

OPTIMAL SUBGRAPHS IN GEOMETRIC SCALE-FREE RANDOM GRAPHS

BY RICCARDO MICHIELAN ^{1,a}, CLARA STEGEHUIS ^{1,b} MATTHIAS WALTER ^{1,c}

¹University of Twente, Enschede, Netherlands, ^ar.michielan@utwente.nl; ^bc.stegehuis@utwente.nl; ^cm.walter@utwente.nl

Geometric scale-free random graphs are popular models for networks that exhibit as heavy-tailed degree distributions, small-worldness and high clustering. In these models, vertices have weights that cause the heavy-tailed degrees and are embedded in a metric space so that close-by groups of vertices tend to cluster. The interplay between the vertex weights and positions heavily affects the local structure of the random graph, in particular the occurrence of subgraph patterns, but the dependencies in these structures and weights make them difficult to analyze. In this paper we investigate subgraph counts using a *divide et impera* strategy: first counting the number of subgraphs in specific classes of vertices; then computing which class yields maximum contribution. Interestingly, the scaling behavior of induced and general subgraphs in such geometric heavy-tailed random graphs is closely related to the solution of a mixed-integer linear program which also shows that subgraphs appear predominantly on vertices with some prescribed degrees and inter-distances. Finally, we derive precise asymptotics for trees and Hamiltonian subgraphs.

1. Introduction. Network subgraphs have been an object of intense study over the past decades, both from a combinatorial perspective, but more recently also in analyzing real-world network data. Here, subgraph counts have gained attention through so-called motifs: statistically significant subgraphs of networks [19]. The typical setting used to analyze motifs is to compare the frequency of a particular chosen subgraph in a network with that in an associated network null-model or random graph model. This investigation initially started in social networks [9, 10], biological networks [29, 23, 16, 1] and ecological networks [6], but the interest in network motifs rapidly expanded to other fields including finance and communication science [26, 27, 13].

One of the main criticisms of the motif analyses is that the null-model compared to which statistical significance is claimed is often unrealistic. Indeed, in most examples, a networked data set is compared to the configuration model [20]: a random graph model that only matches the exact degree sequence of the network. In such random graphs, any type of non-tree subgraph is rare, in contrast to most real-world networks which are often heavily clustered. Therefore, comparing to the configuration model often yields that every subgraph containing a cycle appears more frequently in the data set than in the random graph model, while such structures are not a rare property but rather a commonality of real-world data. Investigating small subgraphs in more realistic random graph models is therefore of vital importance to better understand which specific subgraph frequencies emerge naturally from realistic network features.

Recurrent properties in large real-world networks are heavy-tailed degree sequences, small distances and large clustering coefficient. Several recent random models create graphs with all these properties by using vertex weights to tune the degree sequence to a power-law distribution, and by introducing an underlying geometry that can explain the clusteredness. Examples of such models are hyperbolic random graphs [15], spatial preferential attachment

MSC2020 subject classifications: Primary 05C60, 05C82; secondary 90C11, 05C80.

Keywords and phrases: subgraph count, network motifs, scale-free graphs, network geometry, mixed-integer linear programming.

models [12], and spatial inhomogeneous random graphs [5, 7, 8]. The local limits of this class of random graphs deviates substantially from the usual locally tree-like structure of many complex network models [32, 33]. Therefore, the occurrence of small subgraphs differs significantly when compared to random graphs where geometric features are absent. However, the geometric setting of these random graph models make them mathematically difficult to analyze, as the geometry ruins edge independence. Furthermore, the power-law degrees create so-called degree-degree correlations, which destroy the product form of the edge probabilities.

For this reason, subgraph counts have mainly been studied in random graphs with more homogeneous degree distribution and no underlying geometric structure [25, 21, 11]. More recently, a novel powerful technique has been introduced to deal with degree inhomogeneity in random graphs for counting subgraph occurrences [30, 34]. This approach allows one to obtain a precise estimate of any subgraph count and at the same time to locate where within the network the subgraph appears most frequently. However, this technique strongly relies on independent edge presences, which do not hold for geometric random graph models.

In this paper, we investigate the number of subgraphs in a model that combines scale-freeness and geometry, the geometric inhomogeneous random graph [5]. We show that most of the subgraphs emerge from set of vertices with specific degrees and inter-distances. We create a new technique to overcome the difficulties in analyzing dependent, geometric edge presences by introducing and analyzing a mixed-integer linear program (MILP) that is able to capture the edge dependencies in its constraints. We use its optimizer to prove the scaling of all subgraph counts. The MILP not only determines the scaling of the subgraph counts, it also gives insights into the typical locations of all possible subgraphs, in terms of the vertex degrees and their distances. We show that via a sequence of related MILPs, we can also deduce the uniqueness and possible symmetries in these optimal locations. To our knowledge, this is one of the first papers that uses mixed integer linear programs to analyze random graphs, demonstrating that they can be a powerful tool to show the dominant behavior of complex random graph models. Furthermore, the method is general, and applies to all possible fixed subgraphs, while to the best of our knowledge, all previous results are restricted to specific subgraphs with extreme symmetry properties, such as cliques [18, 22, 2, 3, 31].

Our methods show that in these complex random graph models, subgraphs often appear at very specific degrees and distances. This highlights the intricate interplay between power-law degrees and geometry: some subgraphs are typically formed by close by vertices of low degrees caused by geometry, while other subgraphs are dominated by high-degree vertices that are typically not close by in the geometry.

1.1. Notation. Throughout this paper we build a random graph model on the set of vertices $[n] = \{1, 2, \dots, n\}$ and study the large n limit. We therefore use classical asymptotic notation, in terms of the graph size n . For any two non-negative functions $f(n), g(n)$ we will write: $f(n) = o(g(n))$ if $\lim_{n \rightarrow \infty} f(n)/g(n) = 0$; $f(n) = O(g(n))$ if $\limsup_{n \rightarrow \infty} f(n)/g(n) < \infty$; $f(n) = \Omega(g(n))$ if $\liminf_{n \rightarrow \infty} f(n)/g(n) > 0$; $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$. Moreover, we will say that a sequence of events $\{\mathcal{E}_n\}_{n \geq 1}$ happens with high probability (w.h.p.) if $\lim_{n \rightarrow \infty} \mathbb{P}(\mathcal{E}_n) = 1$. Finally, we will denote by $X_n \xrightarrow{\mathbb{P}} X$ the convergence in probability of a sequence of random variables, meaning that $\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$, for all $\varepsilon > 0$.

1.2. Model. We now describe the geometric inhomogeneous random graphs (GIRGs) [5]. In this class of models, the vertex set is $V = [n] := \{1, \dots, n\}$ and each vertex i is associated with a weight $w_i > 0$. The weight sequence follows a power law distribution with

exponent $\tau > 0$. We will assume that $\mathbf{w} = \{w_i\}_{i \in [n]}$ are i.i.d. samples of a same random variable W such that

$$(1.1) \quad \mathbb{P}(W > x) = cx^{1-\tau}(1 + o(1)),$$

for all $x \geq w_0$, for some global constants $w_0, c > 0$. In this paper, we assume that $\tau \in (2, 3)$, that is, the weights have finite mean μ , but infinite variance. For simplicity, we assume that $w_0 = 1$. Each vertex i is also assigned a position x_i . As ground space, we consider the d -dimensional torus $\mathbb{T}^d = \mathbb{R}^d / \mathbb{Z}^d$, which can be described as the d -dimensional cube $[0, 1]^d$ where opposite boundaries are identified, endowed with the infinity norm distance

$$\|x - y\| := \max_{j=1, \dots, d} \min \left\{ |x^{(j)} - y^{(j)}|, 1 - |x^{(j)} - y^{(j)}| \right\},$$

for any $x, y \in \mathbb{T}^d$. The vertex positions $\mathbf{x} = \{x_i\}_{i \in [n]}$ are sampled i.i.d uniformly over \mathbb{T}^d .

Any two vertices $u, v \in [n]$ are then independently connected with probability p_{uv} determined by the weights and positions of the vertices

$$(1.2) \quad p_{uv} = p(w_u, w_v, x_u, x_v) = 1 \wedge \left(\frac{w_u w_v}{n\mu \|x_u - x_v\|^d} \right)^\gamma,$$

where $\mu = \mathbb{E}[W]$ and $\gamma > 1$ is a fixed parameter. The choice $\gamma = \infty$ is also allowed, and in this case

$$(1.3) \quad p_{uv} = p(w_u, w_v, x_u, x_v) = \mathbb{1}_{\{w_u w_v > n\mu \|x_u - x_v\|^d\}}.$$

We denote the resulting random graph by $\text{GIRG}^{(n)}$, where dependency on the parameters τ, γ, d is implicit. The connection probability in (1.2)–(1.3) reflects the idea that nearby (similar) vertices are likely to be connected, generating a geometric structure. At the same time, high-degree (popular) vertices are also likely to connect.

Sometimes, it will be useful to consider a *cutoff* version of the GIRG, that is, conditioning $\text{GIRG}^{(n)}$ on the realization of the event

$$(1.4) \quad \mathcal{E}_{\bar{\varepsilon}} := \left\{ \max_{v \in V} w_v < n^{1/(\tau-1)} / \bar{\varepsilon}, \min_{u, v \in V} \|x_u - x_v\| > \bar{\varepsilon} n^{-1/d} \right\},$$

for some $\bar{\varepsilon} > 0$. It is immediate to see that for any $\delta > 0$, there exists $\bar{\varepsilon}$ such that $\mathbb{P}(\mathcal{E}_{\bar{\varepsilon}}) > 1 - \delta$, meaning that $n^{1/(\tau-1)}$ and $n^{-1/d}$ are the natural cutoff for vertex weights and inter-distances in $\text{GIRG}^{(n)}$.

1.3. Subgraph counting. In this paper we analyze the number of copies of fixed subgraphs in $\text{GIRG}^{(n)}$ when the number of vertices $|V| = n$ tends to infinity, while the size k of the subgraph is fixed. As $\text{GIRG}^{(n)}$ is a random graph, the number of occurrences of a specified subgraph is a random variable, and we will show that it obeys a specific scaling in n , with high probability. We now describe the problem more precisely, while introducing some notation.

Let $H = (\mathcal{V}, \mathcal{E})$ be a deterministic graph, with $\mathcal{V} = [k]$ and $k = O(1)$. We denote by $N^{(\text{ind})}(H)$ the number of induced subgraphs of $\text{GIRG}^{(n)}$ that are equal to H , and by $N^{(\text{sub})}(H)$ the number of occurrences of H as general subgraph in $\text{GIRG}^{(n)}$. That is,

$$(1.5) \quad N^{(\text{ind})}(H) = \sum_{v \in V_k} \mathbb{1}_{\{H = \text{GIRG}^{(n)}|_v\}},$$

$$(1.6) \quad N^{(\text{sub})}(H) = \sum_{v \in V_k} \mathbb{1}_{\{H \subseteq \text{GIRG}^{(n)}|_v\}},$$

where $V_k = \{(v_1, \dots, v_k) \subset V\}$ is the collection of all ordered sets of vertices of size k in V and $\text{GIRG}^{(n)}|_v$ denotes the induced subgraph of $\text{GIRG}^{(n)}$ on the ordered vertex set v . The relation $H = \text{GIRG}^{(n)}|_v$ means $(v_i, v_j) \in E$ if and only if $(i, j) \in \mathcal{E}$, whereas $H \subseteq \text{GIRG}^{(n)}|_v$ means $(v_i, v_j) \in E$ whenever $(i, j) \in \mathcal{E}$.

The expected number of copies of H in $\text{GIRG}^{(n)}$ can be written as

$$(1.7) \quad \begin{aligned} \mathbb{E} \left[N^{(\text{ind})}(H) \right] &= \sum_{v \in V_k} \mathbb{P} \left(\text{GIRG}^{(n)}|_v = H \right) \\ &= \frac{n!}{(n-k)!} \int \mathbb{P} \left(\text{GIRG}^{(n)}(\mathbf{w}, \mathbf{x})|_v = H \right) d\mathbb{P}(\mathbf{w}, \mathbf{x}). \end{aligned}$$

A similar expression holds for $N^{(\text{sub})}(H)$, replacing $H = \text{GIRG}^{(n)}|_v$ with $H \subseteq \text{GIRG}^{(n)}|_v$. When H is a k -clique, the integral in (1.7) can be computed explicitly and the expected number of k -cliques is asymptotically linear in n or scales as $n^{k(3-\tau)/2}$, depending on k and τ [18]. However, for general H there is no existing method to solve such integrals because the weights and spatial coordinates make the edge presences of H dependent.

Our main step to overcome this problem is to first focus on counting the number of subgraphs that appear on specific subsets of V_k only. Namely, if $A \subseteq V_k$, we define

$$(1.8) \quad N^{(\text{ind})}(H; A) = \sum_{v \in A} \mathbb{1}_{\{H = \text{GIRG}^{(n)}|_v\}}, \quad N^{(\text{sub})}(H; A) = \sum_{v \in A} \mathbb{1}_{\{H \subseteq \text{GIRG}^{(n)}|_v\}},$$

In the next section, we will show that first computing $N^{(\text{ind})}(H; A)$ and $N^{(\text{sub})}(H; A)$ generates a general approach to estimate all fixed-size subgraph counts H , addressing the subgraph counting problem systematically.

1.4. Organization of the paper. We present our main results in Section 2, where we state asymptotic results for subgraph counts and introduce an optimization problem, together with its mixed-integer linear program reformulation. We then provide the proofs. In Section 3 we prove the existence of *optimal* regions in the GIRG, where the subgraph count is maximal. Finally, in Section 4 we prove sharp results for specific classes of subgraphs.

2. Results. The integral in (1.7) is difficult to solve in general, as the edge presences of H are dependent. Thus, we apply a *divide et impera* strategy: first, we count subgraphs formed on vertices with prescribed weights and inter-positions; then, we derive and solve an optimization problem to determine where the subgraphs are more likely to appear.

More precisely, for any $\zeta \in \mathbb{R}$ and $\varepsilon \in (0, 1)$ we define the interval $I_\varepsilon(n^\zeta) = [\varepsilon n^\zeta, n^\zeta/\varepsilon]$. For any $\alpha = \{\alpha_i\}_{i \in [k]}$ and $\beta = \{\beta_{\{i,j\}}\}_{i,j \in [k], i \neq j}$, we define the set $M_\varepsilon^{(\alpha, \beta)} \subseteq V_k$ as

$$(2.1) \quad M_\varepsilon^{(\alpha, \beta)} = \left\{ (v_1, \dots, v_k) \subset V : w_{v_i} \in I_\varepsilon(n^{\alpha_i}) \forall i, \|x_{v_i} - x_{v_j}\| \in I_\varepsilon(n^{\beta_{\{i,j\}}}) \forall i, j \right\}.$$

The random set $M_\varepsilon^{(\alpha, \beta)}$ contains all the lists of k vertices in $\text{GIRG}^{(n)}$ such that their weights scale as $n^{\alpha_1}, \dots, n^{\alpha_k}$, and their distances on \mathbb{T}^d scale as $n^{\beta_{\{i,j\}}}$ for any $i \neq j$.

