# On the renormalization of random network models

Catanzaro, A.

CHAPTER 3

# Clustering without Geometry in an Independent Edges Network Model

This chapter is based on:
Alessio Catanzaro, Remco van der Hofstad, and Diego Garlaschelli. *"Clustering without geometry in sparse networks with independent edges"*. arXiv preprint arXiv:2603.13159 (2026).


**Abstract**

The coexistence of sparsity and clustering (non-vanishing average fraction of triangles per node) is one of the few structural features that, irrespective of finer details, are ubiquitously observed across large real-world networks – motivating the search for generic models producing sparse clustered graphs. Earlier results suggested that sparse random graphs with independent edges fail to reproduce clustering, unless edge probabilities are assumed to depend on underlying metric distances that, thanks to the triangular inequality, naturally favour triadic closure. This observation has opened a debate on whether clustering implies (latent) geometry in real-world networks. Alternatively, recent models of higher-order networks can replicate clustering by abandoning edge independence. Here we prove mathematically, and confirm numerically, that a sparse random graph with independent edges, recently identified in the context of network renormalization as an invariant model under node aggregation, produces finite clustering without any geometric or higher-order constraint. The underlying mechanism is the presence of an infinite-mean node fitness which also implies a power-law degree distribution and, as a novel phenomenon that we characterize rigorously, the breakdown of self-averaging of various network properties. Therefore, as an alternative to geometry or higher-order dependencies, node aggregation invariance emerges as a basic route to realistic network properties.

## §3.1 Introduction: clustering versus geometry.

While real-world networks differ in various structural details, the vast majority of them are characterized by virtually ubiquitous features, the most popular of which are the broad (i.e. infinite-variance) distribution of node degrees (numbers of links per node), the *(ultra)small-world* property (i.e. average shortest path length growing (at most) logarithmically with the number of nodes), and the coexistence of *sparsity* (vanishing link density) and *local clustering* (finite average density of triangles around nodes) in the limit of infinitely large networks [4, 6]. While several models based on different generic mechanisms can successfully replicate broad degree distributions and short path lengths [4], the coexistence of sparsity and local clustering is much more difficult to replicate – and hence more informative, as it requires specific mechanisms that can discriminate models at a finer level.

The prototypical example of this difficulty is illustrated by *random graph models with independent edges*, where different pairs of nodes are independently connected. In the Erdős-Rényi (ER) model [10], where edges are independent and identically distributed (i.i.d.) with the same connection probability $p$, it is easy to show that the link density (expected fraction of realized links) and the local clustering coefficient (defined as the fraction of triangles around a node, averaged over all nodes) have the same expected value $p$ for networks with a large number $n$ of nodes. This automatically implies that, as $n$ diverges, one cannot simultaneously produce finite local clustering and vanishing link density. To overcome this limitation of the ER model, two main routes have been explored.

One route introduces dependencies between edges. This can happen via explicit triadic interactions [9, 87], which however still lead to comparable levels of density and clustering, or via a monopartite projection[1] of a bipartite network [40], which produces a network of overlapping cliques (the edges of which are all mutually dependent) where local clustering can remain very large even for vanishing density. More recent attempts introduce higher-order networks such as simplicial complexes and hypergraphs [88, 89], where edges belonging to the same simplex or hyperedge are mutually dependent. While these more complicated settings can realize the coexistence of sparsity and clustering at the expense of introducing higher-order dependencies, whether that coexistence can be achieved for more parsimonious models with independent edges remains an open question.

The second route explores that question explicitly by preserving edge independence while introducing heterogeneity. More precisely, one can turn the *unconditional* link independence of the ER model into a *conditional* one, by replacing, for each specific pair of nodes $i$ and $j$ ($i, j = 1, \ldots, n$), the constant connection probability $p$ with a probability $p_{ij}$ that is a function $f(w_i, w_j)$ of the realized values $w_i$, $w_j$ of a certain node-specific variable $w$ (called *fitness* or weight, hidden variable, latent variable, Lagrange multiplier, etc.) that is preliminarly assigned to vertices via an i.i.d.

[1] In monopartite projections of bipartite networks, the original nodes are first initially connected (possibly independently of each other) to 'auxiliary' nodes of a different type (representing for instance the possibile affiliations or memberships of the original nodes), and the auxiliary nodes are then eliminated while connecting the original nodes to each other, if they share connections to the same auxiliary nodes.

draw from a certain probability density function (pdf) $\rho(w)$ [11, 13, 64, 90]. Typical choices of $f(w_i, w_j)$ are monotonically increasing in both arguments (and necessarily symmetric for undirected graphs) – implementing the idea that, the larger the fitness, the larger the probability that a node connects to other nodes[2].

In the studies carried out so far, it turned out that, while several combinations of $f$ and $\rho$ can generate sparse small-world networks with a broad degree distribution [11, 13, 64, 90], they systematically fail to generate finite clustering, casting doubts on whether models with independent edges can replicate local clustering. The only exception that has become popular is that of *random geometric graphs*, where the fitness $w$ acquires the role of a *node coordinate* (which, more in general, can be a $D$-dimensional vector) and the connection function $f$ depends on $w_i$ and $w_j$ through some *metric distance* $d_{ij} = d(w_i, w_j)$ (in which case, $f$ generally loses the property of being monotonic in $w_i$[3]). A remarkable example of this success is the *hyperbolic model* [14, 15, 16, 17, 18, 19, 21] which assumes that nodes are sprinkled uniformly at random in a $D$-dimensional hyperbolic space and links are formed with a probability that is a certain decreasing function of the geodesic distance between nodes. In this setting, the triangular inequality ($d_{ij} \leq d_{ik} + d_{jk}\ \forall i, j, k$) guarantees that pairs of nodes $i, j$ that successfully connect to a third node $k$ are very likely to connect to each other (*triadic closure*), provably leading to finite clustering [91, 92]. Additionally, hyperbolicity ensures that other desirable properties such as power-law degree distributions emerge [16]. These investigations have convincingly concluded that, in random graphs with conditionally independent edges and distance-dependent connection probabilities, *geometry implies clustering*.

On the other hand, the potential reverse (non-obvious) implication that *clustering implies geometry*, i.e. that the presence of finite clustering could suffice to indicate that the network is (with high probability) the realization of a random geometric graph, is highly debated [14, 15, 16, 17, 18, 19, 21]. Whether clustering implies geometry is crucial for network theory (because, if true, it would demand for explanations and interpretations of the hidden geometry of real-world networks [21, 91]), for data science (for instance because techniques such as graph embedding and topological data analysis could then be optimized for real-world network geometries) and for the control of processes on networks (because geometry could aid network navigability [16, 93, 94] and the understanding of how clustering impacts epidemic and other processes [9]).

In this chapter, we prove that finite clustering can emerge independently of geometry or higher-order constraints in random graphs with independent edges conditional on infinite-mean fitness [1, 2]. While the finite-mean case has been systematically found to produce vanishing clustering, we show that in the infinite-mean case clustering is finite and strong. Notably, we find that the infinite-mean fitness also implies the breakdown of self-averaging in certain properties, including clustering itself.

[2]Some popular examples are $f(w_i, w_j) = z w_i w_j$, $f(w_i, w_j) = \Theta(w_i + w_j - z)$, $f(w_i, w_j) = z w_i w_j / (1 + z w_i w_j)$, and $f(w_i, w_j) = 1 - e^{-z w_i w_j}$, where $z > 0$ is a global parameter controlling for the overall link density and $w_i \geq 0\ \forall i$.

[3]For instance, if $w_i$ and $w_j$ are scalar (unidimensional) coordinates and their distance is defined as $d_{ij} = |w_i - w_j|$, then $w_i$ and $w_j$ can simultaneously increase while keeping $d_{ij}$ (and hence $p_{ij}$) unchanged.

## §3.2 The Multi-Scale Model.

We consider the inhomogeneous random graph model introduced in [1] and studied in [2, 3, 95]. Each node $i$ is assigned a positive weight (fitness) $w_i$ $(i = 1, n)$, sampled i.i.d. from a certain pdf $\rho(w)$. Given $w_i, w_j$, nodes $i, j$ connect with probability

$$p_{ij} = 1 - e^{-\delta_n w_i w_j}, \tag{3.1}$$

where the scaling $\delta_n$ can be chosen in such a way that the link density vanishes as $n \to \infty$. The connection probability (3.1) is also known in the literature as the Norros–Reittu model [96] and has been used for various purposes [97, 98], by considering a finite-mean $\rho(w)$. In this work we are however interested in the specific 'annealed' regime considered within the framework of network renormalization [54], where the model emerges as a fixed point of a renormalization flow in the space of models, induced by coarse-graining the network by aggregating nodes into super-nodes and recalculating the induced connection probability between super-nodes [1]. The specific connection rule in Eq. (3.1) makes the model invariant, provided the weight of each block is defined as (a random variable equal to) the sum of the weights of the constituent nodes [1]. If one further demands the invariance of the weight distribution, i.e. that the coarse-grained weights are distributed according to the same pdf as the constituent weights, then $\rho(w)$ should be a positively-supported, one-sided stable law $\tilde{\rho}_\alpha(w)$ with stability parameter $0 < \alpha < 1$ [99, 22] and tails decaying as $w^{-1-\alpha}$ for large $w$, hence with diverging mean and higher moments [1, 3]. In this way, one gets a Multi-Scale Model (MSM) interpretable as describing connections between arbitrarily nested (blocks of) nodes, thereby describing the graph at multiple scales of resolution (coarse-graining) simultaneously [1].

Since a stable law $\tilde{\rho}_\alpha(w)$ with $0 < \alpha < 1$ is known in closed form only for $\alpha = 1/2$ (Lévy distribution), we first follow [3] and replace it with a pure Pareto distribution

$$\rho_\alpha(w) = \alpha w^{-1-\alpha}, \quad w \geq 1,\ 0 < \alpha < 1 \tag{3.2}$$

(which has exactly the same tail behaviour as an $\alpha$-stable distribution and serves as an analytically tractable counterpart [3, 99]) and then discuss the extension to actual $\alpha$-stable laws $\tilde{\rho}_\alpha(w)$ via both analytical and numerical arguments. Given (3.2), one can show that choosing $\delta_n = n^{-1/\alpha}$ implies that the link density vanishes as $\log n/n$, i.e. the average degree grows only as $\log n$, and the degree distribution is power-law [3].

We stress that the MSM can be enriched with additional dyadic features [1, 2] which allow the connection probabilities to depend, for instance, on geometric distances between coordinates assigned to nodes. Nonetheless, to investigate whether clustering can be produced *without* geometry, here we obviously consider the simplest, distance-independent version in Eq. (3.1).

