Clustering spectrum of scale-free networks

Clara Stegehuis, Remco van der Hofstad, A.J.E.M. Janssen, and Johan S.H. van Leeuwaarden

Eindhoven University of Technology, Department of Mathematics and Computer Science,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands

(Dated: September 18, 2017)

Real-world networks often have power-law degrees and scale-free properties such as ultra-small distances and ultra-fast information spreading. In this paper, we study a third universal property: three-point correlations that suppress the creation of triangles and signal the presence of hierarchy. We quantify this property in terms of $\bar{c}(k)$, the probability that two neighbors of a degree-k node are neighbors themselves. We investigate how the clustering spectrum $k \mapsto \bar{c}(k)$ scales with k in the hidden variable model and show that $c(k)$ follows a *universal curve* that consists of three k-ranges where $\bar{c}(k)$ remains flat, starts declining, and eventually settles on a power law $\bar{c}(k) \sim k^{-\alpha}$ with α depending on the power law of the degree distribution. We test these results against ten contemporary real-world networks and explain analytically why the universal curve properties only reveal themselves in large networks.

I. INTRODUCTION

Most real-world networks have power-law degrees, so that the proportion of nodes having k neighbors scales as $k^{-\tau}$ with exponent τ between 2 and 3 [1–4]. Power-law degrees imply various intriguing scale-free network properties, such as ultra-small distances [5, 6] and the absence of percolation thresholds when $\tau < 3$ [7, 8]. Empirical evidence has been matched by random graph null models that are able to explain mathematically why and how these properties arise. This paper deals with another fundamental property observed in many scale-free networks related to three-point correlations that suppress the creation of triangles and signal the presence of hierarchy. We quantify this property in terms of the *clustering spectrum*, the function $k \mapsto \bar{c}(k)$ with $\bar{c}(k)$ the probability that two neighbors of a degree-k node are neighbors themselves.

In *uncorrelated* networks the clustering spectrum $\bar{c}(k)$ remains constant and independent of k. However, the majority of real-world networks have spectra that decay in k, as first observed in technological networks including the Internet [9, 10]. Figure 1 shows the same phenomenon for a social network: YouTube users as vertices, and edges indicating friendships between them [11].
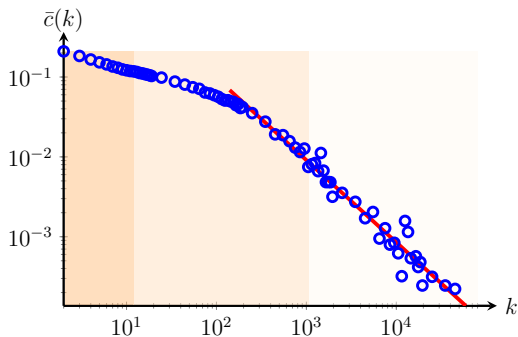

Figure 1. $\bar{c}(k)$ for the YouTube social network

Close inspection suggests the following properties, not only in Fig. 1, but also in the nine further networks in Fig. 2. The right end of the spectrum appears to be of the power-law form $k^{-\alpha}$; approximate values of α give rise to the dashed lines; (ii) The power law is only approximate and kicks in for rather large values of k. In fact, the slope of $\bar{c}(k)$ decreases with k; (iii) There exists a transition point: the minimal degree as of which the slope starts to decline faster and settles on its limiting (large k) value.