Since weights in $\text{GIRG}^{(n)}$ are independently sampled from (1.1), for any $\delta > 0$ there exists $C > 0$ such that

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left(\max_{v \in V} w_v > C n^{1/(\tau-1)} \right) < \delta.$$

Thus, whenever $\alpha_i > \frac{1}{\tau-1}$ for any $i \in [k]$, $M_\varepsilon^{(\alpha, \beta)}$ is empty with high probability. Similarly, as the random positions are independently and uniformly sampled from \mathbb{T}^d , for any $\delta > 0$

there exists $c > 0$ such that

$$\limsup_{n \rightarrow \infty} \mathbb{P} \left(\min_{u, v \in V} \|x_u - x_v\| < cn^{-1/d} \right) < \delta,$$

so that $M_\varepsilon^{(\alpha, \beta)}$ is empty with high probability whenever $\beta_{\{i, j\}} < -\frac{1}{d}$ for some i, j . Moreover, the β variables need to satisfy structural constraints that emerge from the torus norm. Indeed, if $\|x_i - x_j\| \propto n^{\beta_{\{i, j\}}}$ and $\|x_j - x_k\| \propto n^{\beta_{\{j, k\}}}$, then $\|x_i - x_k\| = O(n^{\max(\beta_{\{i, j\}}, \beta_{\{j, k\}})})$, as otherwise the triangle inequality would imply that the distance from i to j to k would be shorter than the distance between i and k . In other words, β should satisfy $\beta_{\{i, j\}} \leq \max(\beta_{\{i, s\}}, \beta_{\{j, s\}}), \forall s \neq i, j$. Thus, we define the *feasible region*

$$(2.2) \quad \mathcal{F} := \left\{ (\{\alpha_i\}_{i \in [k]}, \{\beta_{\{i, j\}}\}_{i, j \in [k]}) : \alpha_i \in [0, 1/(\tau - 1)], \beta_{\{i, j\}} \in [0, -1/d], \right. \\ \left. \beta_{\{i, j\}} \leq \max(\beta_{\{i, s\}}, \beta_{\{j, s\}}), \forall i, j, s \in [k] \right\},$$

so that $M_\varepsilon^{(\alpha, \beta)}$ is empty with high probability whenever $(\alpha, \beta) \notin \mathcal{F}$. In particular, observe that conditionally on the event \mathcal{E}_ε defined in (1.4),

$$N^{(\text{sub})}(H) = N^{(\text{sub})} \left(H; \bigcup_{(\alpha, \beta) \in \mathcal{F}} M_\varepsilon^{(\alpha, \beta)} \right) \quad \text{and} \quad N^{(\text{ind})}(H) = N^{(\text{ind})} \left(H; \bigcup_{(\alpha, \beta) \in \mathcal{F}} M_\varepsilon^{(\alpha, \beta)} \right)$$

almost surely.

2.1. Optimal subgraphs. We now present our results on subgraphs with specific vertex parameters. We denote the occurrences of H as induced or general subgraph in $\text{GIRG}^{(n)}$ with vertices in $M_\varepsilon^{(\alpha, \beta)}$ by $N^{(\text{ind})}(H; M_\varepsilon^{(\alpha, \beta)})$ and $N^{(\text{sub})}(H; M_\varepsilon^{(\alpha, \beta)})$ as in Equation (1.8). Our first lemma shows how the expected subgraph counts in $M_\varepsilon^{(\alpha, \beta)}$ scales in n :

LEMMA 2.1. *For any $H = (\mathcal{V}, \mathcal{E})$ of size k and $(\alpha, \beta) \in \mathcal{F}$, define*

$$(2.3) \quad f_H(\alpha, \beta) := k + (1 - \tau) \sum_{i=1}^k \alpha_i + d \sum_{j=2}^k \min_{i: i < j} \beta_{\{i, j\}} \\ + \gamma \sum_{\substack{i, j \in [k] \\ (i, j) \in \mathcal{E}}} \min(\alpha_i + \alpha_j - d\beta_{\{i, j\}} - 1, 0).$$

Then, for any fixed $\varepsilon \in (0, 1)$,

$$\mathbb{E} \left[N^{(\text{sub})}(H; M_\varepsilon^{(\alpha, \beta)}) \right] = \Theta \left(n^{f_H(\alpha, \beta)} \right),$$

and

$$\mathbb{E} \left[N^{(\text{ind})}(H; M_\varepsilon^{(\alpha, \beta)}) \right] = \begin{cases} \Theta \left(n^{f_H(\alpha, \beta)} \right) & \text{if } \alpha_i + \alpha_j \leq d\beta_{\{i, j\}} + 1, \forall (i, j) \notin \mathcal{E}, \\ 0 & \text{otherwise.} \end{cases}$$

The proof of this result can be found in Section 3. Lemma 2.1 shows that, for different choices of (α, β) , the sets $M_\varepsilon^{(\alpha, \beta)}$ yield different contributions to the expected number of copies of H in $\text{GIRG}^{(n)}$. Next, the natural question is which $M_\varepsilon^{(\alpha, \beta)}$ provides the maximum

contribution. To this end, we formulate an optimization problem for the expected number of general subgraphs and find (α, β) that maximize $f_H(\alpha, \beta)$:

$$(2.4a) \quad \max k + (1 - \tau) \sum_{i=1}^k \alpha_i + d \sum_{j=2}^k \min_{i:i < j} \beta_{\{i,j\}} + \gamma \sum_{\substack{i,j \in [k] \\ (i,j) \in \mathcal{E}}} \min(\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1, 0)$$

$$(2.4b) \quad \text{over} \quad \alpha_i \in [0, 1/(\tau - 1)] \quad \forall i \in [k]$$

$$(2.4c) \quad \beta_{\{i,j\}} \in [-1/d, 0] \quad \forall i, j \in [k] : i < j$$

$$(2.4d) \quad \text{s.t.} \quad \beta_{\{i,j\}} \leq \max(\beta_{\{i,s\}}, \beta_{\{s,j\}}) \quad \forall i, j, s \in [k] : i < j, s \neq i, s \neq j.$$

For the expected number of induced subgraphs, we formulate the same optimization problem as in (2.4), including the extra set of constraints:

$$(2.5) \quad \alpha_i + \alpha_j \leq 1 + d\beta_{\{i,j\}} \quad \forall (i, j) \notin \mathcal{E}$$

It is easily checked that the optimization problems (2.4) and (2.4)+(2.5) are essentially independent of the dimension parameter d :

PROPOSITION 2.2. *A vector (α, β) is feasible or optimal for (2.4) or satisfies (2.5) for a particular dimension $d = \bar{d}$ if and only if the vector $(\alpha, \bar{d} \cdot \beta)$ satisfies the respective property for $d = 1$.*

PROOF. Let $d \geq 1$. Under the transformation $\tilde{\beta} = d \cdot \beta$, $f_H^{(d)}(\alpha, \beta) = f_H^{(1)}(\alpha, \tilde{\beta})$, where $f_H^{(d)}$ denotes the objective function (2.3) for a given d . Furthermore, we can rewrite (2.4c)–(2.5) as

$$\begin{aligned} \tilde{\beta}_{\{i,j\}} &\in [-1, 0], & \forall i, j \in [k] : i < j, \\ \alpha_i + \alpha_j &\leq 1 + \tilde{\beta}_{\{i,j\}}, & \forall (i, j) \notin \mathcal{E}, \end{aligned}$$

which are the same constraints as for the optimization problem with $d = 1$. \square

Another interesting property is that γ does not play any role in the optimization problem for general subgraphs (2.4). Indeed, general subgraphs are always formed to minimize the *edge-energy*:

PROPOSITION 2.3. *The vector (α^*, β^*) is optimal for (2.4) at a given $\gamma > 1$ if and only if it is optimal for $\gamma = \infty$, that is, if and only if (α^*, β^*) maximizes the function*

$$f(\alpha, \beta) := k + (1 - \tau) \sum_{i=1}^k \alpha_i + d \sum_{j=2}^k \min_{i:i < j} \beta_{\{i,j\}}$$

subject to (2.4b), (2.4c), (2.4d) and

$$(2.6) \quad \alpha_i + \alpha_j \geq 1 + d\beta_{\{i,j\}}, \quad \forall (i, j) \in \mathcal{E}.$$

The proof of Proposition 2.3 can be found in Section 3. As an immediate consequence of Lemma 2.1 we get that the expected number of subgraphs is dominated by contributions from the optimal solution (α^*, β^*) :

COROLLARY 2.4. *Let H be a subgraph of size k . Let $\star \in \{\text{ind}, \text{sub}\}$ denote the induced or general subgraph setting. Suppose that the corresponding problem (2.4) or (2.4)+(2.5) has unique solution (α^*, β^*) . Then, for any $(\alpha, \beta) \in \mathcal{F} \setminus \{(\alpha^*, \beta^*)\}$ and $\varepsilon \in (0, 1)$,*

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E} \left[N^{(\star)} \left(H; M_\varepsilon^{(\alpha, \beta)} \right) \right]}{\mathbb{E} \left[N^{(\star)} \left(H; M_\varepsilon^{(\alpha^*, \beta^*)} \right) \right]} = 0.$$

PROOF. Since (α^*, β^*) is the unique solution of (2.4) (or (2.4)+(2.5)) it follows that

$$f_H(\alpha, \beta) < f_H(\alpha^*, \beta^*),$$

for any $(\alpha, \beta) \in \mathcal{F}$ with $(\alpha, \beta) \neq (\alpha^*, \beta^*)$. Then, by Lemma 2.1,

$$\frac{\mathbb{E} \left[N^{(\text{sub})} \left(H; M_\varepsilon^{(\alpha, \beta)} \right) \right]}{\mathbb{E} \left[N^{(\text{sub})} \left(H; M_\varepsilon^{(\alpha^*, \beta^*)} \right) \right]} = \frac{\Theta(n^{f_H(\alpha, \beta)})}{\Theta(n^{f_H(\alpha^*, \beta^*)})} \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

The case of induced subgraphs follows similarly. \square

Moreover, we obtain a stronger concentration result for subgraphs H that achieve their solution on particular values α and β .

THEOREM 2.5. *Under the assumptions of Corollary 2.4, suppose $(\alpha^*, \beta^*) \equiv (1/2, 0)$ or $(\alpha^*, \beta^*) \equiv (0, -1/d)$. Then, for any fixed $\varepsilon \in (0, 1)$,*

$$(2.7) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left(g_1(\varepsilon) \leq \frac{N^{(\star)}(H; M_\varepsilon^{(\alpha^*, \beta^*)})}{n^{f_H(\alpha^*, \beta^*)}} \leq g_2(\varepsilon) \right) = 1$$

for some positive functions g_1, g_2 not depending on n . Moreover, if (α^*, β^*) is the unique solution of the optimization problem, then for any $(\alpha, \beta) \in \mathcal{F} \setminus \{(\alpha^*, \beta^*)\}$ and $\varepsilon \in (0, 1)$,

$$\frac{N^{(\star)}(H; M_\varepsilon^{(\alpha, \beta)})}{N^{(\star)}(H; M_\varepsilon^{(\alpha^*, \beta^*)})} \xrightarrow{\mathbb{P}} 0.$$

The proof of this result can be found in Section 3. Observe that, under the hypothesis of Theorem 2.5, these particular subgraph counts are dominated by the contribution over set of vertices either with degree $\propto \sqrt{n}$ or at distance $\propto n^{1/d}$ from each other. Therefore, in the first case the leading contribution emerges regardless of geometry and is caused by the power-law weights. In the second case, the spatial embedding of the vertices plays a fundamental role, but the power-law weights do not. For this reason, we identify two distinct phases that we call *geometric* and *non-geometric*.

Figures 2, 3, 6 and 7 show that the solution to the optimization problem for general subgraphs is uniquely achieved on $(\alpha, \beta) \equiv (1/2, 0)$ or $(0, -1/d)$ for a large class of subgraphs (all subgraphs with only light gray nodes and dark blue edges, or only orange nodes and yellow edges). This class contains mostly relatively dense subgraphs, and contains for example all cliques. The same happens for induced subgraphs, as shown in Figures 4, 5, 8 and 9. In particular, these dense subgraphs manifest a geometric phase when τ is large, whereas a non-geometric phase for a smaller value τ .

2.2. *Total subgraph counts.* The results in the previous section focus on the occurrence of H on sets of the kind $M_\varepsilon^{(\alpha, \beta)}$, given $(\alpha, \beta) \in \mathcal{F}$. We now deal with the total number of occurrences $N^{(\text{sub})}(H)$ and $N^{(\text{ind})}(H)$. First, an immediate application of Markov inequality shows that the number of occurrences of any H in $\text{GIRG}^{(n)}$ grows at least linearly or super-linearly with n (depending on k and τ), with high probability:

COROLLARY 2.6. *Let H be a subgraph of size k and $\star \in \{\text{ind}, \text{sub}\}$.*

- *When $k \leq \frac{2}{3-\tau}$, there exist $\delta > 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{N^{(\star)}(H)}{n} \leq \delta \right) = 0.$$

- *When $k \geq \frac{2}{3-\tau}$, there exist $\delta > 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{N^{(\star)}(H)}{n^{k(3-\tau)/2}} \leq \delta \right) = 0.$$

PROOF. Observe that $N^{(\star)}(H) \geq N^{(\star)}(H, M_\varepsilon^{(\alpha, \beta)})$ for any $(\alpha, \beta) \in \mathcal{F}$. Moreover, $(\alpha, \beta) \equiv (1/2, 0)$ and $(\alpha, \beta) \equiv (0, -1/d)$ are always in the feasible region for both optimization problems (2.4) and (2.4)–(2.5). Computing the exponent (2.3) yields:

$$\begin{aligned} f_H(\alpha, \beta) &= (3 - \tau)k/2, & \text{if } (\alpha, \beta) \equiv (1/2, 0) \\ f_H(\alpha, \beta) &= 1, & \text{if } (\alpha, \beta) \equiv (0, -1/d) \end{aligned}$$

By (2.7), there exists a function g_1 such that $\mathbb{P} \left(N^{(\star)}(H; M_\varepsilon^{(\alpha, \beta)}) \leq g_1(\varepsilon)n^{f_H(\alpha, \beta)} \right) \rightarrow 0$, as $n \rightarrow \infty$. Hence, we conclude that there exists $\delta > 0$ such that

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(N^{(\star)}(H) \leq \delta n^{\max(1, (3-\tau)k/2)} \right) = 0.$$

□

Next, we derive an upper and a lower bound for the expected number of copies of H in $\text{GIRG}^{(n)}$:

THEOREM 2.7. *Let H be a subgraph of fixed size, let $\star \in \{\text{ind}, \text{sub}\}$ and let f_H^* be the optimal value of the optimization problem (2.4) (if $\star = \text{sub}$) or (2.4)+(2.5) (if $\star = \text{ind}$). Then*

$$(2.8) \quad \liminf_{n \rightarrow \infty} \frac{\mathbb{E} [N^{(\star)}(H)]}{n^{f_H^*}} > 0.$$

Moreover, conditionally on \mathcal{E}_ε defined in (1.4), for any $\delta > 0$,

$$(2.9) \quad \limsup_{n \rightarrow \infty} \frac{\mathbb{E} [N^{(\star)}(H) \mid \mathcal{E}_\varepsilon]}{n^{f_H^* + \delta}} < \infty.$$

In other words, Theorem 2.7 states that the expected subgraph count in $\text{GIRG}^{(n)}$ scales polynomially in the graph size, where the exponent is the solution of (2.4). Observe that from (2.9) it also follows that, for any $\delta > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(N^{(\star)}(H) > n^{f_H^* + \delta} \right) = 0.$$

from direct application of Markov inequality.

