x>
- [9] CHUNG, F. and LU, L. (2002). The average distances in random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA* **99** 15879–15882. MR1944974 <https://doi.org/10.1073/pnas.252631999>
- [10] COLOMER-DE SIMON, P. and BOGUÑÁ, M. (2012). Clustering of random scale-free networks. *Phys. Rev. E* **86** 026120. <https://doi.org/10.1103/PhysRevE.86.026120>
- [11] FALOUTSOS, M., FALOUTSOS, P. and FALOUTSOS, C. (1999). On power-law relationships of the Internet topology. In *ACM SIGCOMM Computer Communication Review* **29** 251–262. ACM, New York.
- [12] GAO, C. and LAFFERTY, J. (2017). Testing network structure using relations between small subgraph probabilities. Preprint. Available at [arXiv:1704.06742](https://arxiv.org/abs/1704.06742) [stat.ME].
- [13] GAO, J., VAN DER HOFSTAD, R., SOUTHWELL, A. and STEGEHUIS, C. (2020). Counting triangles in power-law uniform random graphs. *Electron. J. Combin.* **27**. <https://doi.org/10.37236/9239>
- [14] GARAVAGLIA, A. and STEGEHUIS, C. (2019). Subgraphs in preferential attachment models. *Adv. in Appl. Probab.* **51** 898–926. MR4001020 <https://doi.org/10.1017/apr.2019.36>

- [15] VAN DER HOFSTAD, R. JANSSEN, A. J. E. M. VAN LEEUWAARDEN, J.S.H. and STEGEHUIS, C. (2017). Local clustering in scale-free networks with hidden variables. *Phys. Rev. E* **95** 022307. <https://doi.org/10.1103/PhysRevE.95.022307>
- [16] VAN DER HOFSTAD, R., VAN DER HOORN, P., LITVAK, N. and STEGEHUIS, C. Limit theorems for assortativity and clustering in the configuration model with scale-free degrees. (2020) *Adv. Appl. Probab.* **52** 1035-1084.
- [17] VAN DER HOFSTAD, R. (2017). *Random Graphs and Complex Networks. Vol. 1. Cambridge Series in Statistical and Probabilistic Mathematics* **43**. Cambridge Univ. Press, Cambridge. MR3617364 <https://doi.org/10.1017/9781316779422>
- [18] VAN DER HOFSTAD, R., HOOGHIEMSTRA, G. and VAN MIEGHEM, P. (2005). Distances in random graphs with finite variance degrees. *Random Structures Algorithms* **27** 76–123. MR2150017 <https://doi.org/10.1002/rsa.20063>
- [19] VAN DER HOFSTAD, R. and LITVAK, N. (2014). Degree-degree dependencies in random graphs with heavy-tailed degrees. *Internet Math.* **10** 287–334. MR3259269 <https://doi.org/10.1080/15427951.2013.850455>
- [20] VAN DER HOORN, P. and LITVAK, N. (2015). Upper bounds for number of removed edges in the erased configuration model. In *Algorithms and Models for the Web Graph. Lecture Notes in Computer Science* **9479** 54–65. Springer, Cham. MR3500668 https://doi.org/10.1007/978-3-319-26784-5_5
- [21] JANSON, S. (2009). The probability that a random multigraph is simple. *Combin. Probab. Comput.* **18** 205–225. MR2497380 <https://doi.org/10.1017/S0963548308009644>
- [22] JEONG, H., TOMBOR, B., ALBERT, R., OLTVAI, Z. and BARABÁSI, A.-L. (2000). The large-scale organization of metabolic networks. *Nature* **407** 651–654.
- [23] KRIOUKOV, D., PAPADOPOULOS, F., KITSAK, M., VAHDAT, A. and BOGUÑÁ, M. (2010). Hyperbolic geometry of complex networks. *Phys. Rev. E* (3) **82** 036106. MR2787998 <https://doi.org/10.1103/PhysRevE.82.036106>
- [24] MASLOV, S., SNEPPEN, K. and ZALIZNYAK, A. (2004). Detection of topological patterns in complex networks: Correlation profile of the Internet. *Phys. A* **333** 529–540. <https://doi.org/10.1016/j.physa.2003.06.002>
- [25] MILO, R., ITZKOVITZ, S., KASHTAN, N., LEVITT, R., SHEN-ORR, S., AYZENSHTAT, I. and SHEFFER, M. (2004). Superfamilies of evolved and designed networks. *Science* **303** 1538–1542. [https://doi.org/10.1126/science.1089167](https://doi.org/10.1126/science.1126/science.1089167)
- [26] MILO, R., SHEN-ORR, S., ITZKOVITZ, S., KASHTAN, N. and CHKLOVSKII, D. (2002). Network motifs: Simple building blocks of complex networks. *Science* **298** 824–827. <https://doi.org/10.1126/science.298.5594.824>
- [27] ONNELA, J.-P., SARAMÄKI, J. and KERTÉSZ, J. (2005). Intensity and coherence of motifs in weighted complex networks. *Phys. Rev. E* **71** 065103. <https://doi.org/10.1103/PhysRevE.71.065103>
- [28] OSTILLI, M. (2014). Fluctuation analysis in complex networks modeled by hidden-variable models: Necessity of a large cutoff in hidden-variable models. *Phys. Rev. E* **89** 022807. <https://doi.org/10.1103/PhysRevE.89.022807>
- [29] RAVASZ, E. and BARABÁSI, A.-L. (2003). Hierarchical organization in complex networks. *Phys. Rev. E* **67** 026112. <https://doi.org/10.1103/PhysRevE.67.026112>
- [30] STEGEHUIS, C. (2019). Degree correlations in scale-free random graph models. *J. Appl. Probab.* **56** 672–700. MR4015632 <https://doi.org/10.1017/jpr.2019.45>
- [31] STEGEHUIS, C., VAN DER HOFSTAD, R. and VAN LEEUWAARDEN, J. S. H. (2017). Clustering spectrum of scale-free networks. *Phys. Rev. E* **96** 042309. <https://doi.org/10.1103/physreve.96.042309>
- [32] STEGEHUIS, C., VAN DER HOFSTAD, R., VAN LEEUWAARDEN, J. S. H. and JANSSEN, A. J. E. M. (2019). Variational principle for scale-free network motifs. *Sci. Rep.* **9** 6762. <https://doi.org/10.1038/s41598-019-43050-8>
- [33] TSOURAKAKIS, C. E., PACHOCKI, J. and MITZENMACHER, M. (2017). Scalable motif-aware graph clustering. In *Proceedings of the 26th International Conference on World Wide Web. WWW'17* 1451–1460. <https://doi.org/10.1145/3038912.3052653>
- [34] VÁZQUEZ, A., PASTOR-SATORRAS, R. and VESPIGNANI, A. (2002). Large-scale topological and dynamical properties of the Internet. *Phys. Rev. E* **65** 066130. <https://doi.org/10.1103/PhysRevE.65.066130>
- [35] WUCHTY, S., OLTVAI, Z. N. and BARABÁSI, A.-L. (2003). Evolutionary conservation of motif constituents in the yeast protein interaction network. *Nat. Genet.* **35** 176–179. <https://doi.org/10.1038/ng1242>