For scale-free networks a decaying $\bar{c}(k)$ is taken as an indicator for the presence of modularity and hierarchy [10], architectures that can be viewed as collections of subgraphs with dense connections within themselves and sparser ones between them. The existence of clusters of dense interaction signals hierarchical or nearly decomposable structures. When the function $\bar{c}(k)$ falls off with k, low-degree vertices have relatively high clustering coefficients, hence creating small modules that are connected through triangles. In contrast, high-degree vertices have very low clustering coefficients, and therefore act as bridges between the different local modules. This also explains why $\bar{c}(k)$ is not just a local property, and when viewed as a function of k, measures crucial mesoscopic network properties such as modularity, clusters and communities. The behavior of $\bar{c}(k)$ also turns out to be a good predictor for the macroscopic behavior of the network. Randomizing real-world networks while preserving the shape of the $\bar{c}(k)$ curve produces networks with very similar component sizes as well as similar hierarchical structures as the original network [16]. Furthermore, the shape of $\bar{c}(k)$ strongly influences the behavior of networks under percolation [17]. This places the $\bar{c}(k)$-curve among the most relevant indicators for structural correlations in network infrastructures.

In this paper, we obtain a precise characterization of clustering in the hidden variable model, a tractable random graph null model. We start from an explicit form of the $\bar{c}(k)$ curve for the hidden variable model [18–20]. We obtain a detailed description of the $\bar{c}(k)$-curve in the large-network limit that provides rigorous underpinning of the empirical observations (i)-(iii). We find that the decay rate in the hidden variable model is significantly
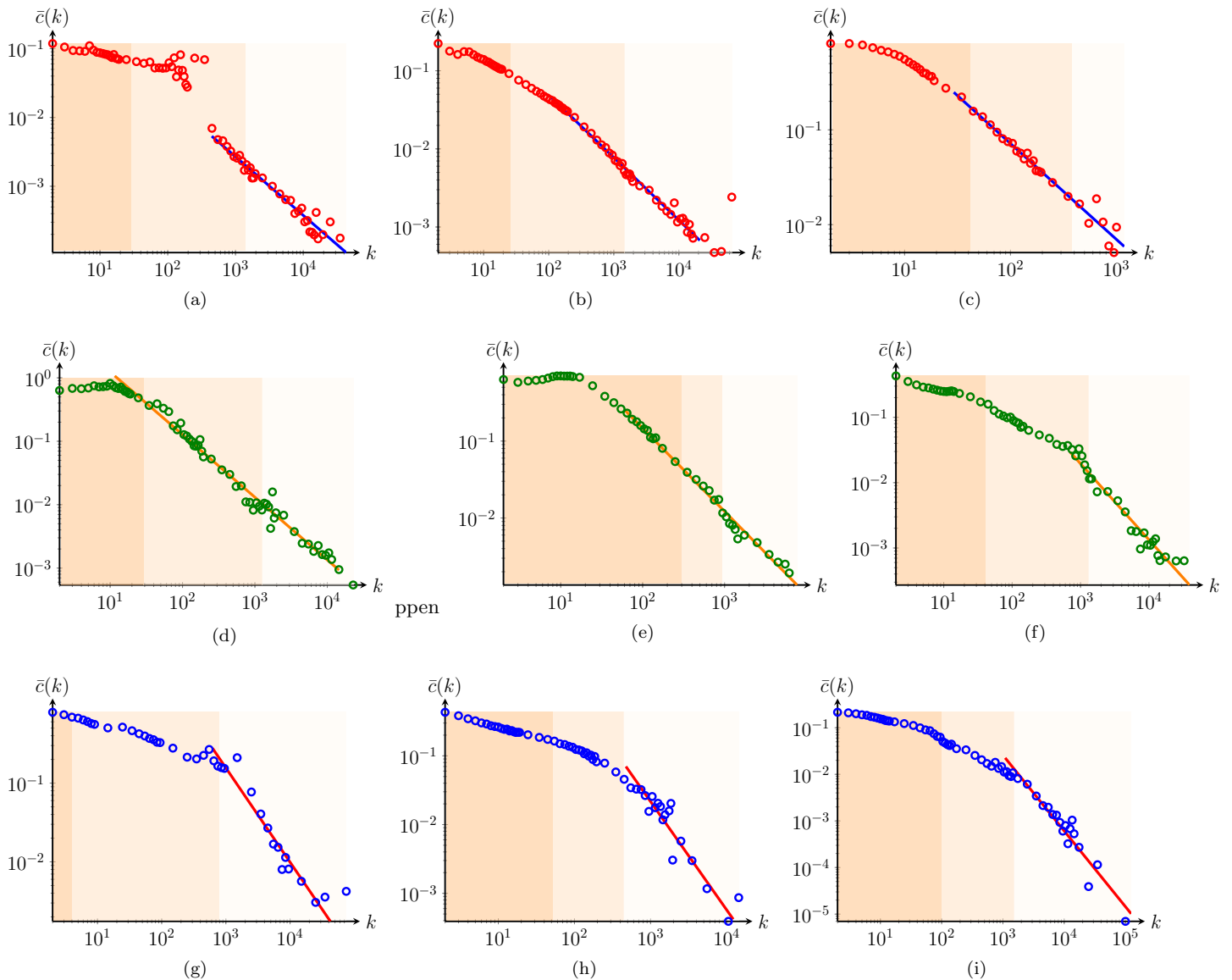
Figure 2. $\bar{c}(k)$ for several information (red), technological (green) and social (blue) real-world networks. (a) Hudong encyclopedia [12], (b) Baidu encyclopedia [12], (c) WordNet [13], (d) TREC-WT10g web graph [14], (e) Google web graph [11], (f) Internet on the Autonomous Systems level [11], (g) Catster/Dogster social networks [15], (h) Gowalla social network [11], (i) Wikipedia communication network [11]. The different shadings indicate the theoretical boundaries of the regimes as in Fig. 3, with N and τ as in Table I.