## §3.3 The Local Clustering Coefficient.

Abstractly, the clustering coefficient quantifies the tendency of a graph to form closed triangles. Globally, it can be defined as the ratio between the total number of closed

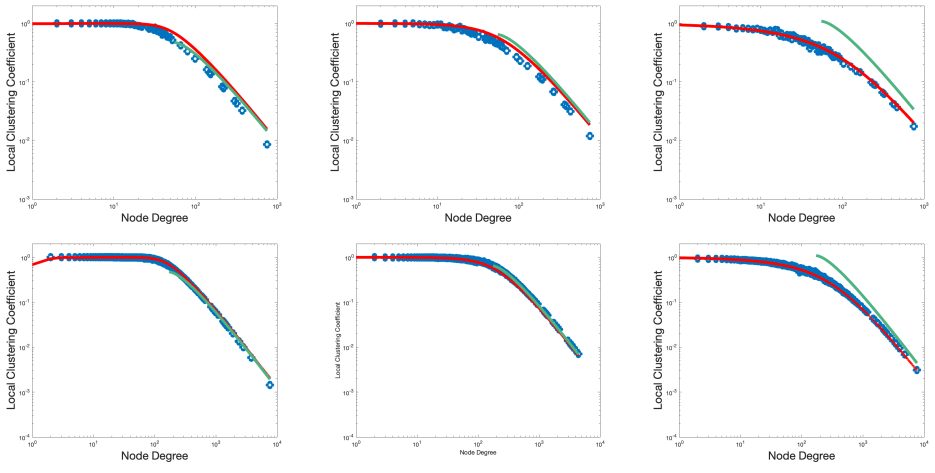


*Figure 3.1:* ***C****lustering functions for different values of* $n$ *and* $\alpha$ *for the MSM model* (3.1) *with Pareto weights* (3.2)*. Blue circles: empirical clustering function* (3.4) *versus the degree* $k$ *computed on actual realized graphs (obtained by sampling the weights once, and sampling the graph once conditionally on the realized weights). Red curves: our analytical expression* (3.6) *for the annealed clustering function evaluated on the actual degree* $k = a\sqrt{n}$*. Green curves: our asymptotic calculation* (3.44) *valid for diverging reduced degrees. We evaluate and plot all functions only for degrees larger than* 1*, to avoid ambiguities in the definition of the clustering coefficient for* $k < 2$*. Top:* $n = 10^3$*, bottom:* $n = 10^4$*. From left to right:* $\alpha = 0.3, 0.5, 0.7$*.*



triangles and the number of potential (either open or closed) triangles (also called cherries, or wedges). In our regime with $0 < \alpha < 1$, this ratio vanishes for the MSM model [3], in line with what observed in sparse real-world networks. Our interest here is the (more distinctive) *local* clustering coefficient $C_v$, which quantifies the tendency of the neighbours of the specific node $v$ to form closed triangles around $v$. For a node $v$ with degree $D_v$, $C_v$ is defined as the ratio between the number $\Delta_v$ of triangles containing $v$ and the number of wedges centered at $v$. In terms of the adjacency matrix $\mathbf{A} = (A_{ij})_{i,j=1}^n$, this reads

$$C_v = \frac{\Delta_v}{\binom{D_v}{2}} = \frac{\sum_{i\neq v}\sum_{j\neq v,i} A_{vi}A_{ij}A_{jv}}{D_v(D_v-1)}, \quad D_v > 1. \tag{3.3}$$

The behaviour of $C_v$ versus $D_v$, as well as the average value $C$ of $C_v$ over nodes, can effectively characterize real-world networks and discriminate among different models – especially in determining whether a finite (expected) value of $C$ can emerge in large sparse graphs, as observed in real-world networks. Note that $C_v$ in Eq. (3.3) is defined only for nodes with $D_v > 1$, as only nodes with at least two neighbours can host a triangle. For nodes with $D_v = 0, 1$, one can either leave $C_v$ undefined (and exclude it from the calculation of $C$) or set conventionally $C_v \equiv 0$ (and include it in the calculation of $C$). We will consider both cases and discuss their implications.

## §3.3.1 Clustering Function.

For a given graph, we consider the average clustering coefficient as a function of the degree $k$ of each node, i.e., the *empirical clustering function*

$$C(k) = \frac{1}{N_k} \sum_{v:\ D_v=k} \frac{\Delta_v}{\binom{k}{2}}, \tag{3.4}$$

where $N_k = \sum_{v:\ D_v=k} \delta_{D_v,k}$ is the number of nodes with degree $k$. In terms of expectations ($\mathbb{E}$) over the model, we introduce the *annealed clustering function* as

$$\bar{C}(k) = \frac{1}{\mathbb{E}[N_k]}\, \mathbb{E}\left[ \sum_{v:\ D_v=k} \frac{\Delta_v}{\binom{k}{2}} \right]. \tag{3.5}$$

One of our main results is a rigorous calculation of $\bar{C}(k)$ in the asymptotic (large $n$) limit, valid for both small and large $k$, which shows excellent agreement with the empirical clustering function $C(k)$ [4]. In this setting, we redefine $\bar{C}(\cdot)$ as a function of the *reduced degree* $a = k/\sqrt{n}$, as this rescaling allows us to prove that, defining $g(x) \equiv 1 - e^{-x}$, the limiting form of $\bar{C}(a)$ as $n \to \infty$ is

$$\bar{C}(a) = \frac{\alpha^2}{a^2} \int_0^\infty \mathrm{d}x \int_0^\infty \mathrm{d}y\, \frac{g(a^{1/\alpha}\tau_\alpha x)}{x^{1+\alpha}} \frac{g(a^{1/\alpha}\tau_\alpha y)}{y^{1+\alpha}} g(xy), \tag{3.6}$$

where $\tau_\alpha \equiv \Gamma(1-\alpha)^{-1/\alpha}$, and $\Gamma(\cdot)$ is the Euler Gamma function. The proof is in section 3.3.2. In Figure 3.1, we compare the numerical evaluation of the integral (3.6) with the direct computation of the empirical clustering function (3.4) on the actual realizations of the random graph. The agreement is remarkably good, already for moderate values of $n$.

## §3.3.2 Limiting expression for local clustering function

In this section we prove that the clustering function of the MSM indeed converges in the large $n$ limit to the expression (3.6).

**Theorem 3.3.1.**
*Let the vertex space of the network be $[n] = \{1, \dots, n\}$. Let $(W_v)_{v\in[n]}$ be i.i.d. Pareto random variables, so that their density equals*

$$\rho_\alpha(w) = \frac{\alpha}{w^{\alpha+1}}, \qquad w \geq 1, \tag{3.7}$$

*with $\alpha \in (0,1)$, so that $\mathbb{E}[W] = \infty$. Conditionally on $(W_v)_{v\in[n]}$, edges are present independently, with the edge between $u, v \in [n]$ being present with probability*

$$p_{uv} = 1 - e^{-W_u W_v / n^{1/\alpha}}. \tag{3.8}$$

[4] We also believe that, through a second-moment analysis, it could be proven that in the large $n$ limit, with high probability, the two clustering functions become the same. We will however not prove this claim in the present work.

*Throughout, we let $(I_{u,v})_{1\le u<v\le n}$ denote the edge indicator variables (or the elements of the adjacency matrix), so that*

$$\mathbb{P}\big(I_{u,v}=1 \mid (W_v)_{v\in[n]}\big) = 1-e^{-W_uW_v/n^{1/\alpha}}, \tag{3.9}$$

*and $(I_{u,v})_{1\le u<v\le n}$ are conditionally independent given $(W_v)_{v\in[n]}$. Let the Annealed Clustering Function be defined as*

$$\bar{C}(k) = \frac{1}{\mathbb{E}\left[N_k\right]}\,\mathbb{E}\left[\sum_{v/D_v=k}\frac{\Delta_v}{\binom{k}{2}}\right] \tag{3.10}$$

*Then, as $n\to\infty$, for $k=a\sqrt{n}$,*

$$\bar{C}_n(a\sqrt{n}) \to \bar{C}(a) = \frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}][1-e^{-xy}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y, \tag{3.11}$$

*where $\tau_\alpha=\Gamma(1-\alpha)^{-1/\alpha}$.*

*Proof.* We compute

$$\mathbb{E}[N_k] = n\,\mathbb{E}\left[\mathbb{P}(D_1=k \mid W_1)\right], \tag{3.12}$$

where we note that, conditionally on $W_1=w$, $D_1\sim \mathrm{Bin}\,(n-1,q_w)$, where

$$q_w = \mathbb{E}\left[1-e^{-wW/n^{1/\alpha}}\right]. \tag{3.13}$$

We also compute the numerator

$$\mathbb{E}\Big[\sum_{v\in[n]}\Delta_v\mathbb{1}_{D_v=k}\Big] = n\,\mathbb{E}\Big[\mathbb{E}\left[\Delta_1\mathbb{1}_{D_1=k}\mid W_1\right]\Big] \tag{3.14}$$

$$= n\binom{n-1}{2}\,\mathbb{E}\Big[\mathbb{P}(D_1'=k-2\mid W_1)\mathbb{P}(I_{1,2}I_{1,3}I_{2,3}=1\mid W_1)\Big], \tag{3.15}$$

where $D_1'\sim\mathrm{Bin}(n-3,q_w)$, and we have used conditional independence. Here, we can identify

$$D_1' = \sum_{u\in[n]\setminus[3]} I_{1,u}. \tag{3.16}$$

We see that all these expectations involve binomial random variables, which we continue to investigate, starting with the conditional success probability $q_w$.

**Analysis of success probabilities** We compute

$$q_w = \mathbb{E}\left[1-e^{-wW/n^{1/\alpha}}\right] \tag{3.17}$$

$$= \int_1^\infty [1-e^{-ws/n^{1/\alpha}}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s \tag{3.18}$$

$$= \big(\frac{w}{n^{1/\alpha}}\big)^\alpha\int_{wn^{-1/\alpha}}^\infty [1-e^{-s}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s$$

$$= \frac{w^\alpha}{n}\int_0^\infty [1-e^{-s}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s - \frac{w^\alpha}{n}\int_0^{wn^{-1/\alpha}}[1-e^{-s}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s.$$

We next use that

$$\int_0^\infty [1-e^{-s}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s = \int_0^\infty \int_0^s e^{-y}\mathrm{d}y\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s \tag{3.19}$$

$$= \int_0^\infty e^{-y}\int_y^\infty \frac{\alpha}{s^{\alpha+1}}\mathrm{d}s\mathrm{d}y \tag{3.20}$$

$$= \int_0^\infty e^{-y}y^{-\alpha}\mathrm{d}y = \Gamma(1-\alpha), \tag{3.21}$$

and

$$\int_0^{wn^{-1/\alpha}} [1-e^{-s}]\frac{\alpha}{s^{\alpha+1}}\mathrm{d}s \le \int_0^{wn^{-1/\alpha}} \frac{\alpha}{s^{\alpha}}\mathrm{d}s \tag{3.22}$$

$$= \frac{\alpha}{\alpha-1}\left(wn^{-1/\alpha}\right)^{1-\alpha} \tag{3.23}$$

$$= \frac{\alpha}{\alpha-1}w^{1-\alpha}n^{1-1/\alpha}. \tag{3.24}$$

This gives excellent control over $q_w$ for all $w = o(n^{1/\alpha})$.

**Binomial local limit approximations** Let $S_m \sim \mathrm{Bin}(m,p)$. Then, a local central limit approximation shows that

$$\mathbb{P}(S_m = k) = \frac{1}{\sqrt{2\pi mp(1-p)}}\mathrm{e}^{-(k-mp)^2/[2mp(1-p)]}(1+o(1)), \tag{3.25}$$

when $|k-mp| = o((mp)^{2/3})$, and uniformly in $k$ and $p$. Such approximations have a long history, we refer to the books by Bollobás [100] and Janson, Łuczak and Ruciński [101]. In particular, the upper bound in (3.25) follows from [100, Theorem 1.2], the lower bound from [100, Theorem 1.5], and an analysis of the error terms present there, showing that they are $o(1)$ when $|n-mp| = o((mp)^{2/3})$.

In our case, we have that $m = n-1$ or $m = n-2$, $p = q_w$ with $w = W_1$, so that

$$\mathbb{P}(D_1 = k \mid W_1 = w) = \frac{1}{\sqrt{2\pi mq_w(1-q_w)}}\mathrm{e}^{-(k-mq_w)^2/[2mq_w(1-q_w)]}(1+o(1)), \tag{3.26}$$

when $|k-mq_w| = o((mq_w)^{2/3})$, and uniformly in $k$ and $w$. Since, for us, $k$ is fixed, while $w$ is being integrated out over, the above means that the main contribution for $W_1 = w$ comes from $w$ for which $mq_w \approx k$, with values of $w$ for which

$$|k-mq_w| \ge C\sqrt{k}\log k \tag{3.27}$$

receiving probability at most

$$\mathrm{e}^{-C\log k} = k^{-C}, \tag{3.28}$$

which is negligible. Thus, we can safely think of $mq_w = k$, which means

$$mq_w = w^\alpha\Gamma(1-\alpha)(1+o(1)) = k, \tag{3.29}$$

so that

$$w = k^{1/\alpha}\Gamma(1-\alpha)^{-1/\alpha}(1+o(1)). \tag{3.30}$$

This implies that, with $m = n-1$,

$$\begin{aligned}\mathbb{E}[N_k] &= n\,\mathbb{E}\left[\mathbb{P}(D_1 = k \mid W_1)\right] \\ &= (1+o(1))n\int_1^\infty \frac{1}{\sqrt{2\pi m q_w(1-q_w)}} \mathrm{e}^{-(k-mq_w)^2/[2mq_w(1-q_w)]}\frac{\alpha}{w^{\alpha+1}}\mathrm{d}w \\ &= (1+o(1))\int_1^\infty \frac{1}{\sqrt{2\pi n q_w}} \mathrm{e}^{-(k-nq_w)^2/[2nq_w]}\frac{\alpha}{w^{\alpha+1}}\mathrm{d}w,\end{aligned} \tag{3.31}$$

as long as $nq_w \gg 1$, and thus, since $k \approx nq_w$, also $k \gg 1$. Since the above integral is dominated by $nq_w$ that are close to $k$, we may replace $nq_w$ by $k$ everywhere, except in the $(k-nq_w)^2$ term. This gives

$$\mathbb{E}[N_k] = (1+o(1))\frac{n}{\sqrt{2\pi k}}\int_1^\infty \mathrm{e}^{-(k-nq_w)^2/[2k]}\frac{\alpha}{w^{\alpha+1}}\mathrm{d}w. \tag{3.32}$$