2.3. *Results on particular subgraph classes.* While the previous results hold in general for any H of fixed size, we can prove sharper convergence for specific classes of subgraphs.

First we focus on the case when H is a tree, in which case the number of copies of H as general subgraph in GIRGs shows the same asymptotic behaviour as in non-geometric scale free random graphs. In particular, it scales the same as in the inhomogeneous random graph (IRG) where the connection probability (1.2) is replaced by $\tilde{p}_{uv} = \Theta\left(1 \wedge \frac{w_u w_v}{\mu n}\right)$. We denote by $\tilde{N}^{(\text{sub})}(H)$ the number of copies of H in $\text{IRG}^{(n)}$ as general subgraph, for which the behavior of subgraph counts have been estimated in [34].

THEOREM 2.8. *If H is a tree, then there exist $m, M > 0$ such that*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(m \leq \frac{\tilde{N}^{(\text{sub})}(H)}{N^{(\text{sub})}(H)} \leq M \right) = 1.$$

The intuition behind this result is that the spatial embedding does not significantly influence the occurrence of pattern not containing cycles (i.e., trees). On the contrary, subgraphs that contain a Hamiltonian cycle undergo a geometric phase transition. Let $\mathcal{H}_k = ([k], \mathcal{E}_{\mathcal{H}_k})$ be the Hamiltonian cycle with edge set $\mathcal{E}_{\mathcal{H}_k} = \{(i, i+1)\}_{i \in [k]}$ using the convention that vertex $k+1$ coincides with vertex 1. We say that H is a Hamiltonian pattern, if there exist an automorphism of H that contains \mathcal{H}_k .

THEOREM 2.9. *Let H be a Hamiltonian pattern of size k . Let $\star \in \{\text{ind}, \text{sub}\}$. Then there exist constants $I_G^{(\star)}(H), I_N^{(\star)}(H) \in (0, \infty)$ depending on the parameters of the GIRG such that:*

(i) *When $k < 1 + \frac{1}{3-\tau}$, the optimization problem for H is solved by $\alpha \equiv 0, \beta \equiv -\frac{1}{d}$, and*

$$(2.10) \quad \frac{N^{(\star)}(H)}{n} \xrightarrow{\mathbb{P}} I_G^{(\star)}(H).$$

(ii) *When $k > \frac{2}{3-\tau}$ and k is odd, the optimization problem for H is solved by $\alpha \equiv 1/2, \beta \equiv 0$, and*

$$(2.11) \quad \frac{N^{(\star)}(H)}{n^{k(3-\tau)/2}} \xrightarrow{\mathbb{P}} I_N^{(\star)}(H).$$

2.4. *MILP reformulation.* We now provide a strategy to easily solve the optimization problem (2.4) or (2.4)+(2.5) by reformulating it as a mixed-integer linear optimization problem. Furthermore, we will show that this formulation will allow to check whether a particular solution is unique, checking the assumption of Theorem 2.5.

We first replace every minimum in (2.4) by an auxiliary variable that is bounded from above by the two expressions in the minimum. These variables are ζ_j for $\min_{i:i < j} \beta_{\{i,j\}}$ as well as $\delta_{\{i,j\}}$ for $\min(\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1, 0)$. For the maximum in (2.4d) we introduce a binary decision variable $z_{i,j,s}$ that indicates which of the constraints $\beta_{\{i,j\}} \leq \beta_{\{i,s\}}$ and $\beta_{\{i,j\}} \leq \beta_{\{j,s\}}$ must hold. If that variable equals 1, then (2.12h) and (2.12i) simplify to $\beta_{\{i,j\}} \leq \beta_{\{i,s\}}$ and $\beta_{\{i,j\}} \leq \beta_{\{j,s\}} + 1/d$, where the latter is redundant due to the bounds on the β -variables. Similarly, if $z_{i,j,s} = 0$ then (2.12h) and (2.12i) simplify to $\beta_{\{i,j\}} \leq \beta_{\{i,s\}} + 1/d$ and $\beta_{\{i,j\}} \leq \beta_{\{j,s\}}$, where the former is redundant. The resulting MILP reads as follows.

$$(2.12a) \quad \max (1 - \tau) \sum_{i=1}^k \alpha_i + d \sum_{j=1}^k \zeta_j + \gamma \sum_{\{i,j\} \in \mathcal{E}} \delta_{\{i,j\}}$$

$$\begin{aligned}
(2.12b) \quad & \text{over} \quad \alpha_i \in [0, 1/(\tau - 1)] && \forall i \in [k] \\
(2.12c) \quad & \beta_{\{i,j\}} \in [-1/d, 0] && \forall i, j \in [k] : i < j \\
(2.12d) \quad & \zeta_j \in [-1/d, 0] && \forall j \in \{2, 3, \dots, k\} \\
(2.12e) \quad & \delta_{\{i,j\}} \in [-1, 0] && \forall \{i, j\} \in \mathcal{E} \\
(2.12f) \quad & z_{i,j,s} \in \{0, 1\} && \forall i, j, s \in [k] : i < j, s \neq i, s \neq j \\
(2.12g) \quad & \text{s.t.} \quad \zeta_j \leq \beta_{\{i,j\}} && \forall i, j \in [k] : i < j \\
(2.12h) \quad & \beta_{\{i,j\}} \leq \beta_{\{i,s\}} + (1 - z_{i,j,s})/d && \forall i, j, s \in [k] : i < j, s \neq i, s \neq j \\
(2.12i) \quad & \beta_{\{i,j\}} \leq \beta_{\{j,s\}} + z_{i,j,s}/d && \forall i, j, s \in [k] : i < j, s \neq i, s \neq j \\
(2.12j) \quad & \delta_{\{i,j\}} \leq \alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1 && \forall \{i, j\} \in \mathcal{E}
\end{aligned}$$

Theorems 2.4–2.5 show that for subgraphs for which the optimizer of (2.12) (resp. (2.12)+(2.5)) is unique in terms of its (α, β) variables, there is a specific type of dominant subgraph structure. That is, a copy of H forms most likely on vertices of weights α and inter-distances β .

Hence, we are not just interested in an optimum solution of (2.12) or (2.12)+(2.5), but also wish to know if the determined optimum solution is unique. For linear optimization problems, sensitivity analysis can be used to assert uniqueness. However, here we deal with a MILP and we only care about the uniqueness after projecting the optimal solution onto the α - and β -variables. Therefore, we use the following approach.

First, we solve (2.12) or (2.12)+(2.5) and obtain the optimizer (α^*, β^*) with objective value z^* , restricting to only optimal solutions of the MILP. Second, we add the constraint

$$(2.13) \quad (1 - \tau) \sum_{i=1}^k \alpha_i + d \sum_{j=1}^k \zeta_j + \gamma \sum_{\{i,j\} \in \mathcal{E}} \delta_{\{i,j\}} = z^*$$

to the MILP. Third, for every $i \in [k]$ and each $\sigma \in \{\pm 1\}$ we maximize $\sigma \alpha_i$ subject to (2.12b)–(2.12j) and (2.13) by replacing (2.12a) by the new objective $\sigma \in \{\pm 1\}$. Fourth, for every $i, j \in [k] : i < j$ and each $\sigma \in \{\pm 1\}$ we maximize $\sigma \beta_{\{i,j\}}$ subject to (2.12b)–(2.12j) and (2.13). Constraint (2.13) ensures that every feasible solution to any of these MILPs is optimal for (2.12) (resp. (2.12)+(2.5)). Finally, we conclude that (α^*, β^*) is unique if and only if (α^*, β^*) is also the optimizer of all modified optimization problems.

If any such solution (α', β') is different from (α^*, β^*) then the optimal solution of (2.12) is not unique. Furthermore, if there exists an optimal solution of (2.12) different from (α^*, β^*) then it differs in at least one coordinate, where it is either smaller or larger. Such a solution would constitute a feasible solution to each of these modified optimization problems. Furthermore, the difference in at least one coordinate implies that at least one i or i, j has a higher value of $\sigma \alpha_i$ or $\sigma \beta_{\{i,j\}}$ than (α^*, β^*) . Consequently, in at least one of these modified problems (α^*, β^*) will not be optimal with respect to the modified objective. We conclude that the described approach correctly asserts uniqueness of (α^*, β^*) or outputs at least one alternative optimal solution.

2.5. Numerical results. We implemented (2.12) and (2.12)+(2.5) as well as the approach for testing uniqueness using the open-source MILP-solver SCIP [4] through its Python interface [17].

These results paint an interesting picture of the typical occurrences of all subgraphs of sizes 4 and 5. For example, for small τ , dense subgraphs are often formed by \sqrt{n} -degree

nodes that are far apart, while for larger values of τ , dense subgraphs are often formed by low-degree vertices that are close in location. For less dense subgraphs, the picture is very subgraph-dependent, with some hubs and some lower-degree vertices, and also differences in the typical distances. Also in these sparser subgraphs, a phase transition in τ seems to appear, but the change in distances and weights of the involved vertices is more complex.

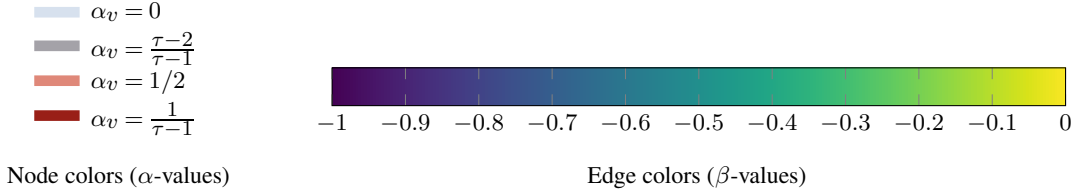


Fig 1: Legend for other figures. Coloring of nodes and edges according to α - and β -values, respectively. For subgraph plots, dashed edges indicate edges that are not in the graph, while for induced subgraphs plots, these are drawn as dotted edges. In the figures, gray indicates that there also exist different optimal solutions.

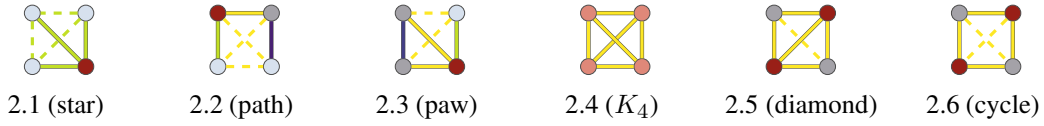


Fig 2: Optimal **subgraph** structures of all subgraphs with four nodes for **small** $\tau \in (2, 2.5)$, $d = 1$ and $\gamma = 2$. The legend is depicted in Figure 1.

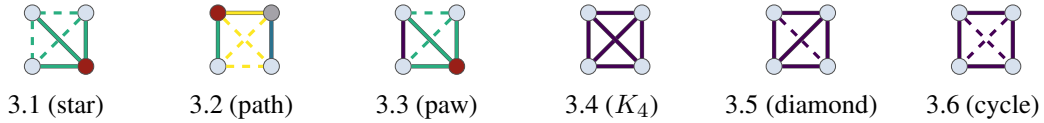


Fig 3: Optimal **subgraph** structures of all subgraphs with four nodes for **large** $\tau \in (2.5, 3)$, $d = 1$ and $\gamma = 2$. The legend is depicted in Figure 1.

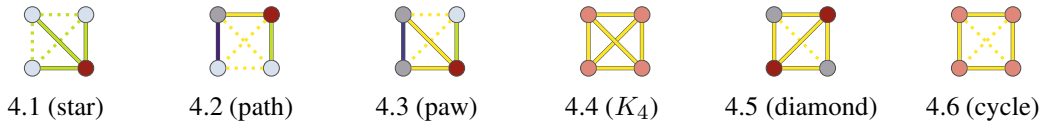


Fig 4: Optimal **induced subgraph** structures of all subgraphs with four nodes for **small** $\tau \in (2, 2.5)$, $d = 1$ and $\gamma = 2$. The legend is depicted in Figure 1.

2.6. Discussion.

Number of subgraphs outside the feasible region. In Section 2.1 we derived the number of subgraphs that appear on sets of the kind $M_\varepsilon^{(\alpha, \beta)}$ in $\text{GIRG}^{(n)}$, where $(\alpha, \beta) \in \mathcal{F}$ is in the feasible region defined in (2.2). As already observed, whenever the variables (α, β) are chosen outside \mathcal{F} , the corresponding set $M_\varepsilon^{(\alpha, \beta)}$ is empty with high probability. Thus, in this case the random variable $N^{(*)}(H; M_\varepsilon^{(\alpha, \beta)})$ converges to 0 in probability as n grows large. However, its expectation does not necessarily converge to 0 and, in fact, it could grow much larger

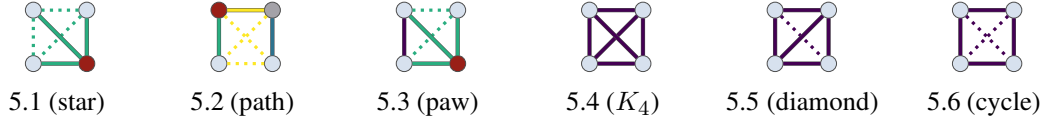


Fig 5: Optimal **induced subgraph** structures of all subgraphs with four nodes for **large** $\tau \in (2.5, 3)$, $d = 1$ and $\gamma = 2$. The legend is depicted in Figure 1.

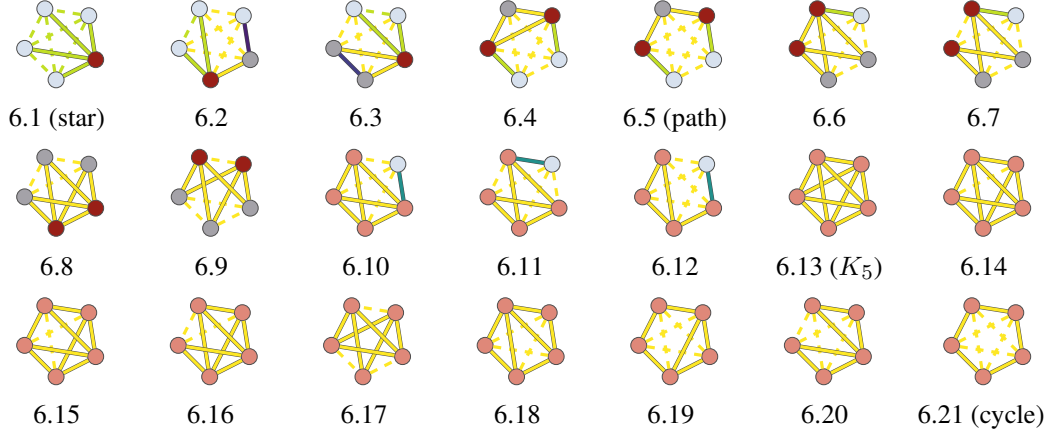


Fig 6: Optimal **subgraph** structures of all subgraphs with five nodes for **small** $\tau \in (2, 2.6)$, $d = 1$ and $\gamma = 2$. The legend is depicted in Figure 1.