different from the exponent $\bar{c}(k) \sim k^{-1}$ that has been found in a hierarchical graph model [10] as well as in the preferential attachment model [21] and a preferential attachment model with enhanced clustering [22]. Furthermore, we show that before the power-law decay of $\bar{c}(k)$ kicks in, $\bar{c}(k)$ first has a constant regime for small k, and a logarithmic decay phase. This characterizes the entire clustering spectrum of the hidden variable model.

This paper is structured as follows. Section II introduces the random graph model and its local clustering coefficient. Section III presents the main results for the clustering spectrum. Section IV explains the shape of the

clustering spectrum in terms of an energy minimization argument, and Section V quantifies how fast the limiting clustering spectrum arises as function of the network size. We conclude with a discussion in Section VI and present all mathematical derivations of the main results in the appendix.

II. HIDDEN VARIABLES

As null model we employ the hidden variable model [18, 23–26]. Given N nodes, hidden variable models

are defined as follows. Associate to each node a hidden variable h drawn from a given probability distribution function

$$\rho(h) = Ch^{-\tau} \qquad (1)$$

for some constant C. Next join each pair of vertices independently according to a given probability $p(h, h')$ with h and h' the hidden variables associated to the two nodes. Many networks can be embedded in this hidden-variable framework, but particular attention goes to the case in which the hidden variables have themselves the structure of the degrees of a real-world network. In that case the hidden-variable model puts soft constraints on the degrees, which is typically easier to analyze than hard constraints as in the configuration model [4, 27–29]. Chung and Lu [30] introduced the hidden variable model in the form

$$p(h, h') \sim \frac{hh'}{N\langle h \rangle}, \qquad (2)$$

so that the expected degree of a node equals its hidden variable.

We now discuss the structural and natural cutoff, because both will play a crucial role in the description of the clustering spectrum. The structural cutoff is defined as the largest possible upper bound on the degrees required to guarantee single edges, while the natural cutoff characterizes the maximal degree in a sample of N vertices. For scale-free networks with exponent $\tau \in (2,3]$ the structural cutoff scales as $\sqrt{N}$ while the natural cutoff scales as $N^{1/(\tau-1)}$, which gives rise to structural negative correlations and possibly other finite-size effects. If one wants to avoid such effects, then the maximal value of the product hh' should never exceed $N\langle h \rangle$, which can be guaranteed by the assumption that the hidden degree h is smaller than the structural cutoff $h_s = \sqrt{N\langle h \rangle}$. While this restricts $p(h, h')$ in (2) within the interval $[0, 1]$, banning degrees larger than the structural cutoff strongly violates the reality of scale-free networks, where degrees all the way up to the natural cutoff $(N\langle h \rangle)^{1/(\tau-1)}$ need to be considered. We therefore work with (although many asymptotically equivalent choices are possible; see [31] and Appendix A)