We next substitute $nq_w \approx w^\alpha\Gamma(1-\alpha)$ to arrive at

$$\begin{aligned}\mathbb{E}[N_k] &= (1+o(1))\frac{n}{\sqrt{2\pi k}}\int_1^\infty \mathrm{e}^{-(k-w^\alpha\Gamma(1-\alpha))^2/[2k]}\frac{\alpha}{w^{\alpha+1}}\mathrm{d}w \\ &= (1+o(1))\frac{n\alpha}{w_k^{\alpha+1}\sqrt{2\pi k}}\int_1^\infty \mathrm{e}^{-(k-w^\alpha\Gamma(1-\alpha))^2/[2k]}\mathrm{d}w,\end{aligned} \tag{3.33}$$

where $w_k$ solves $w^\alpha\Gamma(1-\alpha) = k$, and thus

$$w_k = k^{1/\alpha}\Gamma(1-\alpha)^{-1/\alpha}. \tag{3.34}$$

We then let $z = w^\alpha\Gamma(1-\alpha)$, so that $\mathrm{d}z = \alpha w^{\alpha-1}\Gamma(1-\alpha)\mathrm{d}w$, so that

$$\mathrm{d}w = \Gamma(1-\alpha)^{-1}(z/\Gamma(1-\alpha))^{(1-\alpha)/\alpha}\mathrm{d}z = z^{(1-\alpha)/\alpha}\Gamma(1-\alpha)^{-1/\alpha}\mathrm{d}z, \tag{3.35}$$

so that

$$\begin{aligned}\mathbb{E}[N_k] &= (1+o(1))\frac{n\alpha}{w_k^{\alpha+1}\sqrt{2\pi k}}\int_1^\infty \mathrm{e}^{-(k-z)^2/[2k]}z^{(1-\alpha)/\alpha}\Gamma(1-\alpha)^{-1/\alpha}\mathrm{d}z \\ &= (1+o(1))\frac{n\alpha}{w_k^{\alpha+1}\sqrt{2\pi k}}\Gamma(1-\alpha)^{-1/\alpha}k^{(1-\alpha)/\alpha}\int_1^\infty \mathrm{e}^{-(k-z)^2/[2k]}\mathrm{d}z \\ &= (1+o(1))\frac{n\alpha}{w_k^{\alpha+1}\sqrt{2\pi k}}k^{(1-\alpha)/\alpha}\Gamma(1-\alpha)^{-1/\alpha}\sqrt{2\pi k} \\ &= n(1+o(1))\alpha k^{-(\alpha+1)/\alpha}\Gamma(1-\alpha)^{(\alpha+1)/\alpha}k^{(1-\alpha)/\alpha}\Gamma(1-\alpha)^{-1/\alpha} \\ &= n(1+o(1))\frac{\alpha\Gamma(1-\alpha)}{k^2}.\end{aligned} \tag{3.36}$$

This confirms that $\mathbb{E}[N_k] = \Theta(1)$ when $k = \Theta(\sqrt{n})$, so that one cannot expect that $N_k/\mathbb{E}[N_k] \xrightarrow{p} 1$ for such values of $k$.

Since (3.25) depends on $k$, for $k$ large, in a relatively weak way, this means that, for values of $w$ close to $w_k$,

$$\mathbb{P}(D_1' = k-2 \mid W_1 = w) \approx \mathbb{P}(D_1 = k \mid W_1 = w). \tag{3.37}$$

Thus, since such factors appear both in $\mathbb{E}[N_k]$ and $\mathbb{E}\left[\sum_{v\in[n]} \Delta_v \mathbb{1}_{D_v=k}\right]$, they cancel out asymptotically, so that

$$\frac{\mathbb{E}\left[\sum_{v\in[n]} \Delta_v \mathbb{1}_{D_v=k}\right]}{\mathbb{E}[N_k]} \approx n\binom{n-1}{2}\mathbb{P}(I_{1,2}I_{1,3}I_{2,3}=1 \mid W_1 = w_k). \tag{3.38}$$

Recalling (3.10), this then shows that

$$\bar{C}_n(k) = (1+o(1))\frac{\binom{n-1}{2}}{\binom{k}{2}}\mathbb{P}(I_{1,2}I_{1,3}I_{2,3}=1 \mid W_1 = w_k). \tag{3.39}$$

We now take $k = a\sqrt{n}$, and obtain

$$\bar{C}_n(a\sqrt{n}) = (1+o(1))\frac{n\alpha^2}{a^2}\int_1^\infty\int_1^\infty [1-e^{-xw_k/n^{1/\alpha}}][1-e^{-yw_k/n^{1/\alpha}}][1-e^{-xy/n^{1/\alpha}}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y. \tag{3.40}$$

We first note that, for $k = a\sqrt{n}$,

$$w_{a\sqrt{n}} = n^{1/(2\alpha)}a^{1/\alpha}\Gamma(1-\alpha)^{-1/\alpha} \equiv n^{1/(2\alpha)}a^{1/\alpha}\tau_\alpha. \tag{3.41}$$

We rescale $x$ and $y$ by $n^{-1/2\alpha}$ to obtain

$$\bar{C}_n(a\sqrt{n}) = (1+o(1))\frac{\alpha^2}{a^2}\int_{n^{-1/2\alpha}}^\infty\int_{n^{-1/2\alpha}}^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}][1-e^{-xy}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \tag{3.42}$$

$$= (1+o(1))\frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}][1-e^{-xy}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y,$$

as predicted in (3.11).

## §3.4 Hub and leaf behaviour.

To refine the characterization of clustering in the model, we investigate the limiting behavior of the annealed clustering function for nodes in two distinct regimes of the degree spectrum: *leaf* and *hub* nodes, corresponding to poorly and highly connected ones, respectively. For leaf ($L$) nodes, defined as nodes with vanishing reduced degree $a \to 0$, we find that the function in Eq. (3.6) has the asymptotic plateau

$$\bar{C}_L \equiv \lim_{a\to 0}\bar{C}(a) = 1, \tag{3.43}$$

which implies that in the low-degree limit the local neighborhood of a node is almost fully clustered (excluding of course nodes with $k = 0, 1$). This behaviour can be

intuitively explained in terms of an interconnected core of hubs coexisting with a majority of low-degree nodes, whose few links are largely connected with the hubs. Therefore, low-degree nodes have a high chance to close triangles between neighbours because the hubs are densely connected among themselves. For hub ($H$) nodes, defined as nodes with diverging reduced degree $a \to \infty$, we find (see SI) that the function in Eq. (3.6) decays as

$$\bar{C}(a) \sim \bar{C}_H(a) \equiv 2\Gamma(1-\alpha)\frac{\log a}{a^2}, \quad a \to \infty. \tag{3.44}$$

Figure 3.1 shows that the asymptotics (3.43) and (3.44) are in excellent agreement both with the empirical clustering function (3.4) computed on actual realizations of the random graph and with the full analytical expression (3.6) for the annealed clustering function (3.5). We now prove these claims.



## §3.4.1 Asymptotics of the clustering function

**Lemma 3.4.1 ($\bar{C}(a)$ is bounded by 1).**
*For all $a > 0$, $\bar{C}(a) \leq 1$.*

*Proof.* Estimating $1 - e^{-xy} \leq 1$ bounds $\bar{C}(a)$ by

$$\bar{C}(a) \leq \frac{\alpha^2}{a^2}\Big(\int_0^\infty [1 - e^{-xa^{1/\alpha}\tau_\alpha}]x^{-(\alpha+1)}\mathrm{d}x\Big)^2, \tag{3.45}$$

for which we note that

$$\begin{aligned}
\int_0^\infty [1 - e^{-xa^{1/\alpha}\tau_\alpha}]x^{-(\alpha+1)}\mathrm{d}x &= \int_0^\infty \int_0^{xa^{1/\alpha}\tau_\alpha} e^{-y}\mathrm{d}y x^{-(\alpha+1)}\mathrm{d}x && (3.46)\\
&= \int_0^\infty e^{-y}\int_{ya^{-1/\alpha}/\tau_\alpha}^\infty x^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y && (3.47)\\
&= \frac{1}{\alpha}\int_0^\infty e^{-y}(ya^{-1/\alpha}/\tau_\alpha)^{-\alpha}\mathrm{d}y\\
&= \frac{a\tau_\alpha^\alpha}{\alpha}\int_0^\infty e^{-y}y^{-\alpha}\mathrm{d}y = \frac{a\tau_\alpha^\alpha}{\alpha}\Gamma(1-\alpha)\\
&= \frac{a}{\alpha},
\end{aligned}$$

using that $\tau_\alpha = \Gamma(1-\alpha)^{-1/\alpha}$. Thus, indeed, $\bar{C}(a) \leq 1$.

We next investigate $\bar{C}(a)$ for $a \ll 1$:

**Lemma 3.4.2 ($\bar{C}(a)$ is close to 1 for $a$ small).**
*As $a \searrow 0$,*

$$\bar{C}(a) = 1 + o(1). \tag{3.48}$$

*Proof.* For $a \ll 1$, we write

$$\bar{C}(a) = \frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \tag{3.49}$$
$$-\frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}]e^{-xy}(xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y.$$

The first integral equals 1 by Lemma 3.4.1, so that

$$\bar{C}(a) = 1-\bar{C}_1(a), \tag{3.50}$$

where

$$\bar{C}_1(a) = \frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}]e^{-xy}(xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y. \tag{3.51}$$

We first bound the integrals where either $x$ or $y$ is in $[0, a^{-\varepsilon}]$, for some $\varepsilon > 0$ sufficiently small, as

$$\begin{aligned}\frac{2\alpha^2}{a^2}&\int_0^\infty\int_0^{a^{-\varepsilon}} [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}]e^{-xy}(xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \\ &\le \frac{2\alpha^2}{a^2}\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}]x^{-(\alpha+1)}\mathrm{d}x a^{1/\alpha}\int_0^{a^{-\varepsilon}} y^{-\alpha}\mathrm{d}x \\ &\le \frac{2\alpha}{a(\alpha-1)}a^{1/\alpha-\varepsilon(1-\alpha)} = o(1),\end{aligned} \tag{3.52}$$

using (3.46) and the fact that $1/\alpha - \varepsilon(1-\alpha) > 1$ for $\varepsilon > 0$ sufficiently small since $\alpha \in (0,1)$. Thus,

$$\bar{C}_1(a) = o(1) + \frac{\alpha^2}{a^2}\int_{a^{-\varepsilon}}^\infty\int_{a^{-\varepsilon}}^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}]e^{-xy}(xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y, \tag{3.53}$$

which is bounded by $e^{-a^{-2\varepsilon}} = o(1)$. This shows that $\bar{C}_1(a) = o(1)$ for $a \searrow 0$, and thus proves that $\bar{C}(a) = 1 + o(1)$.