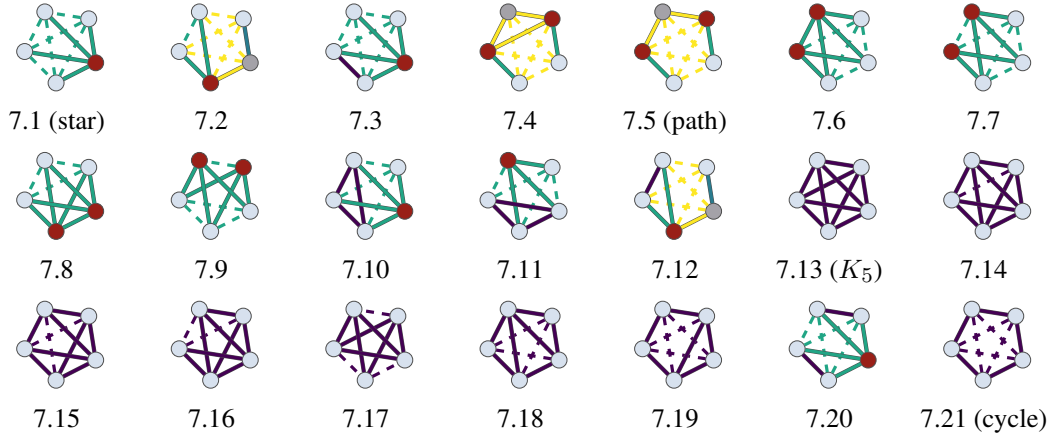


Fig 7: Optimal **subgraph** structures of all subgraphs with five nodes for **large** $\tau \in (2.6, 3)$, $d = 1$ and $\gamma = 2$. The legend is depicted in Figure 1.

than for any other feasible $(\alpha, \beta) \in \mathcal{F}$. One example is the wedge pattern $W = (\mathcal{V}, \mathcal{E})$, with $\mathcal{V} = \{1, 2, 3\}$, $\mathcal{E} = \{(1, 2), (2, 3)\}$, for which it is possible to show that with high probability $N^{(*)}(H)$ is bounded by $o(n^{2/(\tau-1)+\delta})$ for arbitrary $\delta > 0$, whereas $\mathbb{E}[N^{(*)}(H)] = \Omega(n^{4-\tau})$ due to the infinite-variance weights. This means that concentration results such as in Theorem 2.5 or Theorem 2.9 do not hold in general for all H . Therefore, investigating the limit of these non-concentrating subgraphs more closely could be interesting for further research.

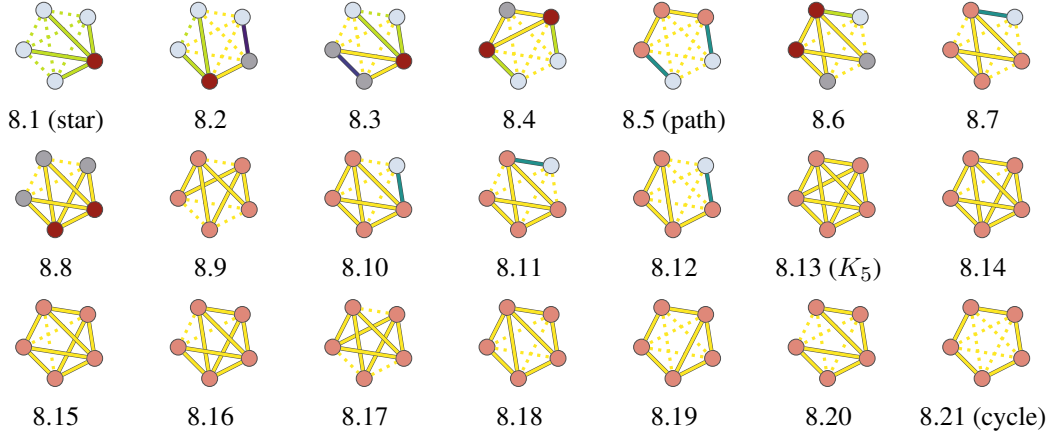


Fig 8: Optimal **induced subgraph** structures of all subgraphs with five nodes for **small** $\tau \in (2, 2.6)$, $d = 1$ and $\gamma = 2$. The legend is depicted in Figure 1.

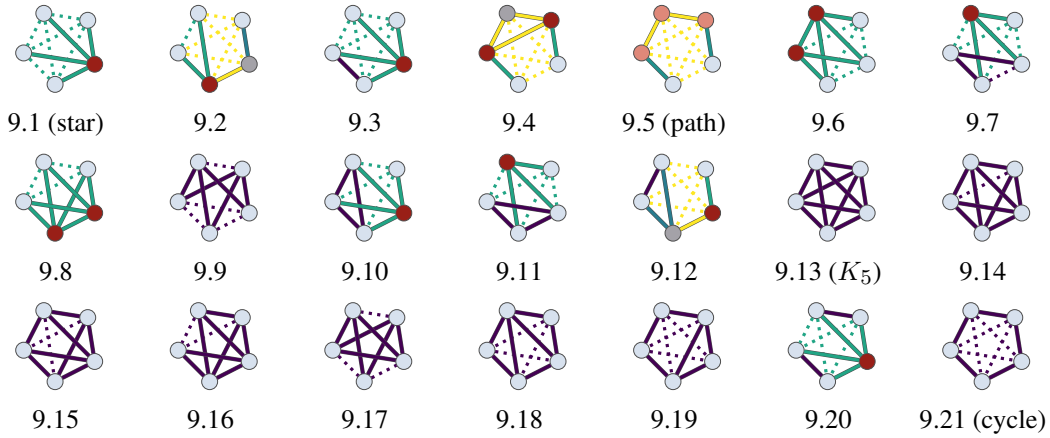


Fig 9: Optimal **induced subgraph** structures of all subgraphs with five nodes for **large** $\tau \in (2.6, 3)$, $d = 1$ and $\gamma = 2$. The legend is depicted in Figure 1.

Phase transition on τ . Theorem 2.9 shows that the occurrence of any Hamiltonian pattern H falls into one of two distinct geometric/non-geometric regimes, linking the pattern size and the power-law exponent. If τ is large, H appears more frequently on low-degree vertices at small distances; when τ is small, then hubs of weight order \sqrt{n} produce a large number of copies of H . Figures 2,3 and 6, 7 illustrate this phase transition for subgraphs of sizes $k = 4$ and 5. Often, the unique optimum is achieved at $(1/2, 0)$ or $(0, -1/d)$ (depending on τ) by the most dense subgraphs. When the maximum is achieved on $(1/2, 0)$, it implies that, surprisingly, the way these subgraph form is not particularly affected by the geometry of the random graph. Theorem 2.9 contains a gap for the geometric/non-geometric transition between $(1/2, 0)$ and $(0, -1/d)$ when $k \in [1 + \frac{1}{3-\tau}, \frac{2}{3-\tau}]$. We conjecture that the phase transition is sharply separated by the curve $k = 2/(3 - \tau)$, as this would match existing results for clique counts [18].

Values of α and β . When $k \neq 2/(3 - \tau)$, that is outside the critical window, we conjecture that whenever there exist unique optimal α -values of (2.4) or (2.4)+(2.5) they take value in a finite set, $\alpha_i \in \{0, \frac{\tau-2}{\tau-1}, \frac{1}{2}, \frac{1}{\tau-1}\}$. Experimentally, Figures 2–5 and 6–9 also show that this

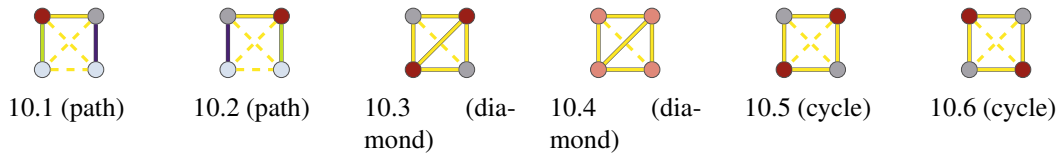


Fig 10: **Several different optimal subgraph** structures of all subgraphs with four nodes for **small** $\tau \in (2, 2.5)$, $d = 1$ and $\gamma = 2$ for which the optimal solutions was not unique. The legend is depicted in Figure 1.

is the case for subgraphs up to size 5 (results for subgraphs of size 6 are not included in this paper, but are also consistent. The values 0 and $1/(\tau - 1)$ result from the domain constraint of alpha to be in $[0, 1/(\tau - 1)]$. As (2.4) optimizes a piecewise linear function, the optimizers over α are at the end points of the interval $[0, 1/(\tau - 1)]$ or at the points where $\alpha_i + \alpha_j - d\beta_{ij} = 1$. This still leaves many possible options for α , but we conjecture that the optimizers that are caused by these constraints only have $\alpha = 1/2$ or $\alpha = (\tau - 2)/(\tau - 1)$. Proving this is not immediate unfortunately, as the triangle inequality for the β values makes it difficult to make small perturbations to optimal solutions. Under this conjecture, the optimal β -values would also only be given by 0, $-1/d$ (end points), and values where $\alpha_i + \alpha_j - d\beta_{ij} = 1$. For the conjectured values of α , this would result in $d \cdot \beta_i \in \left\{ -1, \frac{2-\tau}{\tau-1}, -\frac{1}{2}, -\frac{1}{\tau-1}, \frac{\tau-3}{\tau-1}, \frac{\tau-3}{2\tau-2}, 0 \right\}$.

Unique optimizers and predominant subgraphs. The results of Corollary 2.4 and Theorem 2.5 hold for subgraphs with unique optimizers of (2.4) or (2.4)+(2.5) is unique. This is the case for almost all subgraphs on 4 or 5 vertices. However, Figure 10 shows that non-unique optimizers are usually caused by symmetries in the subgraph. In that case, there is still a range of optimal subgraph structures, but rather than being defined on the vertices, they are defined on the edges. For example, on the cycle of size 4 for $\tau < 5/2$, an optimal solution satisfies that the product of the vertex weights incident to any edge is of order n . Due to symmetry, this can arise in many different ways, and the extremal options are shown in Figure 10, but in fact, the number of optimal solutions is infinite.

The reason that uniqueness of the optimizer is so important is the following. The optimal value of (2.4) or (2.4)+(2.5) identifies the order of magnitude of subgraph counts within a specific region. Theorems 2.4–2.5 then show that there exist optimal region(s), defined by the pair (α^*, β^*) , maximizing the number of subgraphs in it. However, uniqueness of the optimizer does not immediately imply that the subgraphs in such maximal regions dominate all other regions combined. Practically, to demonstrate that subgraphs in the optimal region dominate all others, we need to show that for arbitrary $p_0 \in (0, 1)$ there exists $\varepsilon > 0$ such that

$$(2.14) \quad \lim_{n \rightarrow \infty} \mathbb{P} \left(\mathbf{v} \in M_\varepsilon^{(\alpha^*, \beta^*)} \mid H \subseteq \text{GIRG}^{(n)} \mid \mathbf{v} \right) \geq p_0.$$

While this is not true in general, it is not difficult to show that (2.14) holds under the hypothesis of Theorem 2.9, with a slight adaptation of its proof in Section 4.

Cycles of even size. For Hamiltonian patterns of even size, Theorem 2.9(ii) does not hold. The main reason is that for cycles of even size k the solution to (2.4) is not unique. Indeed, in this case, two mirrored extreme optima appear, see Figure 10. The convex combination of these two solution generates an infinite set of solutions. Introducing extra edges or considering the induced optimization problem may break the symmetry and result in a unique solution for the optimization problem. In this case, a more refined strategy is necessary to obtain similar results as in (2.11).

Efficient subgraph counting algorithms. A very rich line of research deals with the design of efficient algorithms to count or approximate subgraphs in networks [14, 28, 24]. The knowledge of optimal set of vertices where subgraphs form more frequently may be used to improve significantly the efficiency of approximating algorithms. Indeed, one could use the following procedure: first, compute the (possibly unique) solution (α^*, β^*) of the general or induced optimization problem for a given pattern H ; next, generate the induced graph over the set of vertices with degrees in $\cup_{\alpha \in \alpha^*} [\varepsilon n^\alpha, n^\alpha / \varepsilon]$, for some sensitivity parameter $\varepsilon \in (0, 1)$; finally, run the algorithm for counting the occurrence of H on the induced graph. As an example, consider the case when the optimum is uniquely achieved for $\alpha^* \equiv 1/2$. In this case, the size of the induced graph on vertices with order degree $\sim \sqrt{n}$ is $\Theta(n^{(3-\tau)/2})$; then, the number of edges in the induced graph is $O(n^{3-\tau})$, which is sub-linear with the size of the initial graph. Consequently, it is possible to define an approximated version of any node-based counting algorithm for H improving the running time by $\Omega(n^{(\tau-1)/2})$, or of any edge-based counting algorithm improving the running time by $\Omega(n^{\tau-2})$. Moreover, the approximation error can be made arbitrarily small by tuning the parameter ε , at increased computational cost.

3. Proofs for generic subgraphs. Here we demonstrate results of Sections 2.1–2.2. First we prove Proposition 2.3, showing that the parameter γ does not play any role in the optimization problem 2.4. Next, we prove the central Lemma 2.1, showing that the subgraph counts over sets of the form $M_\varepsilon^{(\alpha, \beta)}$ are particularly easy to compute. To demonstrate Theorem 2.5 we need to show that the subgraph counts are self-averaging random variables where we use a second moment method. Finally, we will prove upper and lower bounds for the total number of subgraphs in Theorem 2.7.

PROOF OF PROPOSITION 2.3. Suppose that $\alpha_i + \alpha_j < 1 + d\beta_{\{i,j\}}$ for some i, j . We will show that the objective function f_H defined by (2.4a) achieves a higher value when we decrease $\beta_{\{i,j\}}$ by a sufficiently small quantity $x > 0$. When decreasing $\beta_{\{i,j\}}$, we need to ensure that all constraints given by the triangular inequality are still satisfied. To do so, we apply the following policy that keeps track of what β_{ik} or β_{jk} values are the same as β_{ij} , while one other distance in the triple $\{i, j, k\}$ is larger. This would imply that decreasing β_{ij} can only be feasible when also decreasing β_{ik} or β_{jk} :

1. Start by declaring all edges (a, b) with $a, b \in [k]$ inactive.
2. Activate edge (i, j) and consider all (unordered) triplets of distinct elements (i, j, h) with $h \in [k]$:
 - If $\beta_{\{i,j\}} \leq \beta_{\{i,h\}} = \beta_{\{j,h\}}$, keep (i, h) and (j, h) inactive.
 - If $\beta_{\{i,h\}} < \beta_{\{i,j\}} = \beta_{\{j,h\}}$, switch (j, h) to active, and keep (i, h) inactive.
 - If $\beta_{\{j,h\}} < \beta_{\{i,j\}} = \beta_{\{i,h\}}$, switch (i, h) to active, and keep (j, h) inactive.
3. For each newly activated edge (i', j') , repeat the operation in step 2, considering all triplets (i', j', h) .

By construction, decreasing β with a sufficiently small amount $x > 0$ over all active edges is feasible with respect to the triangular inequality.

W.l.o.g. assume $\alpha_1 + \alpha_2 < 1 + d\beta_{\{1,2\}}$. Decrease the values in β over the activated edges by $x > 0$ sufficiently small, accordingly to the activation policy described above, starting from edge $(1, 2)$. Observe that, for each $t \geq 3$, there are exactly two possibilities:

- $\min_{s < t} \beta_{\{s,t\}} < \beta_{\{1,2\}}$. Let s' be a vertex that achieves the minimum. Then, the policy keeps the edge (s', t) inactive, as it only activates edges with β value equal to $\beta_{\{1,2\}}$;
- $\min_{s < t} \beta_{\{s,t\}} \geq \beta_{\{1,2\}}$. Denote by (s', t) the edge achieving the minimum. In this case, observe that vertex t is at least at distance $\Omega(n^{\beta_{\{1,2\}}})$ from edge $(1, 2)$. In turn, it follows from our activation policy that t is at distance $\Omega(n^{\beta_{\{1,2\}}})$ from any other activated edge. In particular, this implies that the edge (s', t) will never be activated.