For $a \to \infty$, the integral should be dominated by small $x, y$, so that one might consider to expand $1 - e^{-xy} = xy$, to obtain

$$\begin{aligned}\bar{C}(a) &= (1+o(1))\frac{\alpha^2}{a^2}\int_0^\infty\int_0^\infty [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}](xy)^{-\alpha}\mathrm{d}x\mathrm{d}y \\ &= (1+o(1))\frac{\alpha^2}{a^2}\big(a^{1/\alpha}\tau_\alpha\big)^{2\alpha-2}\Big(\int_0^\infty [1-e^{-x}]x^{-\alpha}\mathrm{d}x\Big)^2.\end{aligned} \tag{3.54}$$

However, the integral that is squared is infinite, so that we need to proceed alternatively. In fact, this analysis is quite intricate, and a careful split is necessary. The main result is as follows:

**Lemma 3.4.3 (Large $a$ asymptotics of $\bar{C}(a)$).**
*As $a \nearrow \infty$,*

$$\bar{C}(a) = 2\Gamma(1-\alpha)\frac{\log a}{a^2} + O(a^{-2}). \tag{3.55}$$

*Proof.* We again split the integral conveniently. We split the integral depending on whether $x$ or $y$ is in $[0, a^{-1/\alpha}/\tau_\alpha]$ or not. This gives

$$\bar{C}(a) = \bar{C}_1(a) + \bar{C}_2(a) + \bar{C}_3(a), \tag{3.56}$$

where

$$\bar{C}_1(a) = \frac{\alpha^2}{a^2} \int_{a^{-1/\alpha}/\tau_\alpha}^{\infty} \int_{a^{-1/\alpha}/\tau_\alpha}^{\infty} [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}][1-e^{-xy}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y, \tag{3.57}$$

$$\bar{C}_2(a) = \frac{\alpha^2}{a^2} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} [1-e^{-xa^{1/\alpha}\tau_\alpha}][1-e^{-ya^{1/\alpha}\tau_\alpha}][1-e^{-xy}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y. \tag{3.58}$$

We bound, again using (3.46) and $1 - e^{-xy} \le xy$,

$$\begin{aligned}\bar{C}_2(a) &\le \frac{\alpha^2}{a^2} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} [1 - e^{-xa^{1/\alpha}\tau_\alpha}][1 - e^{-ya^{1/\alpha}\tau_\alpha}](xy)^{-\alpha}\mathrm{d}x\mathrm{d}y \qquad (3.59)\\ &= \frac{\alpha^2}{a^2}\Big( \int_{0}^{a^{-1/\alpha}/\tau_\alpha} y^{-\alpha}\mathrm{d}y\Big)^2 \\ &= \Theta(1)a^{-2}(a^{-1/\alpha})^{2(1-\alpha)} = o(a^{-2}).\end{aligned}$$

We continue with $\bar{C}_3(a)$, which is the most involved. We bound $1 - e^{-ya^{1/\alpha}\tau_\alpha} \le ya^{1/\alpha}$ to obtain

$$\bar{C}_3(a) \le \frac{2\alpha^2 a^{1/\alpha}}{a^2} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} \int_{a^{-1/\alpha}/\tau_\alpha}^{\infty} (xy)^{-(\alpha+1)} y[1 - e^{-xy}]\mathrm{d}x\mathrm{d}y. \tag{3.60}$$

We bound $1 - e^{-xy} \le (xy) \wedge 1$. The integral where $x \le 1/y$ is bounded by

$$\Theta(1)a^{-2+1/\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} y^{-\alpha+1} \int_{0}^{1/y} x^{-\alpha}\mathrm{d}x\mathrm{d}y \le \Theta(1)a^{-2+1/\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} \mathrm{d}y \le \Theta(1)a^{-2}. \tag{3.61}$$

Similarly, the integral where $x > 1/y$ is bounded by

$$\Theta(1)a^{-2+1/\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} y^{-\alpha} \int_{1/y}^{\infty} x^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \le \Theta(1)a^{-2+1/\alpha} \int_{0}^{a^{-1/\alpha}/\tau_\alpha} \mathrm{d}y \le \Theta(1)a^{-2}. \tag{3.62}$$

This shows that $\bar{C}_3(a) = O(a^{-2})$.

We finally compute $\bar{C}_1(a)$, which we rewrite as

$$\begin{aligned}\bar{C}_1(a) &= \alpha^2\tau_\alpha^{2\alpha} \int_{1}^{\infty} \int_{1}^{\infty} [1 - e^{-x}][1 - e^{-y}][1 - e^{-xya^{-2/\alpha}/\tau_\alpha^2}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \qquad (3.63)\\ &= \tau_\alpha^{2\alpha}\, \mathbb{E}[1 - e^{-W_1W_2a^{-2/\alpha}/\tau_\alpha^2}] + \bar{e}_1(a),\end{aligned}$$

where we let $\bar{e}_1(a)$ denote the contribution due to $e^{-x}$ and $e^{-y}$. We use [3, (4.6)], which states that

$$\mathbb{E}[1 - e^{-\delta W_1 W_2}] = (1 + o(1))\alpha\Gamma(1-\alpha)\delta^\alpha \log(1/\delta). \tag{3.64}$$

This gives

$$\bar{C}_1(a) = \tau_\alpha^{2\alpha}\alpha\tau_\alpha^{-2\alpha}\frac{2}{\alpha}\Gamma(1-\alpha)\frac{\log a}{a^2} + \bar{e}_1(a) = 2\Gamma(1-\alpha)\frac{\log a}{a^2} + \bar{e}_1(a). \tag{3.65}$$

Finally,

$$\begin{aligned}\bar{e}_1(a) &\leq 2\alpha^2\tau_\alpha^{2\alpha}\int_1^\infty\int_1^\infty e^{-x}[1 - e^{-xya^{-2/\alpha}/\tau_\alpha^2}](xy)^{-(\alpha+1)}\mathrm{d}x\mathrm{d}y \\ &= 2\tau_\alpha^{2\alpha}\,\mathbb{E}\left[e^{-W_1}(1 - e^{-W_1W_2a^{-2/\alpha}/\tau_\alpha^2})\right] = O(a^{-2}).\end{aligned} \tag{3.66}$$

This can be seen by letting $X$ be defined by

$$\mathbb{P}(X \in [a,b]) = \frac{1}{\mathbb{E}\left[e^{-W_1}\right]}\,\mathbb{E}\left[e^{-W_1}\mathbb{1}_{\{X\in[a,b]\}}\right], \tag{3.67}$$

so that

$$\bar{e}_1(a) = \Theta(1)\,\mathbb{E}\left[1 - e^{-XWa^{-2/\alpha}/\tau_\alpha^2}\right]. \tag{3.68}$$

Since $X$ has thin tails (it has an exponential moment), the random variable $XW$ is regularly varying with exponent $\alpha$, so that

$$\mathbb{E}\left[1 - e^{-XWa^{-2/\alpha}/\tau_\alpha^2}\right] = \Theta(a^{-2}), \tag{3.69}$$

as claimed. The above follows from the fact that $XW$ is again heavy-tailed with tail exponent $\alpha$, as proved in [161, Proposition 5.2].

## §3.5 Average Clustering Coefficient.

The previous results finally allow us to estimate the overall expected value $\bar{C}$ of the average local clustering coefficient $C$ by averaging the annealed clustering function $\bar{C}(a)$ over the distribution $P(a)$ of the reduced degree $a$ as

$$\bar{C} \equiv \int da\ \bar{C}(a)P(a) \simeq \int_{a\in L} \mathrm{d}a\ \bar{C}_L(a)P(a) + \int_{a\in H} \mathrm{d}a\ \bar{C}_H(a)P(a)$$

where we have formally split the integral into two contributions, produced by leaf ($a \in L$) and hub ($a \in H$) nodes respectively, consistently with the two regimes characterized asymptotically in (3.43) and (3.44). A rigorous definition of this split is provided in next section 3.5.1, where we also show that, in order to compute $\bar{C}$ exactly, it is not necessary to identify precisely the crossover value(s) of $a$ separating the two regions. Notably, it turns out that the integral over $a \in H$ in (3.70) vanishes, implying that hubs no not contribute to the overall clustering. The remaining contribution

of the leaf nodes $a \in L$ is always finite, but its precise value depends on whether (following our initial discussion on the two possible definitions of $C$) nodes with degree $k < 2$ are considered as having $C_{k<2} = 0$ and hence included in the interval $L$ (in which case $\bar{C}_L(a) \equiv 0$ for $0 \leq a \leq 1/\sqrt{n}$) or not (in which case $C_{k<2}$ is undefined and not included in the integration). In next section 3.5.1 we find that, depending on the choice,

$$\lim_{n\to\infty} \bar{C} = \begin{cases} 1 & \text{if } C_{k<2} \text{ undefined} \\ 1 - r_{0/1} & \text{if } C_{k<2} = 0 \end{cases}, \tag{3.70}$$

where $r_{0/1}$ is the expected fraction of nodes with degree 0 or 1. We plot the two versions of the average local clustering coefficient in Fig. 3.2, confirming the agreement with the respective theoretical result. This result is quite remarkable, as it means that the network is fully clustered, in the sense that for almost all nodes that can form triangles, their neighbours are connected and actually close the triangle.



## §3.5.1 Computation of Average Clustering Coefficient

The limiting asymptotics for hubs and leaves allows us to derive an estimate of the average clustering coefficient of the network $\bar{C}$. Indeed, one needs to integrate the annealed clustering function $\bar{C}(a)$ over the distribution $P(a)$ of the reduced degree $a$, and from figure 3.1 it is clear that one can split the integral into the two leaf ($L$) and hub ($H$) regimes as

$$\bar{C} = \int_{a_{min}}^{\infty} \mathrm{d}a \bar{C}(a) P(a) = \int_{a_{min}}^{a^*} \mathrm{d}a \bar{C}_L(a) P(a) + \int_{a^*}^{\infty} \mathrm{d}a \bar{C}_H(a) P(a), \tag{3.71}$$

where $C_H(a) \sim 2\Gamma(1-\alpha)\frac{\log a}{a^2}$ (for $a > a^*$) is the asymptotic clustering function of hubs, while $C_L(a) \sim 1$ (for $a_{min} < a < a^*$) is the asymptotic plateau for leaves, $a_{min}$ is the minimal reduced degree above which the clustering function starts to have meaning, and $a^*$ is an appropriate crossover value between the two regimes. This crossover is visible in Figure 3.1, although its value is not easily identified. Luckily, as we will show in the following computations, the actual value of $a^*$ is not necessary for the computation of the clustering to go through, and we can proceed without specifying it.

The main issue in the evaluation of the integral in Eq. (3.71) is that the degree distribution is not known everywhere but only in the tail [3], where it has the asymptotic behaviour $\Gamma(1-\alpha)k^{-2}$. Transforming to the distribution $P(a)$ of the reduced degree $a = \frac{k}{\sqrt{n}}$,

$$P(a) \sim \frac{\Gamma(1-\alpha)}{\sqrt{n}} a^{-2}. \tag{3.72}$$

Nonetheless, the first integral can be evaluated by realizing that by definition it evaluates to the fraction of leaves nodes or, equivalently

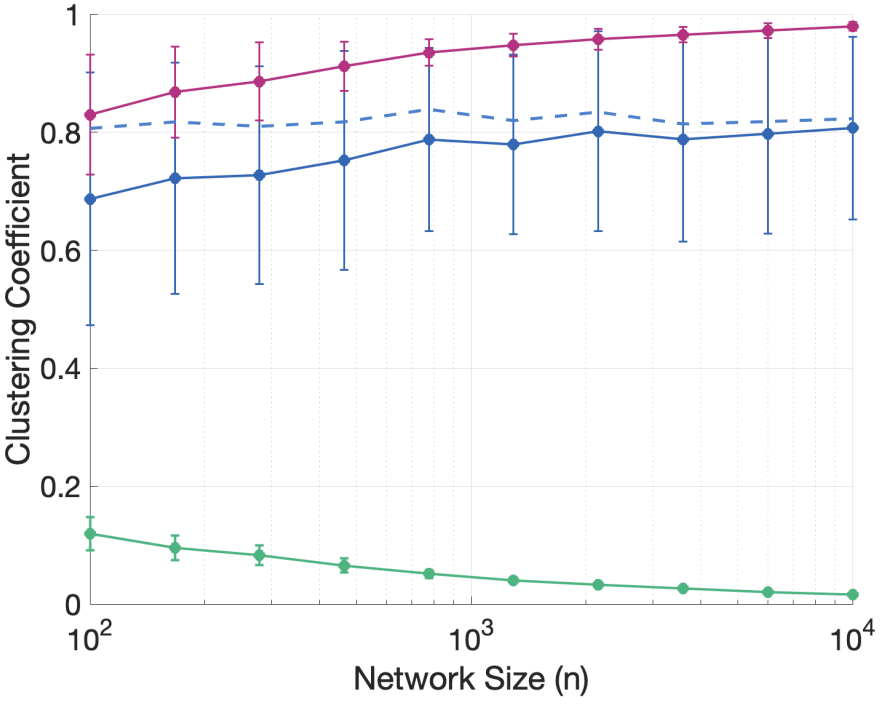


*Figure 3.2: Average local clustering coefficient $C$ versus network size $n$, for $\alpha = 0.5$. For each sampling of the $n$ node weights, a single adjacency matrix is realized, and then weights are redrawn (up to $10^3$ times for each $n$). In purple the clustering coefficient defined excluding the nodes with degree 0 and 1. In blue, the solid line is the average clustering coefficient evaluated by setting the clustering function to zero for $k = 0, 1$. We highlight that the error bars over this quantity do not shrink as $n$ increases. The dashed blue line is 1 minus the average of $r_{0/1}$ over realizations. Finally, in green the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, and we notice that in this case the error shrinks as desired.*

$$\int_{a_{min}}^{a^*} da \bar{C}_L(a) P(a) = \int_{a_{min}}^{a^*} da P(a) = \begin{cases} 1 - \int_{a^*}^{\infty} da P(a) & \text{if } C_{k<2} \text{ undefined} \\ 1 - r_{0/1} - \int_{a^*}^{\infty} da P(a) & \text{if } C_{k<2} = 0 \end{cases} \tag{3.73}$$

and the last term can be recomprised in the second integral as it is evaluated on the same interval, so that the average clustering coefficient has the expression

$$\bar{C} = \begin{cases} 1 + \int_{a^*}^{\infty} P(a)(\bar{C}_H(a) - 1) & \text{if } C_{k<2} \text{ undefined} \\ 1 - r_{0/1} + \int_{a^*}^{\infty} da P(a)(\bar{C}_H(a) - 1) & \text{if } C_{k<2} = 0 \end{cases} \tag{3.74}$$