In other words, whenever we decrease $\beta_{\{1,2\}}$, the triangular inequality constraints are satisfied and the quantities $\min_{s<t} \beta_{\{s,t\}}$ are kept fixed, when applying our decreasing policy.

Finally, observe that when decreasing $\beta_{\{1,2\}}$ (and eventually other $\beta_{\{s,t\}}$ for some $s, t \in [k]$) by $x > 0$, the third sum in the objective function f_H increases at least by γdx . On the other hand, among the $k - 1$ terms in the second sum of the objective function f_H , only the first (corresponding to index $j = 2$) decreases by a quantity dx , whereas all the others remain unchanged. Therefore, decreasing $\beta_{\{1,2\}}$ increases the value of f_H by $(\gamma - 1)dx > 0$. \square

PROOF OF LEMMA 2.1. The definitions in (1.8) yield

$$\mathbb{E}[N^{(\text{ind})}(H; M_\varepsilon^{(\alpha, \beta)})] = \sum_{\mathbf{v} \in V_k} \mathbb{P}(H = \text{GIRG} | \mathbf{v} | \mathbf{v} \in M_\varepsilon^{(\alpha, \beta)}) \mathbb{P}(\mathbf{v} \in M_\varepsilon^{(\alpha, \beta)})$$

$$\mathbb{E}[N^{(\text{sub})}(H; M_\varepsilon^{(\alpha, \beta)})] = \sum_{\mathbf{v} \in V_k} \mathbb{P}(H \subseteq \text{GIRG} | \mathbf{v} | \mathbf{v} \in M_\varepsilon^{(\alpha, \beta)}) \mathbb{P}(\mathbf{v} \in M_\varepsilon^{(\alpha, \beta)}).$$

Given $\mathbf{v} \in M_\varepsilon^{(\alpha, \beta)}$, each vertex $v_i \in \mathbf{v}$ has weight in $I_\varepsilon(n^{\alpha_i})$ with a probability which is proportional to $n^{(1-\tau)\alpha_i}$ due to (1.1), independently of all the other vertices in \mathbf{v} . The probability that the inter-distance between any two selected vertices $v_i, v_j \in \mathbf{v}$ falls in $I_\varepsilon(n^{\beta_{\{i,j\}}})$ is slightly more complicated. We calculate this probability starting from vertex v_1 , which is w.l.o.g. given a fixed location on the torus. Next, consider v_2 : the region in which the distance between v_2 and v_1 is in $I_\varepsilon(n^{\beta_{\{1,2\}}})$ is proportional to $n^{d\beta_{\{1,2\}}}$, as the positions are sampled uniformly. We then proceed to vertex v_3 . The distance from v_3 to v_1 and v_2 should be inside $I_\varepsilon(n^{\beta_{\{1,3\}}})$ and $I_\varepsilon(n^{\beta_{\{2,3\}}})$, hence the area in which v_3 satisfies this is proportional to $\min(n^{d\beta_{\{1,3\}}}, n^{d\beta_{\{2,3\}}})$. Iterating this procedure, combined with the fact that positions are sampled uniformly on the torus, yields that the probability that the vertex set $\mathbf{v} \in M_\varepsilon^{(\alpha, \beta)}$ is proportional to

$$(3.1) \quad \mathbb{P}(\mathbf{v} \in M_\varepsilon^{(\alpha, \beta)}) \propto \prod_{i \in [k]} n^{(1-\tau)\alpha_i} \prod_{j \in [k]} \min_{i:i < j} n^{d\beta_{\{i,j\}}}.$$

Furthermore, since $v_1, \dots, v_k \in M_\varepsilon^{(\alpha, \beta)}$, the connection probability between v_i and v_j is proportional to $\min(n^{\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1}, 1)^\gamma$. This yields

$$(3.2) \quad \mathbb{P}(H \subseteq \text{GIRG} | \mathbf{v} | \mathbf{v} \in M_\varepsilon^{(\alpha, \beta)}) \propto \prod_{i,j:(i,j) \in \mathcal{E}} \min(n^{\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1}, 1)^\gamma.$$

Then, (3.1) and (3.2), and the fact that \mathbf{v} can be chosen in $n!/(n-k)!$ ways, gives

$$\mathbb{E}[N^{(\text{sub})}(H; M_\varepsilon^{(\alpha, \beta)})] \propto n^{f_H(\alpha, \beta)},$$

where

$$f_H(\alpha, \beta) = k + (1-\tau) \sum_{i=1}^k \alpha_i + d \sum_{j=2}^k \min_{i:i < j} \beta_{\{i,j\}} + \gamma \sum_{i,j:(i,j) \in \mathcal{E}} \min(\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1, 0).$$

For $\mathbb{E}[N^{(\text{ind})}(H; M_\varepsilon^{(\alpha, \beta)})]$, we compute $\mathbb{P}(H = \text{GIRG} | \mathbf{v} | \mathbf{v} \in M_\varepsilon^{(\alpha, \beta)})$, similarly to (3.2) as

$$\begin{aligned} \mathbb{P}(H = \text{GIRG} | \mathbf{v} | \mathbf{v} \in M_\varepsilon^{(\alpha, \beta)}) &\propto \prod_{(i,j) \in \mathcal{E}} \min(n^{\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1}, 1)^\gamma \prod_{(i,j) \notin \mathcal{E}} (1 - \min(n^{\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1}, 1)^\gamma) \\ &\propto \prod_{(i,j) \in \mathcal{E}} \min(n^{\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1}, 1)^\gamma \prod_{(i,j) \notin \mathcal{E}} \max(1 - n^{\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1}, 0)^\gamma. \end{aligned}$$

In particular, whenever $\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1 > 0$ for any $(i, j) \notin \mathcal{E}$, $\mathbb{P}(H = \text{GIRG} \mid \mathbf{v}) = 0$. Indeed, if a pair $(i, j) \notin \mathcal{E}$ existed such that $\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1 > 0$, then the edge (v_i, v_j) would be present with probability one, by (1.2). On the other hand, if $\alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1 \leq 0$ for all $(i, j) \notin \mathcal{E}$, then the expected number of copies of H as induced subgraph is asymptotically equal to the expected number of copies of H as general subgraph and

$$\mathbb{E}[N^{(\text{ind})}(H; M_\varepsilon^{(\alpha, \beta)})] \propto \begin{cases} n^{f_H(\alpha, \beta)} & \text{if } \alpha_i + \alpha_j - d\beta_{\{i,j\}} - 1 \leq 0, \forall (i, j) \notin \mathcal{E}, \\ 0 & \text{otherwise} \end{cases}$$

follows, which concludes the proof. \square

Next, we prove that when the optimal solution is either $(\alpha^*, \beta^*) \equiv (1/2, 0)$ or $(\alpha^*, \beta^*) \equiv (0, -1/d)$, the convergence of expectations for the number of subgraphs in the optimal region can be sharpened to a convergence in probability. We first introduce an auxiliary lemma.

LEMMA 3.1. *Let H be a graph, $\star \in \{\text{ind}, \text{sub}\}$ and $(\alpha, \beta) \equiv (1/2, 0)$ or $(\alpha, \beta) \equiv (0, -1/d)$. Then,*

$$\lim_{n \rightarrow \infty} \frac{\text{Var}\left(N^{(\star)}(H; M_\varepsilon^{(\alpha, \beta)})\right)}{\mathbb{E}\left[N^{(\star)}(H; M_\varepsilon^{(\alpha, \beta)})\right]^2} = 0.$$

PROOF. For any fixed combination of k vertices \mathbf{v} we define the events

$$A_{\mathbf{v}} = \{H \subseteq \text{GIRG}^{(n)} \mid \mathbf{v}\}, \quad B_{\mathbf{v}} = \{\mathbf{v} \in M_\varepsilon^{(\alpha, \beta)}\},$$

so that $N^{\text{sub}}(H, M_\varepsilon^{(\alpha, \beta)}) = \sum_{\mathbf{v} \in V_k} \mathbb{1}_{\{A_{\mathbf{v}}, B_{\mathbf{v}}\}}$. Then, we write its variance as

$$\begin{aligned} \text{Var}\left(N^{\text{sub}}(H, M_\varepsilon^{(\alpha, \beta)})\right) &= \text{Var}\left(\sum_{\mathbf{v} \in V_k} \mathbb{1}_{\{A_{\mathbf{v}}, B_{\mathbf{v}}\}}\right) \\ &= \sum_{\mathbf{v}, \mathbf{u} \in V_k} \text{Cov}\left(\mathbb{1}_{\{A_{\mathbf{v}}, B_{\mathbf{v}}\}}, \mathbb{1}_{\{A_{\mathbf{u}}, B_{\mathbf{u}}\}}\right) \\ &= \sum_{\mathbf{v}, \mathbf{u} \in V_k} \mathbb{P}(A_{\mathbf{v}}, B_{\mathbf{v}}, A_{\mathbf{u}}, B_{\mathbf{u}}) - \mathbb{P}(A_{\mathbf{v}}, B_{\mathbf{v}})\mathbb{P}(A_{\mathbf{u}}, B_{\mathbf{u}}). \end{aligned}$$

Observe that if $\mathbf{v} \cap \mathbf{u} = \emptyset$, then the covariance between $\mathbb{1}_{\{A_{\mathbf{v}}, B_{\mathbf{v}}\}}$ and $\mathbb{1}_{\{A_{\mathbf{u}}, B_{\mathbf{u}}\}}$ is 0. Therefore, we restrict to the case $|\mathbf{v} \cap \mathbf{u}| = s$, with $s \geq 1$. Furthermore, for all pairs \mathbf{v}, \mathbf{u} we use the bound

$$\mathbb{P}(A_{\mathbf{v}}, B_{\mathbf{v}}, A_{\mathbf{u}}, B_{\mathbf{u}}) - \mathbb{P}(A_{\mathbf{v}}, B_{\mathbf{v}})\mathbb{P}(A_{\mathbf{u}}, B_{\mathbf{u}}) \leq \mathbb{P}(A_{\mathbf{v}}, B_{\mathbf{v}}, A_{\mathbf{u}}, B_{\mathbf{u}}) \leq \mathbb{P}(B_{\mathbf{v}}, B_{\mathbf{u}}).$$

Following the same computations as in the proof of Lemma 2.1 leading to Equation (3.1), and assuming that \mathbf{v} and \mathbf{u} intersect in s elements, we obtain

$$\begin{aligned} \mathbb{P}(B_{\mathbf{v}}, B_{\mathbf{u}}) &= \Theta(n^{k(1-\tau)\alpha + (k-1)d\beta}) \cdot \Theta(n^{(k-s)(1-\tau)\alpha + (k-s)d\beta}) \\ &= \Theta(n^{(2k-s)(1-\tau)\alpha + (2k-s-1)d\beta}), \end{aligned}$$

where $(\alpha, \beta) = (1/2, 0)$ or $(\alpha, \beta) = (0, -1/d)$. Since there are in total $\Theta(n^{2k-s})$ ways to choose the vertices in \mathbf{v} and \mathbf{u} , we obtain

$$\begin{aligned} \sum_{\mathbf{v}, \mathbf{u}: |\mathbf{v} \cap \mathbf{u}| = s} \text{Cov}\left(\mathbb{1}_{\{A_{\mathbf{v}}, B_{\mathbf{v}}\}}, \mathbb{1}_{\{A_{\mathbf{u}}, B_{\mathbf{u}}\}}\right) &\leq \Theta(n^{2k-s}) \cdot \Theta(n^{(2k-s)(1-\tau)\alpha^* + (2k-s-1)d\beta^*}) \\ &= \Theta(n^{(2k-s) + (2k-s)(1-\tau)\alpha^* + (2k-s-1)d\beta^*}), \end{aligned}$$

and taking the sum over all possible values for s ,

$$\text{Var} \left(N^{\text{sub}}(H; M_\varepsilon^{(\alpha, \beta)}) \right) = O \left(\max_{s \in 1, \dots, k} n^{(2k-s)+(2k-s)(1-\tau)\alpha+(2k-s-1)d\beta} \right).$$

By Lemma 2.1, $\mathbb{E} \left[N^{\text{sub}}(H; M_\varepsilon^{(\alpha, \beta)}) \right] = \Omega(n^{k+k(1-\tau)\alpha+(k-1)d\beta})$. Therefore

$$\begin{aligned} \frac{\text{Var} \left(N^{\text{sub}}(H; M_\varepsilon^{(\alpha, \beta)}) \right)}{\mathbb{E} \left[N^{\text{sub}}(H; M_\varepsilon^{(\alpha, \beta)}) \right]^2} &= \frac{O(\max_{s \in 1, \dots, k} n^{(2k-s)+(2k-s)(1-\tau)\alpha+(2k-s-1)d\beta})}{\Omega(n^{2k+2k(1-\tau)\alpha+2(k-1)d\beta})} \\ &= O \left(\max_{s \in 1, \dots, k} n^{-s-s(1-\tau)\alpha+(1-s)d\beta} \right). \end{aligned}$$

Now, $g(s) = -s - s(1-\tau)\alpha + (1-s)d\beta$ is negative for all $s \geq 1$, as

- when $(\alpha, \beta) = (0, -1/d)$ then $g(s) = -s - (1-s) = -1$,
- when $(\alpha, \beta) = (1/2, 0)$ then $g(s) = -s - s(1-\tau)/2 = -s(3-\tau)/2$.

This proves the claim. In the case of induced subgraph, the proof works by replacing N^{sub} with N^{ind} and defining instead $A_v := \{H = \text{GIRG}^{(n)}|_v\}$ at the beginning of the proof. \square

PROOF OF THEOREM 2.5. The Chebyshev inequality states that for any random variable X with finite variance and $\delta > 0$,

$$\mathbb{P} \left(\left| \frac{X}{\mathbb{E}[X]} - 1 \right| \geq \delta \right) \leq \frac{\text{Var}(X)}{\delta^2 \mathbb{E}[X]^2}.$$

Combining this with Lemma 3.1 yields that for any $\delta > 0$

$$\mathbb{P} \left(\left| \frac{N^{(*)}(H; M_\varepsilon^{(\alpha^*, \beta^*)})}{\mathbb{E} \left[N^{(*)}(H; M_\varepsilon^{(\alpha^*, \beta^*)}) \right]} - 1 \right| \geq \delta \right) \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

Combining this with Lemma 2.1 proves (2.7).