We can evaluate this last integral as

$$\begin{aligned}\int_{a^*}^{\infty} da P(a)(\bar{C}_H(a)-1) &= \int_{a^*}^{\infty} da \left[2\Gamma(1-\alpha)\frac{\log a}{a^2}-1\right]\frac{\Gamma(1-\alpha)}{a^2\sqrt{n}} \\ &= \frac{\Gamma(1-\alpha)}{\sqrt{n}}\frac{6\Gamma(1-\alpha)\log a^* + 2\Gamma(1-\alpha) - 9(a^*)^2}{9(a^*)^3} \\ &= o(1),\end{aligned} \tag{3.75}$$

and see that it is suppressed by a square root of $n$ term, whatever the value of $a^*$.

## §3.6 Lack of self-averaging.

The result in Eq.(3.70) is calculated conditionally on the realized weights $\{w_i\}_{i=1}^n$. It shows that, if nodes with $k < 2$ are excluded, $\bar{C}$ converges to the (deterministic) maximum value 1, with vanishing fluctuations. By contrast, as we show now, for finite but large $n$ the fractions of nodes with degree zero and one concentrate around functions of the rescaled total weight $S_n = \delta_n \sum_{j=1}^n w_j$, which in turn converges in distribution to an $\alpha$-stable random variable and therefore remains a fluctuating quantity. If one looks instead at isolated and pendant nodes, the fractions concentrate around functions of the rescaled total weight $S_n = \delta_n \sum_{j=1}^n w_j$, which in turn converges in distribution to an $\alpha$-stable random variable. This remains a fluctuating quantity even in the large $n$ limit, as proven in the followings and confirmed numerically in Figure 3.2. This means that, for given $n$ and $\alpha$, different realizations of the weights produce different values of $r_{0/1}$ and of $\bar{C}$, if nodes with $k < 2$ are included. Therefore, under this second definition, $\bar{C}$ converges in distribution to the random variable $1 - r_{0/1}$, remaining macroscopically fluctuating even in the large graph limit. This behaviour follows from the infinite-mean nature of the weights and is a notable manifestation of the (rather unique) lack of self-averaging in macroscopic network properties (in particular the fractions of nodes with degree $k < 2$), which is not typically observed in models with independent edges and finite-mean weights. Notably, the lack of self-averaging extends to further 'low degree' nodes. This means that any macroscopic property that is a function of the full degree distribution, and not just its tail, will in general keep memory of the fluctuations of the fractions of low-degree nodes. Figure 3.3 confirms that other macroscopic properties (spectral gap, largest connected component size) besides the clustering coefficient keep fluctuating in the thermodynamic limit. We now prove these claims, and show these results are confirmed numerically in Figs. 3.2 3.3.

### §3.6.1 Computation of $r_{0/1}$

We start by computing the fraction of nodes that are disconnected from the graph as

$$r_0 \equiv \frac{1}{n}\sum_{i=1}^{n}\prod_{j:\, j\neq i}(1-p_{ij}) = \frac{1}{n}\sum_{i=1}^{n} e^{-\delta_n w_i \sum_{j\neq i} w_j}, \tag{3.76}$$

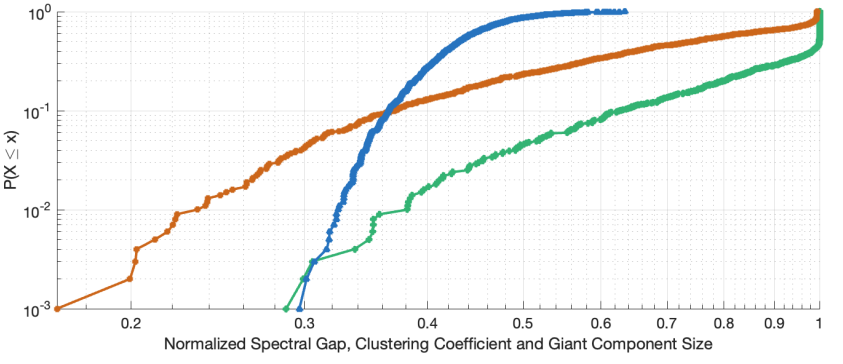


*Figure 3.3:* ***Cumulative distributions of network properties across*** $10^3$ ***realizations with resampled weights*** *(*$n = 10^4$*,* $\alpha = 0.5$*). Orange: average local clustering coefficient including isolated and pendant nodes. Green: fraction of nodes in the largest connected component. Blue: normalized spectral gap (difference between the two largest eigenvalues of the adjacency matrix, divided by the largest one).*



where we recall that $\delta_n = n^{-1/\alpha}$. Since the $j = i$ term in the sum in the exponential makes little difference, we can approximate

$$r_0 \approx \frac{1}{n}\sum_{i=1}^{n} e^{-\delta_n w_i \sum_j w_j}. \tag{3.77}$$

We next use that most of the weights have normal size, and they contribute most to the sum over $i \in \{1, \dots, n\}$. Thus,

$$r_0 \approx \mathbb{E}\left[e^{-WS_n} \mid S_n\right], \tag{3.78}$$

where we take the expectation over the Pareto variable $W$ in the above, we abbreviate

$$S_n = \delta_n \sum_{j=1}^{n} w_j, \tag{3.79}$$

and condition on this total rescaled weight. We note that the random variable $S_n$ converges in distribution to an $\alpha$-stable distribution, since the random variables $(w_i)_{i=1}^n$ are Pareto variables with infinite mean. Thus,

$$r_0 \approx \mathbb{E}\left[e^{-W\mathcal{S}} \mid \mathcal{S}\right], \tag{3.80}$$

where $\mathcal{S}$ is an $\alpha$-stable random variable. Equation (3.80) requires some interpretation. What it means is that the proportion of vertices with $r_0$ degree 0, given the vertex weights $(w_i)_{i=1}^n$, is close to the rhs of (3.78), i.e., it holds that $r_0 \approx \mathbb{E}\left[e^{-WS_n} \mid S_n\right]$, where the rescaled total weight $S_n$ is defined in (3.79). In particular, this means that, when simulating the network several times (and thus redrawing the vertex weights), the proportion of vertices with degree 0 *fluctuates macroscopically*, even in the large

graph limit. This is a type of *non-self-averaging* phenomenon that is rather unusual in network science. The simulations shown in Figures 3.6 3.7 3.8 confirm that (3.80) is an excellent approximation of $r_0$, and that it remains truly random in the large-graph limit.

We can also use (3.80) us to relate our approximation to [3, Proposition 4], in which $\mathbb{E}[r_0]$ is shown to be approximated by

$$\mathbb{E}[r_0] \approx \mathbb{E}\left[e^{-\Gamma(1-\alpha)W^\alpha}\right], \tag{3.81}$$

which indeed equals $\mathbb{E}\left[e^{-W\mathcal{S}}\right]$, since, for every $a \geq 0$,

$$\mathbb{E}\left[e^{-a\mathcal{S}}\right] = e^{-\Gamma(1-\alpha)a^\alpha}, \tag{3.82}$$

which follows from [3, Theorem 5]. For a practical computation of the approximation of $r_0$, we use the fact, again for $a \geq 0$,

$$\mathbb{E}\left[e^{-aW}\right] = \alpha a^\alpha \Gamma(-\alpha, a), \tag{3.83}$$

where we recall that $\Gamma(r, a)$ is the incomplete gamma function. Applying this to (3.78), we are led to

$$r_0 \approx \alpha S_n^\alpha \Gamma(-\alpha, S_n). \tag{3.84}$$

.

We show in Figures 3.6 3.7 3.8 that (3.84) is in very good accordance with the actual fraction of disconnected nodes (central panels).

We extend the above analysis to the proportion $r_1$ of vertices of degree 1. We start from

$$\begin{aligned} r_1 &\equiv \frac{1}{n}\sum_{i,k:\, i\neq k} p_{ik} \prod_{j:\, j\neq i,k} (1-p_{ij}) \\ &= \frac{1}{n}\sum_{i,k:\, i\neq k} (1-e^{-\delta_n w_i w_k}) e^{-\delta_n w_i \sum_{j:\, j\neq i,k} w_j} \\ &\approx \sum_k \mathbb{E}\left[(1-e^{-W\delta_n w_k}) e^{-W\delta_n \sum_{j:\, j\neq k} w_j}\right], \end{aligned} \tag{3.85}$$

where the asymptotics follows from the fact that the main contribution in the sum over $i$ arises from normal weight vertices, so that the average over $i$ can be replaced by an expectation with respect to $W$.

Note that the dominant contributions to the sum over $k$ in (3.85) correspond to those $k$ for which $w_k$ is of the order $1/\delta_n = n^{1/\alpha}$, which means that these are the maximal-weight vertices. This will be the guiding principle to derive the asymptotics of the above formula.

For simplicity, and without loss of generality, we will assume that the weights are *ordered* from large to small, so that $w_1 \geq w_2 \cdots \geq w_n$. We can rewrite

$$\delta_n \sum_{j:\, j\neq k} w_j = S_n - \delta_n w_k = S_n - y_k^{(n)}, \tag{3.86}$$

where we recall $S_n$ from (3.79), and we define the rescaled vertex weights $(y_k^{(n)})_{k\geq 1}$ as

$$y_k^{(n)} = \delta_n w_k. \tag{3.87}$$

In terms of these definitions, we obtain the finite-size approximation

$$r_1 \approx \sum_k \mathbb{E}\left[(1-e^{-Wy_k^{(n)}})e^{-W(S_n-y_k^{(n)})} \mid (y_k^{(n)})_{k\geq 1}\right]. \tag{3.88}$$

This approximation is inspired by the asymptotics properties of $r_1$. Indeed, as $n\to\infty$, we have the convergence in distribution

$$\Big(S_n, (y_k^{(n)})_{k\geq 1}\Big)\Big(\sum_{j\geq 1}\Gamma_j^{-1/\alpha}, (\Gamma_k^{-1/\alpha})_{k\geq 1}\Big), \tag{3.89}$$

where  denotes convergence in distribution in the product topology, $(\Gamma_k)_{k\geq 1}$ are gamma random variables, i.e.,

$$\Gamma_k = \sum_{i=1}^{k} E_i, \tag{3.90}$$

where $(E_i)_{i\geq 1}$ are independent and identically distributed exponential random variables with mean 1. Because of this, we see that the random variables in (3.88) converge in distribution in the large graph limit, so that

$$r_1 \sum_k \mathbb{E}\left[(1-e^{-Wy_k})e^{-W(\mathcal{S}-y_k)} \mid (y_k)_{k\geq 1}\right], \tag{3.91}$$

where

$$\Big(\mathcal{S}, (y_k)_{k\geq 1}\Big) \equiv \Big(\sum_{j\geq 1}\Gamma_j^{-1/\alpha}, (\Gamma_k^{-1/\alpha})_{k\geq 1}\Big). \tag{3.92}$$

Thus, as for $r_0$, also $r_1$ converges in *distribution*, but it remains to fluctuate even in the large graph limit.