Next, let $(\alpha, \beta) \in \mathcal{F} \setminus \{(\alpha^*, \beta^*)\}$. Lemma 2.1 shows that there exists a function h_2 such that

$$\frac{\mathbb{E} \left[N^{(*)}(H; M_\varepsilon^{(\alpha, \beta)}) \right]}{n^{f_H(\alpha, \beta)}} \leq h_2(\varepsilon).$$

Then the Markov inequality yields

$$(3.3) \quad \begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left(N^{(*)}(H; M_\varepsilon^{(\alpha, \beta)}) \geq \delta n^{f_H(\alpha^*, \beta^*)} \right) &\leq \lim_{n \rightarrow \infty} \frac{\mathbb{E} \left[N^{(*)}(H; M_\varepsilon^{(\alpha, \beta)}) \right]}{\delta n^{f_H(\alpha^*, \beta^*)}} \\ &\leq \lim_{n \rightarrow \infty} \frac{h_2(\varepsilon) n^{f_H(\alpha, \beta)}}{\delta n^{f_H(\alpha^*, \beta^*)}} = 0, \end{aligned}$$

for any $\delta > 0$. Moreover, by (2.7), there exists a function g_1 such that

$$(3.4) \quad \mathbb{P} \left(N^{(*)}(H; M_\varepsilon^{(\alpha^*, \beta^*)}) \geq g_1(\varepsilon) n^{f_H(\alpha^*, \beta^*)} \right) \rightarrow 1, \quad \text{as } n \rightarrow \infty.$$

Therefore, combining (3.3) and (3.4), we obtain

$$\mathbb{P} \left(\frac{N^{(*)}(H; M_\varepsilon^{(\alpha, \beta)})}{N^{(*)}(H; M_\varepsilon^{(\alpha^*, \beta^*)})} \geq \delta \right) \rightarrow 0, \quad \text{as } n \rightarrow \infty,$$

which concludes the proof of the theorem. \square

Finally, we demonstrate the result for the total subgraph counts. The lower bound follows immediately from Lemma 2.1, whereas the upper bound involves building a proper covering of the feasible region \mathcal{F} .

PROOF OF THEOREM 2.7. Lower bound. The lower bound in (2.8) follows from the fact that, if (α^*, β^*) is a solution of the general or induced optimization problem, then

$$\mathbb{E} \left[N^{(*)}(H) \right] \geq \mathbb{E}_n \left[N^{(*)}(H; M_{\bar{\varepsilon}}^{(\alpha^*, \beta^*)}) \right] = \Omega(n^{f_H^*})$$

by Lemma 2.1.

Upper bound. To prove the upper bound in (2.9), we construct a covering for the feasible region. On \mathcal{E} , we have

$$N^{(*)}(H) = N^{(*)} \left(H; \bigcup_{(\alpha, \beta) \in \mathcal{F}} M_{\bar{\varepsilon}}^{(\alpha, \beta)} \right) \quad w.h.p.$$

for some $\bar{\varepsilon} > 0$, where \mathcal{F} is the feasible region defined in (2.2).

Covering $[n^a, n^b]$, for $a, b \in \mathbb{R}$ with intervals of the type $I_{\bar{\varepsilon}}(n^\zeta)$, with $\zeta \in [a, b]$ translates to covering $[a \log(n), b \log(n)]$ in logarithmic scale with intervals of the type $[\zeta \log(n) + \log(\bar{\varepsilon}), \zeta \log(n) - \log(\bar{\varepsilon})]$ of length $-2 \log(\bar{\varepsilon})$. Therefore, such a covering has cardinality $\left\lceil \frac{(b-a) \log(n)}{-2 \log(\bar{\varepsilon})} \right\rceil$. In particular, vertex weights take values in the interval $[1, n^{1/(\tau-1)}]$, which can be covered with $\left\lceil \frac{\log(n)}{-2(\tau-1) \log(\bar{\varepsilon})} \right\rceil$ sub-intervals of the form $I_{\bar{\varepsilon}}(n^\zeta)$. Similarly, all inter-distances $\|x_{v_i} - x_{v_j}\|$ take value in the interval $[n^{-1/d}, 1]$ which is covered by $\left\lceil \frac{\log(n)}{-2d \log(\bar{\varepsilon})} \right\rceil$ sub-intervals. In other words, $\bigcup_{(\alpha, \beta) \in \mathcal{F}} M_{\bar{\varepsilon}}^{(\alpha, \beta)}$ admits a finite covering over a set $\mathcal{C} \subset \mathcal{F}$, with cardinality

$$|\mathcal{C}| = \left\lceil \frac{\log(n)}{-2(\tau-1) \log(\bar{\varepsilon})} \right\rceil^k \left\lceil \frac{\log(n)}{-2d \log(\bar{\varepsilon})} \right\rceil^{\binom{k}{2}} = \Theta(\log(n)^{k(k+1)/2}).$$

Then,

$$N^{(*)}(H) \leq \sum_{(\alpha, \beta) \in \mathcal{C}} N^{(*)}(H; M_{\bar{\varepsilon}}^{(\alpha, \beta)}), \quad w.h.p.$$

and from Lemma 2.1 we conclude that

$$\mathbb{E}_n \left[N^{(*)}(H) \right] = \sum_{(\alpha, \beta) \in \mathcal{C}} \mathbb{E}_n \left[N^{(*)}(H; M_{\bar{\varepsilon}}^{(\alpha, \beta)}) \right] \leq |\mathcal{C}| \cdot \Theta(n^{f_H^*}) = o(n^{f_H^* + \delta})$$

for any $\delta > 0$. □

4. Proofs for specific subgraphs. In this section we prove results of Section 2.3.

4.1. *Proof for trees.* We need to show Theorem 2.8, that is, tree counts coincide asymptotically in GIRG and IRG models. In particular we aim to compare $\text{GIRG}^{(n)}$, where the connection probability between any pair of nodes i, j is $p_{ij} = \min \left\{ 1, \frac{w_i w_j}{n \mu \|x_i - x_j\|^d} \right\}^\gamma$, with IRG_n , where the connection probability is $\tilde{p}_{ij} = \Theta \left(1 \wedge \frac{w_i w_j}{\mu n} \right)$. We start by proving that in trees the spatial edge dependency disappears.

LEMMA 4.1. *Let $T = (V_T, E_T)$ be a tree. Let $(X_i)_{i \in V_T}$ be a sequence of i.i.d. uniformly distributed random variables on the d -dimensional torus \mathbb{T}^d . For all $i, j \in V_T$ let $f_{ij} : [0, 1/2] \rightarrow \mathbb{R}$ be a real function. Then,*

$$(4.1) \quad \mathbb{E} \left[\prod_{(i,j) \in E_T} f_{ij}(\|X_i - X_j\|) \right] = \prod_{(i,j) \in E_T} \mathbb{E} [f_{ij}(\|X_i - X_j\|)].$$

PROOF. The proof relies on the fact that, if Y_1 and Y_2 are uniformly distributed on \mathbb{T}^d and g is any real function, then

$$\mathbb{E}[g(\|Y_1 - Y_2\|)] = \mathbb{E}[g(\|Y_1\|)].$$

Indeed, conditioning on Y_2 yields

$$\begin{aligned} \mathbb{E} [\mathbb{E}[g(\|Y_1 - Y_2\|)|Y_2]] &= \int_{\mathbb{T}^d} \left(\int_{\mathbb{T}^d} g(\|y_1 - y_2\|) dy_1 \right) dy_2 \\ &= \int_{\mathbb{T}^d} \left(\int_{\mathbb{T}^d} g(\|y_1\|) dy_1 \right) dy_2 = \mathbb{E}[g(\|Y_1\|)], \end{aligned}$$

where the middle equality follows by definition of the infinity norm and symmetry of the uniform distribution.

We prove (4.1) by induction on the number of edges $|E_T|$. If $|E_T| = 1$, the equation is trivial. Then, suppose that $|E_T| = m > 1$. Let $L \in V_T$ be a leaf of the tree T , and let $P \in V_T$ be its unique parent. If we denote by T^* the subtree of T obtained by removing the edge (P, L) , then

$$\mathbb{E} \left[\prod_{(i,j) \in E_T} f_{ij}(\|X_i - X_j\|) \right] = \mathbb{E} \left[f_{PL}(\|X_P - X_L\|) \prod_{(i,j) \in E_{T^*}} f_{ij}(\|X_i - X_j\|) \right].$$

Conditioning on X_P results in

$$\begin{aligned} \mathbb{E} \left[\prod_{(i,j) \in E_T} f_{ij}(\|X_i - X_j\|) \right] &= \mathbb{E} \left[\mathbb{E} \left[f_{PL}(\|X_P - X_L\|) \prod_{(i,j) \in E_{T^*}} f_{ij}(\|X_i - X_j\|) \middle| X_P \right] \right] \\ &= \mathbb{E} [f_{PL}(\|X_L\|)] \mathbb{E} \left[\mathbb{E} \left[\prod_{(i,j) \in E_{T^*}} f_{ij}(\|X_i - X_j\|) \middle| X_P \right] \right] \\ &= \mathbb{E} [f_{PL}(\|X_P - X_L\|)] \mathbb{E} \left[\prod_{(i,j) \in E_{T^*}} f_{ij}(\|X_i - X_j\|) \right]. \end{aligned}$$

Applying the induction hypothesis on T^* concludes the proof. \square

PROOF OF THEOREM 2.8. The number of copies of H inside the $\text{GIRG}^{(n)}$ as general subgraph is

$$(4.2) \quad N^{(\text{sub})}(H) = \sum_{\mathbf{v} \in V_k} \mathbb{1}_{\{H \subseteq \text{GIRG}^{(n)}|_{\mathbf{v}}\}}.$$

The connection probabilities in $\text{GIRG}^{(n)}$ depend on the vertex weights $\mathbf{w} = \{w_i\}_{i \in [k]}$ and positions $\mathbf{x} = \{x_i\}_{i \in [k]}$. Therefore,

$$\mathbb{P}\left(H \subseteq \text{GIRG}^{(n)} \mid \mathbf{v}\right) = \mathbb{E}_{\mathbf{w}, \mathbf{x}} \left[\prod_{(i,j) \in \mathcal{E}} p(w_{v_i}, w_{v_j}, x_{v_i}, x_{v_j}) \right].$$

Observe that $p(w_{v_i}, w_{v_j}, x_{v_i}, x_{v_j})$ is a function of the distance $\|x_{v_i} - x_{v_j}\|$. Thus, for each $(i, j) \in E$ we can define $f_{ij}(\|x_{v_i} - x_{v_j}\|) := p(w_{v_i}, w_{v_j}, x_{v_i}, x_{v_j})$. Then, the marginal probability given \mathbf{w} that $H \subseteq \text{GIRG}^{(n)} \mid \mathbf{v}$ is

$$\begin{aligned} \mathbb{P}\left(H \subseteq \text{GIRG}^{(n)} \mid \mathbf{v}\right) &= \mathbb{E}_{\mathbf{x}} \left[\prod_{(i,j) \in \mathcal{E}} f_{ij}(\|x_{v_i} - x_{v_j}\|) \right] \\ &= \prod_{(i,j) \in \mathcal{E}} \mathbb{E}_{\mathbf{x}} [f_{ij}(\|x_{v_i} - x_{v_j}\|)], \end{aligned}$$

where the last equality follows from Lemma 4.1, as H is a tree. Moreover, by [5, Lemma 3.3],

$$(4.3) \quad \tilde{p}_{ij} := \mathbb{E}_{\mathbf{x}} [f_{ij}(\|x_{v_i} - x_{v_j}\|)] = \Theta\left(\frac{w_{v_i} w_{v_j}}{\mu n} \wedge 1\right)$$

and combining (4.3) with (4.2),

$$N^{(\text{sub})}(H) \stackrel{d}{\sim} \sum_{\mathbf{v} \in V_k} \prod_{(i,j) \in \mathcal{E}} \text{Be}(\tilde{p}_{ij}).$$

This concludes the proof, as the right hand side of the latter equation is a random variable counting the occurrences of H as general subgraph inside the IRG with weights \mathbf{w} and connection probabilities \tilde{p}_{ij} . \square

4.2. Proof for Hamiltonian patterns. Finally, we prove Theorem 2.9. We will study the geometric and non-geometric regimes $k < 1 + \frac{1}{3-\tau}$ and $k > \frac{2}{3-\tau}$ separately. The proof strategy is similar in the two cases. First we show that the expectations of $N^{(\text{sub})}(H)$ and $N^{(\text{ind})}(H)$ converge, when properly renormalized. The main challenge is to show that when a subgraph is Hamiltonian, this renormalized expectation is finite. Next, we will show that the two random variables concentrate around their means, which will lead to the main result.

Geometric regime $k < 1 + \frac{1}{3-\tau}$. For any $H = (\mathcal{V}, \mathcal{E})$, the expectation of $N^{(\text{sub})}(H)$ is

$$\begin{aligned} \mathbb{E}\left[N^{(\text{sub})}(H)\right] &= \frac{n!}{(n-k)!} \int \mathbb{P}(H \subseteq \text{GIRG}(\mathbf{w}, \mathbf{x}) \mid \mathbf{v}) \, d\mathbb{P}(\mathbf{w}, \mathbf{x}) \\ (4.4) \quad &= \frac{n!}{(n-k)!} \int_{[w_0, \infty)^k} \prod_{i=1}^k f_W(w_{v_i}) \\ &\quad \times \int_{(\mathbb{T}^d)^k} \prod_{(i,j) \in \mathcal{E}} \left(\frac{w_{v_i} w_{v_j}}{n\mu \|x_{v_i} - x_{v_j}\|^d} \wedge 1 \right)^\gamma \, d\mathbf{x} \, d\mathbf{w}, \end{aligned}$$

where f_W is the probability distribution function of the power law random variable W in (1.1). The spatial coordinates x_{v_1}, \dots, x_{v_k} take values in \mathbb{T}^d . However, we may assume

that v_1 has a fixed position x_{v_1} (w.l.o.g. the origin of the torus) and obtain

$$\begin{aligned}
\mathbb{E} \left[N^{(\text{sub})}(H) \right] &= \frac{n!}{(n-k)!} \int_{[w_0, \infty)^k} \prod_{i=1}^k f_W(w_{v_i}) \\
&\quad \times \left(\int_{(\mathbb{T}^d)^{(k-1)} } \prod_{(i,j) \in \mathcal{E}} \left(\frac{w_{v_i} w_{v_j}}{n\mu \|x_{v_i} - x_{v_j}\|^d} \wedge 1 \right)^\gamma dx_{v_2} \cdots dx_{v_k} \right) d\mathbf{w} \\
&= \frac{n!}{(n-k)!} \frac{1}{n^{k-1}} \int_{[w_0, \infty)^k} \prod_{i=1}^k f_W(w_{v_i}) \\
&\quad \times \left(\int_{(n^{1/d}\mathbb{T})^{d(k-1)} } \prod_{(i,j) \in \mathcal{E}} \left(\frac{w_{v_i} w_{v_j}}{\mu \|\tilde{x}_{v_i} - \tilde{x}_{v_j}\|^d} \wedge 1 \right)^\gamma d\tilde{x}_{v_2} \cdots d\tilde{x}_{v_k} \right) d\mathbf{w} \\
(4.5) \quad &=: \frac{n!}{(n-k)!} \frac{1}{n^{k-1}} J_n^{(\text{sub})}(H)
\end{aligned}$$

where we have used the substitution $\tilde{x}_{v_i} = x_{v_i}/n^{1/d}$ for $i = 2, \dots, k$, and $n^{1/d}\mathbb{T}$ is the torus rescaled by a factor $n^{1/d}$.

In particular, as n goes to infinity, $J_n^{(\text{sub})}(H)$ converges to the integral

$$\begin{aligned}
(4.6) \quad J^{(\text{sub})}(H) &:= \int_{[w_0, \infty)^k} \prod_{i=1}^k f_W(w_{v_i}) \int_{(\mathbb{R}^d)^{k-1}} \prod_{(i,j) \in \mathcal{E}} \left(\frac{w_{v_i} w_{v_j}}{\mu \|z_{v_i} - z_{v_j}\|_\infty^d} \wedge 1 \right)^\gamma \\
&\quad dz_{v_2} \cdots dz_{v_k} d\mathbf{w},
\end{aligned}$$

whenever $J^{(\text{sub})}(H)$ is defined.

In a similar way, we derive the expected number of induced subgraphs,

$$(4.7) \quad \mathbb{E} \left[N^{(\text{ind})}(H) \right] = \frac{n!}{(n-k)!} \frac{1}{n^{k-1}} J_n^{(\text{ind})}(H).$$

where $J_n^{(\text{ind})}(H)$ is an integral converging to

$$\begin{aligned}
(4.8) \quad J^{(\text{ind})}(H) &:= \int_{[w_0, \infty)^k} \prod_{i=1}^k f_W(w_{v_i}) \int_{(\mathbb{R}^d)^{k-1}} \prod_{(i,j) \in \mathcal{E}} \left(\frac{w_{v_i} w_{v_j}}{\mu \|z_{v_i} - z_{v_j}\|_\infty^d} \wedge 1 \right)^\gamma \\
&\quad \times \prod_{(i,j) \notin \mathcal{E}} 1 - \left(\frac{w_{v_i} w_{v_j}}{\mu \|z_{v_i} - z_{v_j}\|_\infty^d} \wedge 1 \right)^\gamma dz_{v_2} \cdots dz_{v_k} d\mathbf{w},
\end{aligned}$$

provided that $J^{(\text{ind})}(H)$ is defined. We need to prove that when $k < 1 + 1/(3 - \tau)$,

$$\frac{N^{(\text{sub})}(H)}{n} \xrightarrow{p} J^{(\text{sub})}(H), \quad \frac{N^{(\text{ind})}(H)}{n} \xrightarrow{p} J^{(\text{ind})}(H).$$

Then, we have all tools to prove convergence for the expectations of $N^{(\text{sub})}(H)/n$, $N^{(\text{ind})}(H)/n$.

PROPOSITION 4.2. *Under the hypothesis of Theorem 2.9(i), we have*

$$\frac{\mathbb{E} \left[N^{(\text{sub})}(H) \right]}{n} \longrightarrow J^{(\text{sub})}(H), \quad \frac{\mathbb{E} \left[N^{(\text{ind})}(H) \right]}{n} \longrightarrow J^{(\text{ind})}(H)$$

where $J^{(\text{sub})}(H)$ and $J^{(\text{ind})}(H)$ are the integrals defined in (4.6) and (4.8).

PROOF. From equation (1.7), we can write

$$\begin{aligned} \mathbb{E}[N^{(\text{sub})}(H)] &= \frac{n!}{(n-k)!} \int \mathbb{P}(H \subset \text{GIRG}(\mathbf{w}, \mathbf{x})|_v) \, d\mathbb{P}(\mathbf{w}, \mathbf{x}) \\ &= \frac{n!}{(n-k)!} \int_{[w_0, \infty)^k} \prod_{i=1}^k f_W(w_{v_i}) \int_{(\mathbb{T}^d)^k} \prod_{(i,j) \in \mathcal{E}} p(w_{v_i}, w_{v_j}, x_i, x_j) \, d\mathbf{x} \, d\mathbf{w} \end{aligned}$$

If H is a graph of size k that contains a Hamiltonian cycle, we can assume (up to a permutation of the vertices) that H contains the Hamiltonian cycle \mathcal{H}_k with edge set $\mathcal{E}_{\mathcal{H}_k} = \{(i, i+1)\}_{i \in [k]}$. Moreover, introducing a factor k , we may assume that w_{v_1} is the largest weight in \mathbf{w} . Then, we upper bound the probability for the Hamiltonian pattern H to appear on the vertex set v with the probability that the open path from node v_1 to node v_k appears:

$$\begin{aligned} \mathbb{E}[N^{(\text{sub})}(H)] &= \frac{n!}{(n-k)!} \int_{[w_0, \infty)^k} \prod_{i=1}^k f_W(w_{v_i}) \int_{(\mathbb{T}^d)^k} \prod_{(i,j) \in \mathcal{E}} p(w_{v_i}, w_{v_j}, x_i, x_j) \, d\mathbf{x} \, d\mathbf{w} \\ &\leq \frac{n!k}{(n-k)!} \int_{[w_0, \infty) \times [w_0, w_{v_1})^{k-1}} \prod_{i=1}^k f_W(w_{v_i}) \\ &\quad \times \int_{(\mathbb{T}^d)^k} \prod_{i \in [k-1]} p(w_{v_i}, w_{v_{i+1}}, x_i, x_{i+1}) \, d\mathbf{x} \, d\mathbf{w} \end{aligned}$$

Now, from Lemma 4.1 we can average out the position from the connection probability. Since

$$\mathbb{E}_{\mathbf{x}} [p_{ij}] = \Theta\left(\frac{w_{v_i} w_{v_j}}{\mu n} \wedge 1\right), \text{ we have} \quad (4.9)$$

$$\begin{aligned} \mathbb{E}[N^{(\text{sub})}(H)] &= O\left(\frac{n!k}{(n-k)!} \int_{[w_0, \infty) \times [w_0, w_{v_1})^{k-1}} \prod_{i=1}^k f_W(w_{v_i}) \prod_{i \in [k-1]} \frac{w_{v_i} w_{v_{i+1}}}{\mu n} \, d\mathbf{w}\right) \\ &= O\left(n^k \int_{[w_0, \infty) \times [w_0, w_{v_1})^{k-1}} \prod_{i=1}^k w_{v_i}^{-\tau} \prod_{i \in [k-1]} \frac{w_{v_i} w_{v_{i+1}}}{\mu n} \, d\mathbf{w}\right) \\ &= O\left(n \int_{[w_0, \infty) \times [w_0, w_{v_1})^{k-1}} w_{v_1}^{1-\tau} w_{v_2}^{2-\tau} \cdots w_{v_{k-1}}^{2-\tau} w_{v_k}^{1-\tau} \, d\mathbf{w}\right) \\ &= O\left(n \int_{[w_0, \infty)} w_{v_1}^{1-\tau+(3-\tau)(k-2)} \, dw_{v_1}\right). \end{aligned}$$

The latter integral converges whenever the exponent is smaller than -1 , which is the case since $k < 1 + \frac{1}{3-\tau}$ by hypothesis.

On the other hand, $\mathbb{E}[N^{(\text{sub})}(H)] \geq \mathbb{E}[N^{(\text{sub})}(H; M_\varepsilon^{(\alpha, \beta)})]$, for any $\varepsilon, \alpha, \beta$. Setting $\alpha \equiv 0$ and $\beta \equiv -1/d$, we have from (3)

$$(4.10) \quad \mathbb{E}[N^{(\text{sub})}(H)] = \Omega(n)$$

Now, combining (4.9)–(4.10) yields $\mathbb{E}[N^{(\text{sub})}(H)] = \Theta(n)$, that is,

$$\limsup_{n \rightarrow \infty} \frac{\mathbb{E}[N^{(\text{sub})}(H)]}{n} < \infty, \quad \liminf_{n \rightarrow \infty} \frac{\mathbb{E}[N^{(\text{sub})}(H)]}{n} > 0.$$

Then, from equation (4.5),

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\mathbb{E}[N^{(\text{sub})}(H)]}{n} &= \lim_{n \rightarrow \infty} \frac{n!}{n^k(n-k)!} J_n^{(\text{sub})}(H) \\ &= \lim_{n \rightarrow \infty} J_n^{(\text{sub})}(H) = J^{(\text{sub})}(H) \in (0, \infty). \end{aligned}$$

Finally, in order to prove convergence for the expected number of induced Hamiltonian patterns, we first observe that $\mathbb{E}[N^{(\text{ind})}(H)] \leq \mathbb{E}[N^{(\text{sub})}(H)]$. Indeed, if $H = \text{GIRG}|_{\mathbf{v}}$ for some \mathbf{v} , then also $H \subseteq \text{GIRG}|_{\mathbf{v}}$. Moreover, since $(\boldsymbol{\alpha}, \boldsymbol{\beta}) \equiv (0, -1/d)$ is in the feasible region for the optimization problem (2.4)–(2.5), the expected number of copies of H as induced subgraph is asymptotically lower bounded by $\Omega(n)$. Then, again, since $\mathbb{E}[N^{(\text{ind})}(H)] = \Theta(n)$, we conclude that the integral $J^{(\text{ind})}(H)$ is well defined, using the same argument as for $\mathbb{E}[N^{(\text{sub})}(H)]$ above, but replacing $J_n^{(\text{sub})}(H)$, $J^{(\text{sub})}(H)$ with $J_n^{(\text{ind})}(H)$, $J^{(\text{ind})}(H)$. \square

Non-geometric regime $k > \frac{2}{3-\tau}$. Using a similar approach as done in the geometric regime, we derive an explicit formula for $\mathbb{E}[N^{(\text{sub})}(H)]$ and $\mathbb{E}[N^{(\text{ind})}(H)]$. From (4.4), using the substitution $y_{v_i} = w_{v_i}/\sqrt{n}$ for all $i = 1, \dots, k$, we have

$$\begin{aligned} \mathbb{E}[N^{(\text{sub})}(H)] &= \frac{n!}{(n-k)!} \int_{[w_0, \infty)^k} \prod_{i=1}^k f_W(w_{v_i}) \\ &\quad \times \int_{(\mathbb{T}^d)^k} \prod_{(i,j) \in \mathcal{E}} \left(\frac{w_{v_i} w_{v_j}}{n\mu \|x_{v_i} - x_{v_j}\|^d} \wedge 1 \right)^\gamma \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{w} \\ &= \frac{n!}{(n-k)!} n^{k(1-\tau)/2} \int_{[n^{-1/2}w_0, \infty)^k} \prod_{i=1}^k y_{v_i}^{-\tau} \\ &\quad \times \int_{(\mathbb{T}^d)^k} \prod_{(i,j) \in \mathcal{E}} \left(\frac{y_{v_i} y_{v_j}}{\mu \|x_{v_i} - x_{v_j}\|^d} \wedge 1 \right)^\gamma \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{y} \\ (4.11) \quad &=: \frac{n!}{(n-k)!} n^{k(1-\tau)/2} I_n^{(\text{sub})}(H) \end{aligned}$$

As $n \rightarrow \infty$, the integral $I_n^{(\text{sub})}$ converges to

$$(4.12) \quad I^{(\text{sub})}(H) := \int_{(0, \infty)^k} \prod_{i=1}^k y_{v_i}^{-\tau} \int_{(\mathbb{T}^d)^k} \prod_{(i,j) \in \mathcal{E}} \left(\frac{y_{v_i} y_{v_j}}{\mu \|x_{v_i} - x_{v_j}\|^d} \wedge 1 \right)^\gamma \mathrm{d}\mathbf{x} \mathrm{d}\mathbf{y},$$

whenever $I^{(\text{sub})}$ is defined. Similarly, we derive from (1.7) that

$$\mathbb{E}[N^{(\text{ind})}(H)] = \frac{n!}{(n-k)!} n^{k(1-\tau)/2} I_n^{(\text{ind})}(H),$$

where $I_n^{(\text{ind})}(H)$ is an integral converging to

$$(4.13) \quad I^{(\text{ind})}(H) := \int_{(0,\infty)^k} \prod_{i=1}^k y_{v_i}^{-\tau} \int_{(\mathbb{T}^d)^k} \prod_{(i,j) \in \mathcal{E}} \left(\frac{y_{v_i} y_{v_j}}{\mu \|x_{v_i} - x_{v_j}\|^d} \wedge 1 \right)^\gamma \\ \times \prod_{(i,j) \notin \mathcal{E}} 1 - \left(\frac{y_{v_i} y_{v_j}}{\mu \|x_{v_i} - x_{v_j}\|^d} \wedge 1 \right)^\gamma \, d\mathbf{x} \, d\mathbf{y},$$

provided that $I^{(\text{ind})}(H)$ is defined. We now prove convergence for the expectations of $N^{(\text{sub})}(H)/n^{k(3-\tau)/2}$ and $N^{(\text{ind})}(H)/n^{k(3-\tau)/2}$.

PROPOSITION 4.3. *Under the hypothesis of Theorem 2.9(ii), we have*

$$\frac{\mathbb{E} [N^{(\text{sub})}(H)]}{n^{k(3-\tau)/2}} \longrightarrow I^{(\text{sub})}, \quad \frac{\mathbb{E} [N^{(\text{ind})}(H)]}{n^{k(3-\tau)/2}} \longrightarrow I^{(\text{ind})}$$

where $I^{(\text{sub})}(H)$ and $I^{(\text{ind})}(H)$ are the integrals defined in (4.12) and (4.13).

PROOF. We start by proving that $I^{(\text{sub})}(H)$ and $I^{(\text{ind})}(H)$ are well defined. By the assumptions of Theorem 2.9, we can assume without loss of generality that $\mathcal{H}_k \subseteq H$. Then,

$$I^{(\text{ind})}(H) \leq I^{(\text{sub})}(H) \leq I^{(\text{sub})}(\mathcal{H}_k).$$

The first inequality is obtained by upper bounding the product over $(i, j) \notin \mathcal{E}$ in (4.13) with 1; whereas in the second inequality we upper bound all the terms in the product over $(i, j) \in \mathcal{E}$, when $(i, j) \notin \mathcal{E}_{\mathcal{H}_k}$ in (4.12) with 1. Therefore, $I^{(\text{sub})}(H)$ and $I^{(\text{ind})}(H)$ are well defined as soon as $I^{(\text{sub})}(\mathcal{H}_k) < \infty$. Since k is odd, we can apply the substitution of variables $u_{v_i} = y_{v_i} y_{v_{i+1}}$ and rewrite $I^{(\text{sub})}(\mathcal{H}_k)$ as

$$I^{(\text{sub})}(\mathcal{H}_k) = \int_{(0,\infty)^k} \prod_{i=1}^k y_{v_i}^{-\tau} \int_{(\mathbb{T}^d)^k} \prod_{i=1}^k \left(\frac{y_{v_i} y_{v_{i+1}}}{\mu \|x_{v_i} - x_{v_{i+1}}\|^d} \wedge 1 \right)^\gamma \, d\mathbf{x} \, d\mathbf{y} \\ = \int_{(0,\infty)^k} \prod_{i=1}^k u_{v_i}^{-(\tau+1)/2} \int_{(\mathbb{T}^d)^k} \prod_{i=1}^k \left(\frac{u_{v_i}}{\mu \|x_{v_i} - x_{v_{i+1}}\|^d} \wedge 1 \right)^\gamma \, d\mathbf{x} \, d\mathbf{u}.$$

Next, due to the symmetry of the Hamiltonian cycle \mathcal{H}_k , we assume without loss of generality that u_{v_k} is the maximum among all u_{v_i} , for $i = 1, \dots, k$. Then

$$I^{(\text{sub})}(\mathcal{H}_k) = k \int_{(0, u_{v_k})^{k-1} \times (0, \infty)} \prod_{i=1}^k u_{v_i}^{-(\tau+1)/2} \int_{(\mathbb{T}^d)^k} \prod_{i=1}^k \left(\frac{u_{v_i}}{\mu \|x_{v_i} - x_{v_{i+1}}\|^d} \wedge 1 \right)^\gamma \, d\mathbf{x} \, d\mathbf{u} \\ \leq k \int_{(0, u_{v_k})^{k-1} \times (0, \infty)} \prod_{i=1}^k u_{v_i}^{-(\tau+1)/2} \int_{(\mathbb{T}^d)^k} \prod_{i=1}^{k-1} \left(\frac{u_{v_i}}{\mu \|x_{v_i} - x_{v_{i+1}}\|^d} \wedge 1 \right)^\gamma \, d\mathbf{x} \, d\mathbf{u},$$

where in the last inequality we have used that $\left(\frac{u_{v_k}}{\mu \|x_{v_k} - x_{v_1}\|^d} \wedge 1 \right) \leq 1$. Now, since we are left with an integral over the probability of an open path, we can use Lemma 4.1 and approximate each connection probability by

$$\int_{(\mathbb{T}^d)^k} \prod_{i=1}^{k-1} \left(\frac{u_{v_i}}{\mu \|x_{v_i} - x_{v_{i+1}}\|^d} \wedge 1 \right)^\gamma \, d\mathbf{x} = \Theta \left(\prod_{i=1}^{k-1} (u_{v_i} \wedge 1) \right).$$

Then,

$$I^{(\text{sub})}(\mathcal{H}_k) \leq k \int_{(0, u_{v_k})^{k-1} \times (0, \infty)} \prod_{i=1}^k u_{v_i}^{-(\tau+1)/2} \prod_{i=1}^{k-1} \Theta(u_{v_i} \wedge 1) \, d\mathbf{u}.$$

This integrand is separable in all variables, and we distinguish two cases.