For a practical computation of the approximation of $r_0$, we again use (3.83) on the representation

$$r_1 \approx \sum_{k=1}^{n} \mathbb{E}\left[e^{-W(S_n-y_k^{(n)})} - e^{-WS_n} \mid (y_k^{(n)})_{k\geq 1}\right], \tag{3.93}$$

to arrive at

$$r_1 \approx \alpha \sum_{k=1}^{n}\Big[(S_n-y_k^{(n)})^{\alpha}\Gamma(-\alpha, S_n-y_k^{(n)}) - S_n^{\alpha}\Gamma(-\alpha, S_n)\Big], \tag{3.94}$$

which has the advantage that it no longer contains the expectation with respect to $W$. We show in Figures 3.6 3.7 3.8 that (3.94) is in very good accordance with the actual fraction of disconnected nodes (bottom panels).

## §3.7 Lack of self-averaging in the fraction of next low-degree nodes

The fact that both $r_0$ and $r_1$ do not converge to a limiting deterministic value, but rather *in distribution* to a random variable whose variance is non-vanishing even when $n$ tends to infinity, is a peculiar aspect of the MSM. This property applies not only to isolated or pendant nodes, but also to all the next 'low degree' nodes. Indeed, one can estimate $r_k$, the fraction of nodes whose degree is $k$, generalizing (3.76) and (3.85) as follows:

$$r_k \equiv \frac{1}{n} \sum_{\substack{i_0,i_1,\ldots,i_k \\ (i_r \neq i_s \,\forall r,s)}} p_{i_0 i_1} p_{i_0 i_2} \cdots p_{i_0 i_k} \prod_{j\,:\, j \neq i_0,i_1,\ldots i_k} (1 - p_{i_0 j}) \tag{3.95}$$

$$= \frac{1}{n} \sum_{\substack{i_0,i_1,\ldots,i_k \\ (i_r \neq i_s \,\forall r,s)}} \left(1 - e^{-\delta_n w_{i_0} w_{i_1}}\right) \left(1 - e^{-\delta_n w_{i_0} w_{i_2}}\right) \cdots \left(1 - e^{-\delta_n w_{i_0} w_{i_k}}\right) e^{-\delta_n w_{i_0} \sum_{j\,:\, j \neq i_0,i_1,\ldots i_k} w_j}$$



The key point now is that if one wants to proceed as in the last step of (3.85) and replace the sum over $i_0$ divided by $n$ with the expectation over the whole sample, one should ensure that the sample without $k$ weights is still representative of a Pareto distribution. This happens if $k$ is small compared to $n$. In this case, we can proceed as in (3.85) with the approximation

$$r_k \approx \sum_{\substack{i_1,i_2,\ldots,i_k \\ (i_r \neq i_s \,\forall r,s)}} \mathbb{E}\left[(1 - e^{-W\delta_n w_{i_1}})(1 - e^{-W\delta_n w_{i_2}}) \cdots (1 - e^{-W\delta_n w_{i_k}}) e^{-W\delta_n \sum_{j\,:\, j \neq i_0,i_1,\ldots i_k} w_j}\right]$$

$$= \sum_{\substack{i_1,i_2,\ldots,i_k \\ (i_r \neq i_s \,\forall r,s)}} \mathbb{E}\left[(1 - e^{-W y_{i_1}^{(n)}})(1 - e^{-W y_{i_2}^{(n)}}) \cdots (1 - e^{-W y_{i_k}^{(n)}}) e^{-W\left(S_n - \sum_{j=i_1,i_2,\ldots i_k} y_j^{(n)}\right)}\right].$$

conditional on the realized values of the various $y_j^{(n)}$. Again, as $n \to \infty$, those random variables will converge *in distribution* to the Gamma variables (3.89), and also $r_k$ will be converging *in distribution* to the quantity

$$r_k \sum_{\substack{i_1,i_2,\ldots i_k \\ (i_r \neq i_s \,\forall r,s)}} \mathbb{E}\left[(1 - e^{-W y_{i_1}})(1 - e^{-W y_{i_2}}) \ldots (1 - e^{-W y_{i_k}}) e^{-W\left(\mathcal{S} - \sum_{j=i_1,i_2,\ldots i_k} y_j\right)}\right], \tag{3.96}$$

thus still fluctuating macroscopically even in the large $n$ limit, as we confirm in Figures 3.4 and 3.5. Indeed the figures show that the dispersion of $r_k$ is significant, with the fluctuations shrinking as $k$ (and not $n$) increases, suggesting that, for larger values of the degree, self-averaging is eventually restored. We leave the thorough study of this intriguing phenomenon to future works on the MSM.

## §3.8 Extension to stable weights

In this section we briefly discuss how our results can be extended to the case in which the weights are sampled from a genuine $\alpha$-stable distribution $\tilde{\rho}_\alpha(w)$ with positive

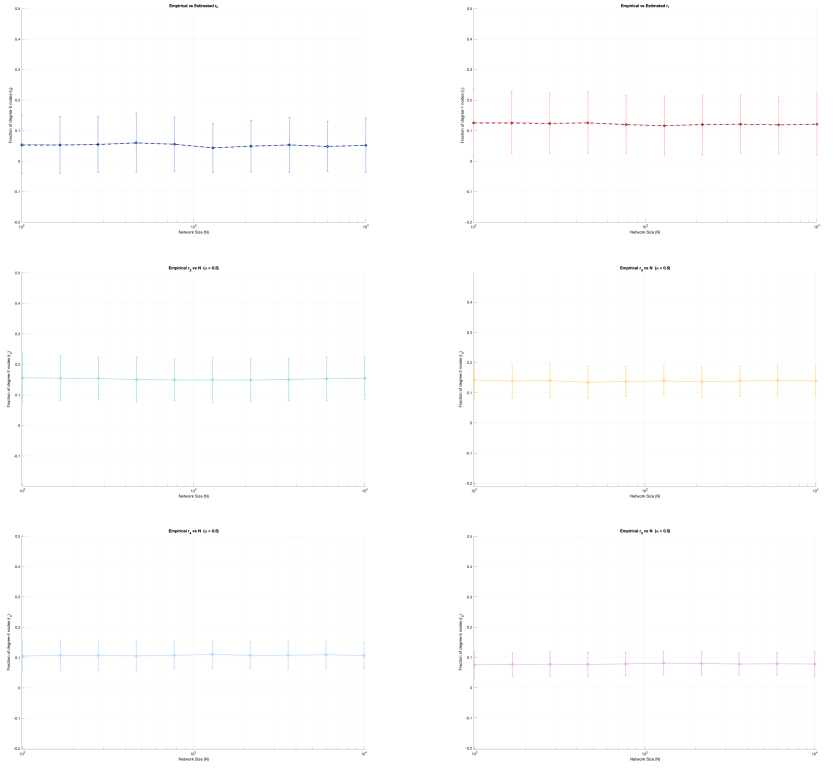

*Figure 3.4:* ***Fractions of low-degree (up to*** $k = 5$***) nodes versus network size, with*** $\alpha = 0.5$**.** *From left to right and top to bottom, we plot the fractions of nodes with a specific low degree (respectively* $k = 0, 1, 2, 3, 4, 5$*) across a sample of* 1000 *different realizations for each value of* $n$*. For the first two figures we include the theoretical estimate we explicitly computed in* (3.84) (3.88)*. We notice that the average value across different values of* $n$ *is roughly constant, while the dispersion of values around that average does not shrink as* $n$ *increases, as one would expect by assuming the self averaging of the quantities. We also notice that the error bars shrink as* $k$ *increases, suggesting that, for larger-degree nodes, self averaging eventually occurs.*

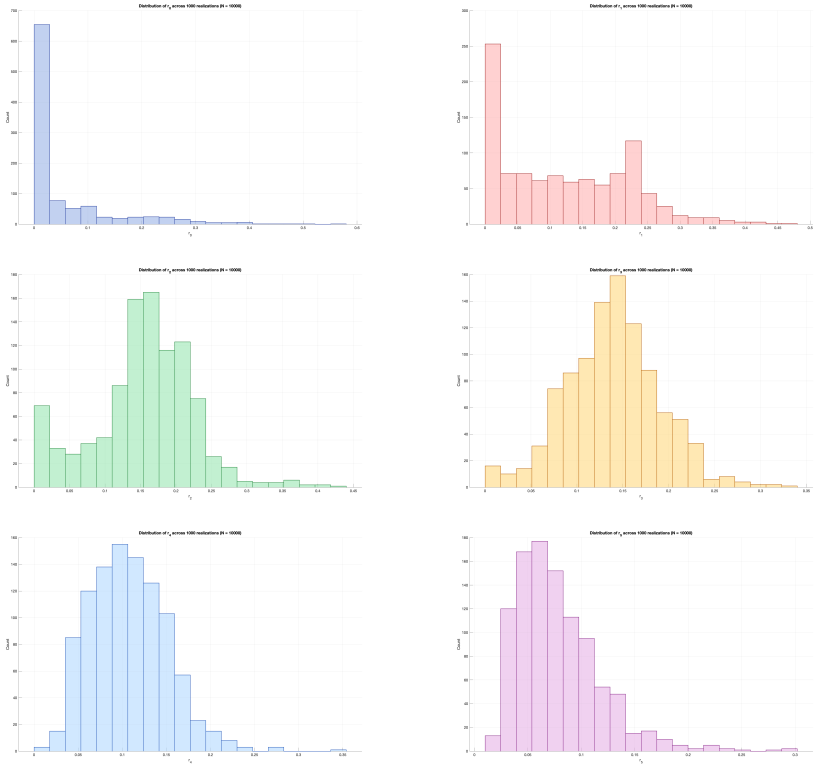

*Figure 3.5:* ***Distribution of the fraction of low-degree (up to*** $k = 5$***) nodes, for*** $\alpha = 0.5$***,*** $n = 1000$ ***(over a sample of*** $1000$ ***realizations).*** *From left to right and top to bottom, we plot the distribution of the fractions of nodes with low degree (respectively* $k = 0, 1, 2, 3, 4, 5$*), across a sample of* $1000$ *different realizations for* $n = 10^4$*. The sample variance* $\sigma_k$ *is respectively* $\sigma_0 = 0.09$*,* $\sigma_1 = 0.1$*,* $\sigma_2 = 0.07$*,* $\sigma_3 = 0.05$*,* $\sigma_4 = 0.04$*,* $\sigma_5 = 0.03$*, indicating that the variance shrinks for larger values of the degree* $k$*.*

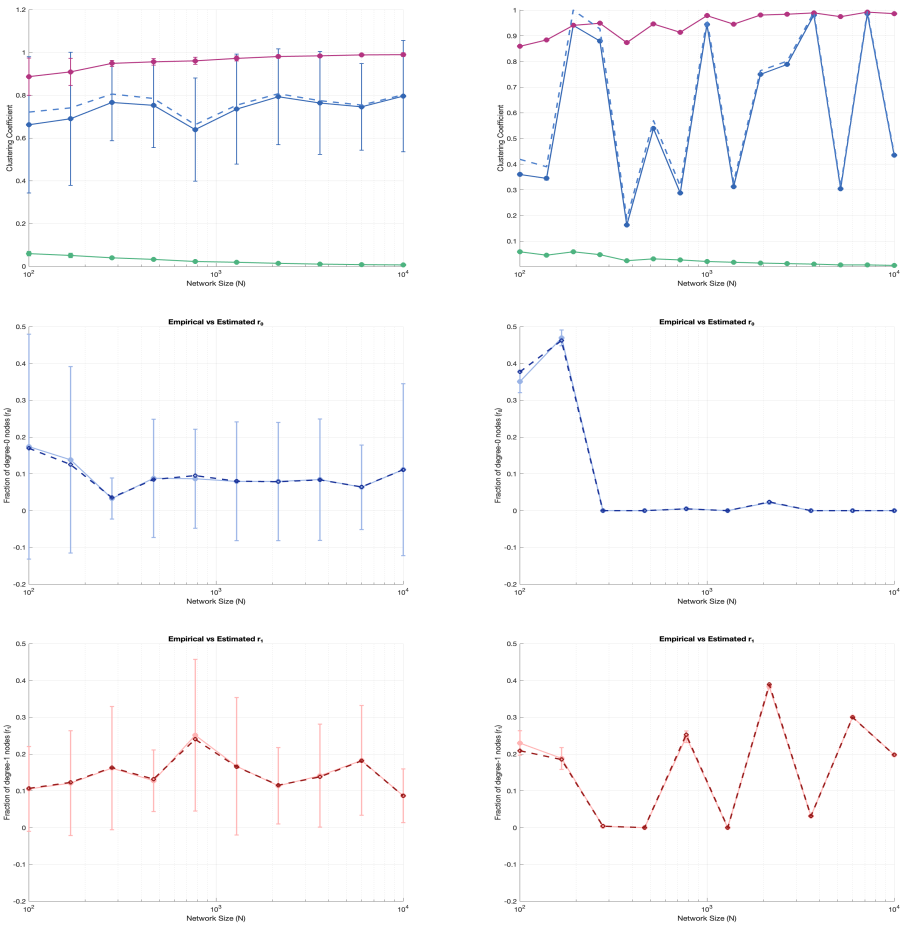


*Figure 3.6: Numerics for $\alpha = 0.3$ (Pareto). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each $n$ weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of $r_0$, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of $r_1$, while the dashed darker one is the approximation we derive in eq (3.88)*

support and diverging mean. Throughout, we follow the formalism introduced in [22], and in particular we adopt the first parametrization defined there, corresponding to $k = 0$.

For stable laws, our arguments can be adapted by replacing the exact expression for the Pareto weight density, $\rho_\alpha(w) = \alpha w^{-(\alpha+1)}$ for $w \geq 1$ in (3.7), with an asymptotic

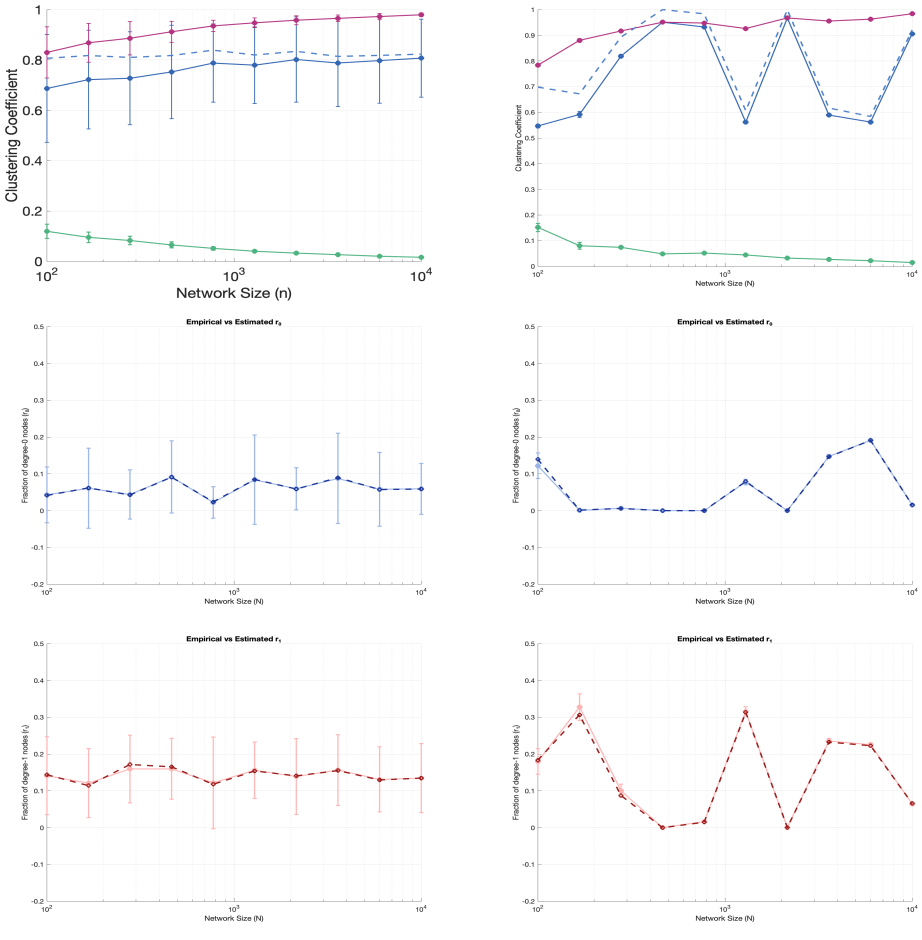


*Figure 3.7: Numerics for $\alpha = 0.5$ (Pareto). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each $n$ weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of $r_0$, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of $r_1$, while the dashed darker one is the approximation we derive in eq (3.88)*

relation. Specifically, Theorem 1.2 of [22] shows that the density of an $\alpha$-stable law satisfies

$$\tilde{\rho}_\alpha(w) = (1 + o(1)) \frac{\alpha \gamma^\alpha (1+\beta) c_\alpha}{w^{\alpha+1}}, \qquad w \geq 1, \tag{3.97}$$

where the tail parameter satisfies $\alpha \in (0, 1)$, $c_\alpha = \frac{\Gamma(\alpha)}{\pi} \sin\left(\frac{\pi\alpha}{2}\right)$, $\gamma$ denotes the scale parameter, and $\beta$ is the skewness parameter of the stable distribution.

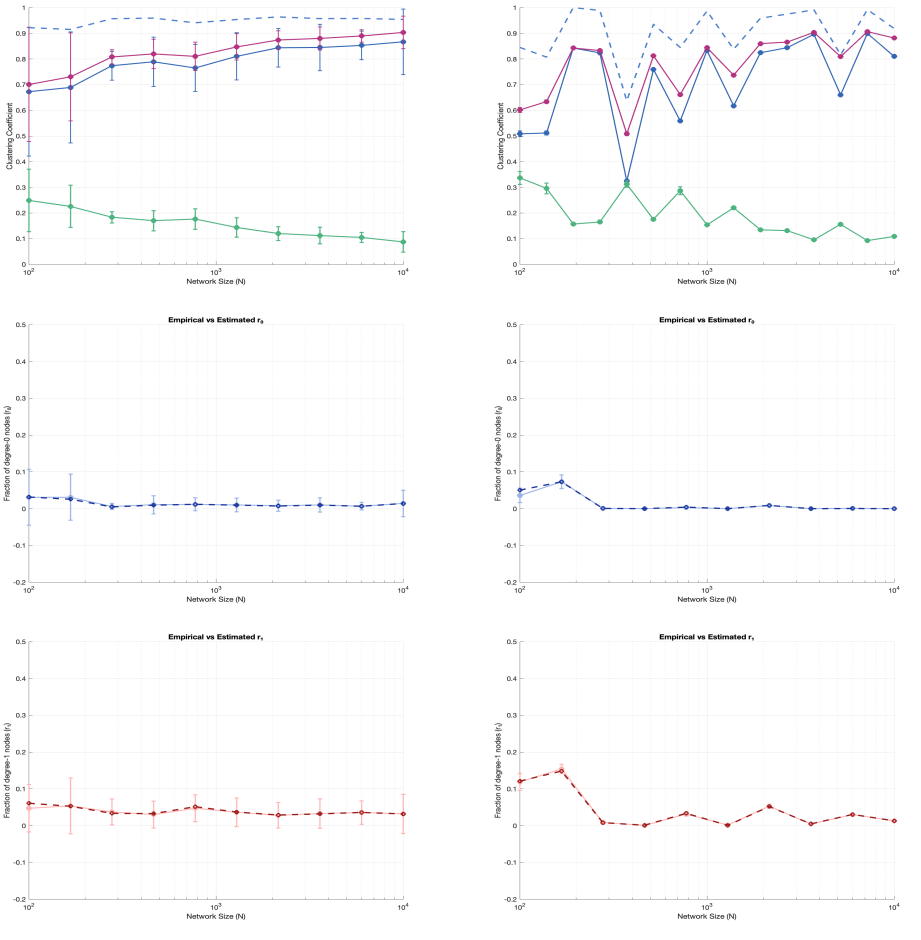


*Figure 3.8: Numerics for $\alpha = 0.7$ (Pareto). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each $n$ weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of $r_0$, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of $r_1$, while the dashed darker one is the approximation we derive in eq (3.88)*

In order to work with a distribution supported on the positive real axis, we choose $\beta = 1$ and set the infimum of the support to zero. This choice fixes the location parameter to be $\delta = \gamma \tan\left(\frac{\pi\alpha}{2}\right)$. The only remaining free parameter is therefore the scale $\gamma$, which we determine by requiring that the tails of the stable and Pareto

distributions coincide asymptotically. Imposing this matching condition yields

$$\gamma = \left[ \frac{\pi}{2\,\Gamma(\alpha)\sin\left(\frac{\pi\alpha}{2}\right)} \right]^{1/\alpha} . \tag{3.98}$$

With this choice of parameters, all the results we derived for Pareto-distributed weights extend directly to the case of stable sampling. In particular, since the integrals appearing in Section 3.3.2 are dominated by contributions from large values of $w$, the asymptotic estimates derived in Section 3.5.1 remain valid. Similarly, the arguments used in Section 3.6.1 to compute the fraction of disconnected nodes and nodes of degree one carry over almost verbatim to the stable case. The only modification is the appearance of a constant $c$ in (3.89), which is now directly related to the scale parameter of the underlying stable law.

Indeed, in the convergence statement (3.89), the partial sum $S_n$ has a strictly $\alpha$-stable distribution for every $n$. Interpreting $S_n$ as the value at time 1 of a stable Lévy process, the sum $c\sum_{j\geq 1}\Gamma_j^{-1/\alpha}$ can be viewed as the collection of jump sizes of the process, ordered by decreasing magnitude. Correspondingly, the rescaled weights $n^{-1/\alpha}W_i$ represent the increments of the process over the time intervals $[(i-1)/n, i/n]$, and the variables $(y_k^{(n)})_{k=1}^n$ in (3.87) can be interpreted as the ordered increments over these intervals. Since, with high probability and for large $n$, each interval contains at most one large jump, it follows that (3.89) continues to hold in this setting, with the first coordinate (corresponding to the sum) having exactly the same distribution as in the Pareto case.

Finally, in Figures 3.10, 3.11, 3.12, and 3.13, we provide numerical evidence confirming that the results presented in the main text remain valid when the weights are drawn from an $\alpha$-stable distribution, in agreement with the arguments given above.

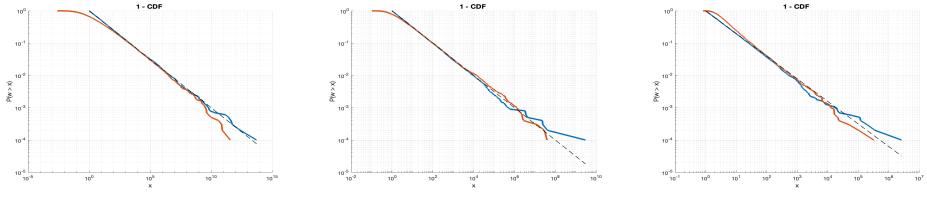


*Figure 3.9:* ***Comparison between pure Pareto and Stable distribution in the tail behaviour****. Using the parameter of* (3.98)*, we can match the tails of the two distributions. From left to right* $\alpha = 0.3, 0.5, 0.7$*, sample of* $10^4$*.*

## §3.9 Concluding Remarks.

This work contributes to the open debate about whether, in models with conditionally independent edges, the coexistence of sparsity and local clustering (along with other desirable features such as a broad degree distribution and the small-world property) can be attained *without geometry*, i.e. without assuming that the connection probability $p_{ij}$ depends on the node attributes $w_i$ and $w_j$ necessarily through a function