- When $u_{v_k} \leq 1$, we have $(u_{v_i} \wedge 1) = u_{v_i}$ for all $i = 1, \dots, k-1$. Then, for all $i = 1, \dots, k-1$,

$$\int_0^{u_{v_k}} (u_{v_i} \wedge 1) u_{v_i}^{-(\tau+1)/2} \, du_{v_i} = \int_0^{u_{v_k}} u^{-(\tau+1)/2} \, du = c_1 \cdot u_{v_k}^{(3-\tau)/2}.$$

for some constant $c_1 > 0$.

- When $u_{v_k} > 1$, we split the integral over u_{v_i} for each $i = 1, \dots, k-1$, into two separate regions

$$\begin{aligned} \int_0^{u_{v_k}} (u_{v_i} \wedge 1) u_{v_i}^{-(\tau+1)/2} \, du_{v_i} &= \int_0^1 u^{-(\tau+1)/2} \, du + \int_1^{u_{v_k}} u^{-(\tau+1)/2} \, du \\ &= c_2 + (c_3 - u_{v_k}^{-(\tau-1)/2}) \leq c_4, \end{aligned}$$

as $\tau \in (2, 3)$, for some constants $c_2, c_3, c_4 > 0$.

Thus, integrating u_{v_k} over the two different regions $(0, 1)$ and $(1, \infty)$ gives:

$$\begin{aligned} I^{(\text{sub})}(\mathcal{H}_k) &\leq k \int_0^\infty \left(\int \int_0^{u_{v_k}} \prod_{i=1}^k u_{v_i}^{-(\tau+1)/2} \prod_{i=1}^{k-1} (u_{v_i} \wedge 1) \, du_{v_1} \cdots du_{v_{k-1}} \right) du_{v_k} \\ &= k \left(\int_0^1 c_1 \cdot u^{\frac{3-\tau}{2}(k-1) - \frac{\tau+1}{2}} \, du + \int_1^\infty c_4 \cdot u^{-\frac{\tau+1}{2}} \, du \right). \end{aligned}$$

The integral over $[1, \infty)$ is finite, since $-\frac{\tau+1}{2} < -1$. Moreover, the integral over $[0, 1]$ is finite as soon as $\frac{3-\tau}{2}(k-1) - \frac{\tau+1}{2} > -1$, which corresponds exactly to the hypothesis $k > \frac{2}{3-\tau}$ of Theorem 2.9(ii). Hence, since $I^{(\text{sub})}(\mathcal{H}_k) < \infty$, both integrals $I^{(\text{sub})}(H)$ and $I^{(\text{ind})}(H)$ are well defined.

Finally, from (4.11) we conclude that

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\mathbb{E}[N^{(\text{sub})}(H)]}{n^{k(3-\tau)/2}} &= \lim_{n \rightarrow \infty} \frac{\frac{n!}{(n-k)!} n^{k(1-\tau)/2} I_n^{(\text{sub})}(H)}{n^{k(3-\tau)/2}} \\ &= \lim_{n \rightarrow \infty} I_n^{(\text{sub})}(H) = I^{(\text{sub})}(H) < \infty, \end{aligned}$$

and similarly for $N^{(\text{ind})}(H)$. □

Concentration around the mean. We have now all tools to demonstrate the concentration result for Hamiltonian patterns.

PROOF OF THEOREM 2.9(I). In this case, $k < 1 + 1/(3 - \tau)$. Set $\star \in \{\text{sub}, \text{ind}\}$ and $(\alpha^*, \beta^*) \equiv (0, -1/d)$. Observe that, for any $\varepsilon > 0$,

$$\mathbb{E}[N^{(\star)}(H)] = \mathbb{E}[N^{(\star)}(H; M_\varepsilon^{(\alpha^*, \beta^*)})] + \mathbb{E}[N^{(\star)}(H; \overline{M}_\varepsilon^{(\alpha^*, \beta^*)})],$$

where $M_\varepsilon^{(\alpha^*, \beta^*)}$ is defined as in (2.1) and $\overline{M}_\varepsilon^{(\alpha^*, \beta^*)}$ is the complementary set. Moreover, there exists $J_\varepsilon(H)$ such that

$$\mathbb{E}[N^{(\star)}(H; M_\varepsilon^{(\alpha^*, \beta^*)})] = n J_\varepsilon(H) (1 + o(1)), \quad \text{and} \quad J_\varepsilon(H) \rightarrow J^{(\star)}(H) \quad \text{as } \varepsilon \rightarrow 0.$$

by Proposition 4.2 combined with the fact that the sequence of random variables $N^{(\star)}(H, M_\varepsilon^{(\alpha^*, \beta^*)})$ indexed by ε converges almost surely to $N^{(\star)}(H)$. Now, by Lemma 3.1 and the Chebychev inequality we can conclude convergence in probability in the optimal set:

$$(4.14) \quad \frac{N^{(\star)}(H; M_\varepsilon^{(\alpha^*, \beta^*)})}{n} \xrightarrow{p} J_\varepsilon(H)$$

On the other hand, the Markov inequality shows that $N^{(\star)}(H; \overline{M}_\varepsilon^{(\alpha^*, \beta^*)})$ is asymptotically upper bounded (in probability) by $(J^{(\star)}(H) - J_\varepsilon(H)) \cdot O(n)$. Then, for any $\delta > 0$,

$$(4.15) \quad \limsup_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\frac{N^{(\star)}(H; \overline{M}_\varepsilon^{(\alpha^*, \beta^*)})}{n} > \delta \right) = 0$$

Finally, combining (4.14) and (4.15) we conclude that

$$\frac{N^{(\star)}(H)}{n} \xrightarrow{p} J^{(\star)}(H).$$

□

PROOF OF THEOREM 2.9(II). The proof follows the same steps as in (i). In this case, since k is odd and $k > 2/(3 - \tau)$, Proposition 4.3 holds, and to obtain expressions similar to (4.14) and (4.15) we need to renormalize by $n^{k(3-\tau)/2}$ and replace $J^{(\star)}(H)$ with $I^{(\star)}(H)$.

□

Funding. The second author was supported by NWO VENI grant 202.001

REFERENCES

- [1] ALON, U. (2007). Network motifs: theory and experimental approaches. *Nature Reviews Genetics* **8** 450–461.
- [2] BLÄSIUS, T., FRIEDRICH, T. and KROHMER, A. (2018). Cliques in hyperbolic random graphs. *Algorithmica* **80** 2324–2344. <https://doi.org/10.1007/s00453-017-0323-3>
- [3] BLÄSIUS, T., KATZMANN, M. and STEGEHUIS, C. (2023). Maximal Cliques in Scale-Free Random Graphs. <https://doi.org/10.48550/ARXIV.2309.02990>
- [4] BOLUSANI, S., BESANÇON, M., BESTUZHEVA, K., CHMIELA, A., DIONÍSIO, J., DONKIEWICZ, T., VAN DOORNALEEN, J., EIFLER, L., GHANNAM, M., GLEIXNER, A. et al. (2024). The SCIP Optimization Suite 9.0. *arXiv preprint arXiv:2402.17702*.
- [5] BRINGMANN, K., KEUSCH, R. and LENGELER, J. (2019). Geometric inhomogeneous random graphs. *Theoretical Computer Science* **760** 35–54. <https://doi.org/10.1016/j.tcs.2018.08.014>
- [6] CONNOR, E. F. and SIMBERLOFF, D. (1979). The assembly of species communities: chance or competition? *Ecology* **60** 1132–1140.
- [7] DEPPEZ, P. and WÜTHRICH, M. V. (2019). Scale-free percolation in continuum space. *Communications in Mathematics and Statistics* **7** 269–308.
- [8] GRACAR, P., HEYDENREICH, M., MÖNCH, C. and MÖRTERS, P. (2022). Recurrence versus transience for weight-dependent random connection models. *Electronic Journal of Probability* **27** 1–31.
- [9] HOLLAND, P. W. and LEINHARDT, S. (1976). Local Structure in Social Networks. *Sociological Methodology* **7** 1–45.
- [10] HOLLAND, P. W. and LEINHARDT, S. (1977). A Method for Detecting Structure in Sociometric Data. In *Social Networks* (S. Leinhardt, ed.) 411–432. Academic Press. <https://doi.org/10.1016/B978-0-12-442450-0.50028-6>
- [11] ITZKOVITZ, S., MILO, R., KASHTAN, N., ZIV, G. and ALON, U. (2003). Subgraphs in random networks. *Physical review E* **68** 026127.
- [12] JACOB, E. and MÖRTERS, P. (2015). SPATIAL PREFERENTIAL ATTACHMENT NETWORKS: POWER LAWS AND CLUSTERING COEFFICIENTS. *The Annals of Applied Probability* **25** 632–662.
- [13] JIANG, S., FIORE, G. A., YANG, Y., FERREIRA JR, J., FRAZZOLI, E. and GONZÁLEZ, M. C. (2013). A review of urban computing for mobile phone traces: current methods, challenges and opportunities. In *Proceedings of the 2nd ACM SIGKDD international workshop on Urban Computing* 1–9.

- [14] KASHTAN, N., ITZKOVITZ, S., MILO, R. and ALON, U. (2004). Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs. *Bioinformatics* **20** 1746–1758.
- [15] KRIOUKOV, D., PAPADOPOULOS, F., KITSACK, M., VAHDAT, A. and BOGUNÁ, M. (2010). Hyperbolic geometry of complex networks. *Physical Review E* **82** 036106.
- [16] MA'AYAN, A., JENKINS, S. L., NEVES, S., HASSELDINE, A., GRACE, E., DUBIN-THALER, B., EUNG-DAMRONG, N. J., WENG, G., RAM, P. T., RICE, J. J. et al. (2005). Formation of regulatory patterns during signal propagation in a mammalian cellular network. *Science* **309** 1078–1083.
- [17] MAHER, S., MILTENBERGER, M., PEDROSO, J. P., REHFELDT, D., SCHWARZ, R. and SER-RANO, F. (2016). PySCIPOpt: Mathematical Programming in Python with the SCIP Optimization Suite. In *Mathematical Software – ICMS 2016* 301–307. Springer International Publishing. https://doi.org/10.1007/978-3-319-42432-3_37
- [18] MICHIELAN, R. and STEGEHUIS, C. (2022). Cliques in geometric inhomogeneous random graphs. *Journal of Complex Networks* **10** cnac002.
- [19] MILO, R., SHEN-ORR, S., ITZKOVITZ, S., KASHTAN, N., CHKLOVSKII, D. and ALON, U. (2002). Network motifs: simple building blocks of complex networks. *Science* **298** 824–827.
- [20] MOLLOY, M. and REED, B. (1995). A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms* **6** 161–180. <https://doi.org/10.1002/rsa.3240060204>
- [21] NOWICKI, K. and WIERMAN, J. C. (1988). Subgraph counts in random graphs using incomplete U-statistics methods. *Discrete Mathematics* **72** 299–310.
- [22] OH, E. and OH, S. (2023). Algorithms for Computing Maximum Cliques in Hyperbolic Random Graphs. In *31st Annual European Symposium on Algorithms (ESA 2023)* (I. L. GÖRTZ, M. FARACH-COLTON, S. J. PUGLISI and G. HERMAN, eds.). *Leibniz International Proceedings in Informatics (LIPIcs)* **274** 85:1–85:15. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany. <https://doi.org/10.4230/LIPIcs.ESA.2023.85>
- [23] PTACEK, J., DEVGAN, G., MICHAUD, G., ZHU, H., ZHU, X., FASOLO, J., GUO, H., JONA, G., BREITKREUTZ, A., SOPKO, R. et al. (2005). Global analysis of protein phosphorylation in yeast. *Nature* **438** 679–684.
- [24] RIBEIRO, P., PAREDES, P., SILVA, M. E. P., APARICIO, D. and SILVA, F. (2021). A Survey on Subgraph Counting: Concepts, Algorithms, and Applications to Network Motifs and Graphlets. *ACM Comput. Surv.* **54**. <https://doi.org/10.1145/3433652>
- [25] RUCIŃSKI, A. (1985). Subgraphs of Random Graphs: A General Approach. In *Random Graphs '83*, (M. Karoński and A. Ruciński, eds.). *North-Holland Mathematics Studies* **118** 221–229. North-Holland. [https://doi.org/10.1016/S0304-0208\(08\)73623-5](https://doi.org/10.1016/S0304-0208(08)73623-5)
- [26] SARACCO, F., DI CLEMENTE, R., GABRIELLI, A. and SQUARTINI, T. (2015). Randomizing bipartite networks: the case of the World Trade Web. *Scientific reports* **5** 10595.
- [27] SARACCO, F., DI CLEMENTE, R., GABRIELLI, A. and SQUARTINI, T. (2016). Detecting early signs of the 2007–2008 crisis in the world trade. *Scientific reports* **6** 30286.
- [28] SCHREIBER, F. and SCHWÖBBERMEYER, H. (2005). Frequency concepts and pattern detection for the analysis of motifs in networks. In *Transactions on computational systems biology III* 89–104. Springer.
- [29] SHEN-ORR, S. S., MILO, R., MANGAN, S. and ALON, U. (2002). Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nature genetics* **31** 64–68.
- [30] STEGEHUIS, C., HOFSTAD, R. v. D. and VAN LEEUWAARDEN, J. S. (2019). Variational principle for scale-free network motifs. *Scientific reports* **9** 6762.
- [31] SÖNMEZ, E. and STEGEHUIS, C. (2024). On the Distances Within Cliques in a Soft Random Geometric Graph. *Journal of Statistical Physics* **191**. <https://doi.org/10.1007/s10955-024-03254-3>
- [32] VAN DER HOFSTAD, R. (2016). *Random graphs and complex networks* **43**. Cambridge university press.
- [33] VAN DER HOFSTAD, R., VAN DER HOORN, P. and MAITRA, N. (2023). Local limits of spatial inhomogeneous random graphs. *Advances in Applied Probability* 1–48.
- [34] VAN DER HOFSTAD, R., VAN LEEUWAARDEN, J. S. H. and STEGEHUIS, C. (2021). Optimal subgraph structures in scale-free configuration models. *The Annals of Applied Probability* **31** 501 – 537. <https://doi.org/10.1214/20-AAP1580>