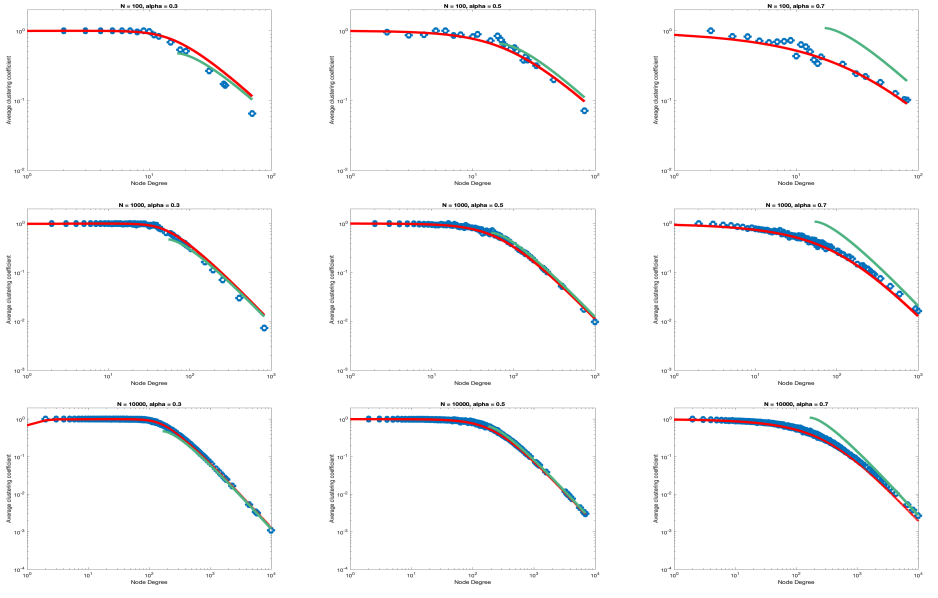


*Figure 3.10:* ***C****lustering functions for different values of $n$ and $\alpha$ for the MSM model* (3.1) *with Stable weights* (3.2)*. Blue circles: empirical clustering function* (3.4) *(versus the reduced degree $a = k/\sqrt{n}$) computed on actual realized graphs sampled from the model (obtained by sampling the weights once, and sampling the graph once conditionally on the realized weights). Red curves: our analytical expression* (3.6) *for the annealed clustering function. Green curves: our asymptotic calculation* (3.44) *valid for diverging reduced degrees. We evaluate and plot all functions only for degrees larger than* 1*, to avoid ambiguities in the definition of the clustering coefficient for $k < 2$. From top to bottom: $n = 10^2, 10^3, 10^4$. From left to right: $\alpha = 0.3, 0.5, 0.7$.*

of some metric distance $d(w_i, w_j)$ [14, 15, 16, 17, 19, 18, 21, 87, 91]. We considered a geometry-free model with independent edges and infinite-mean fitness [1, 2], recently introduced in the context of network renormalization [54], and calculated its full clustering function – rigorously proving the resulting *finite* local clustering in the large $n$ limit and highlighting that clustering does *not* imply geometry. The infinite-mean nature of the fitness arises from a requirement of model invariance under node aggregation [1] and generates other realistic properties including sparsity, a power-law degree distribution [2], and the breakdown of self-averaging of various network properties, which is a novel result characterized here. The model then provides, as an alternative to geometry, the hypothesis of node aggregation invariance as a basic mechanism producing realistic network properties.

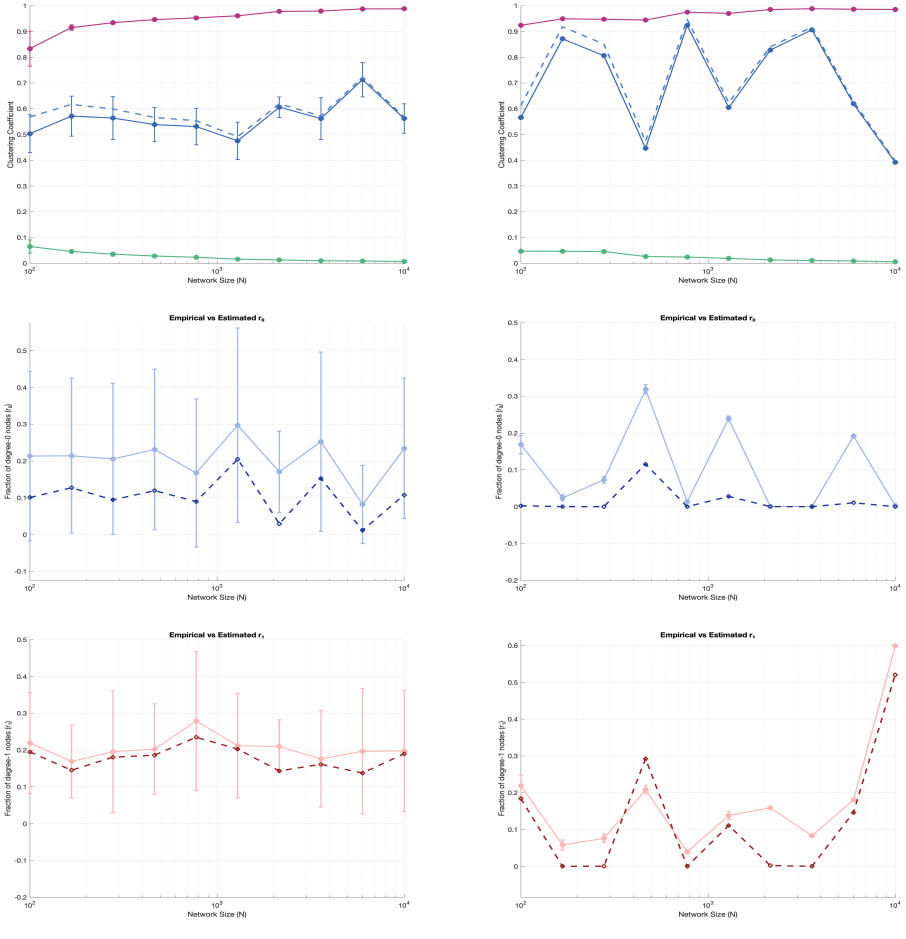


*Figure 3.11: Numerics for $\alpha = 0.3$ (stable). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each $n$ weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of $r_0$, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of $r_1$, while the dashed darker one is the approximation we derive in eq (3.88)*

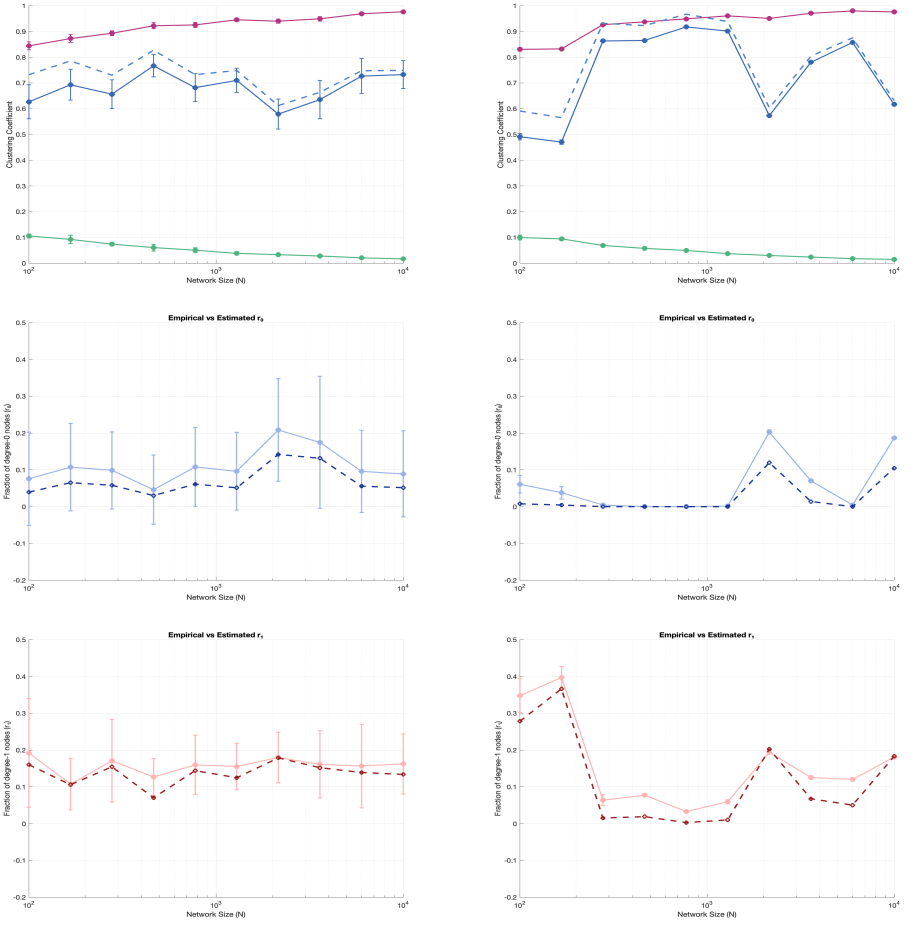


*Figure 3.12: Numerics for $\alpha = 0.5$ (stable). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each $n$ weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of $r_0$, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of $r_1$, while the dashed darker one is the approximation we derive in eq (3.88)*

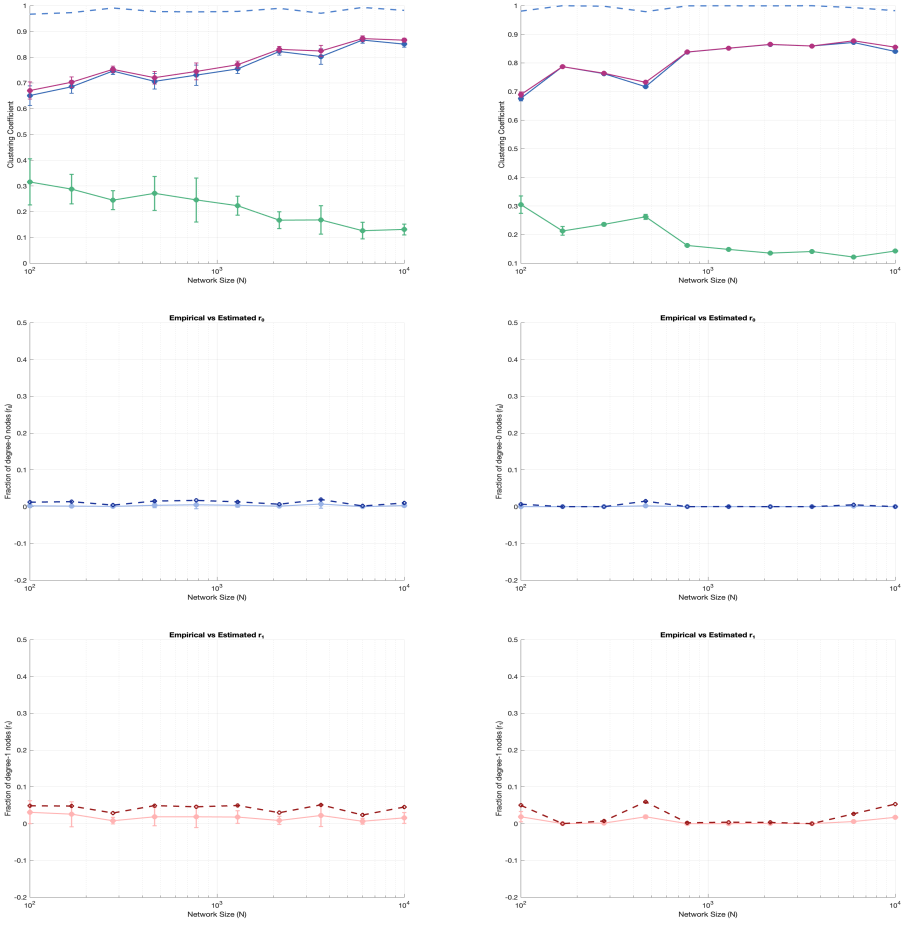


*Figure 3.13: Numerics for $\alpha = 0.7$ (stable). On the left, simulations are done by resampling weights every time an actual network is realized (sample of 10), while on the left for each n weights are extracted only once, and 10 different adjacency matrices are realized from the same weights. On top, the same analysis performed in figure 3.2: in purple the clustering coefficient computed by excluding nodes with $k < 2$; in blue the clustering coefficient including those nodes and $1 - r_{0/1}$, respectively the solid and dashed lines; in green with error bars the difference between $1 - r_{0/1}$ and $\bar{C}$ computed factoring in nodes with $k < 2$, . In the middle, the light blue line with error bars is the actual value of $r_0$, while the dashed darker one is the approximation we derive in eq (3.84). At the bottom, the light red line with error bars is the actual value of $r_1$, while the dashed darker one is the approximation we derive in eq (3.88)*