$$p(h, h') = \min\left(1, \frac{hh'}{N\langle h \rangle}\right), \qquad (3)$$

putting no further restrictions on the range of the hidden variables (and hence degrees).

In this paper, we shall work with $c(h)$, the local clustering coefficient of a randomly chosen vertex with hidden variable h. However, when studying local clustering in real-world data sets, we can only observe $\bar{c}(k)$, the local clustering coefficient of a vertex of degree k. In Appendix C we show that the approximation $\bar{c}(h) \approx c(h)$ is highly accurate. We start from the explicit expression for $c(h)$ [18], which measures the probability that two randomly
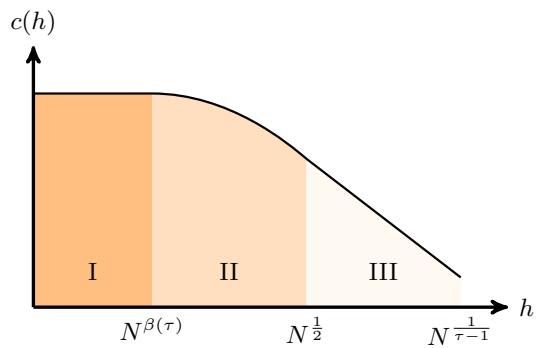

Figure 3. Clustering spectrum $h \mapsto c(h)$ with three different ranges for h: the flat range, logarithmic decay, and the power-law decay.

chosen edges from h are neighbors, i.e.,

$$c(h) = \int_{h'} \int_{h''} p(h'|h)p(h', h'')p(h''|h)\mathrm{d}h''\mathrm{d}h', \qquad (4)$$

with $p(h'|h)$ the conditional probability that a randomly chosen edge from an h-vertex is connected to an h'-vertex and $p(h, h')$ as in (3). The goal is now to characterize the $c(h)$-curve (and hence the $\bar{c}(k)$-curve).

III. UNIVERSAL CLUSTERING SPECTRUM

The asymptotic evaluation of the double integral (4) in the large-N regime reveals three different ranges, defined in terms of the scaling relation between the hidden variable h and the network size N. The three ranges together span the entire clustering spectrum as shown in Fig. 3. The detailed calculations are deferred to Appendix A.

The first range pertains to the smallest-degree nodes, i.e., vertices with a hidden variable that does not exceed $N^{\beta(\tau)}$ with $\beta(\tau) = \frac{\tau-2}{\tau-1}$. In this case we show that

$$c(h) \propto N^{2-\tau} \ln N, \quad h \leq N^{\beta(\tau)}. \qquad (5)$$

In particular, here the local clustering does not depend on the degree and in fact corresponds with the large-N behavior of the global clustering coefficient [31, 32]. Note that the interval $[0, \beta(\tau)]$ diminishes when τ is close to 2, a possible explanation for why the flat range associated with Range I is hard to recognize in some of the real-world data sets.

Range II considers nodes with hidden variables (degrees) above the threshold $N^{\beta(\tau)}$, but below the structural cutoff $\sqrt{N}$. These nodes start experiencing structural correlations, and close inspection of the integral (4) yields

$$c(h) \propto N^{2-\tau}\left(1 + \ln\left(\frac{\sqrt{N}}{h}\right)\right), \quad N^{\beta(\tau)} \leq h \leq \sqrt{N}. \quad (6)$$

This range shows relatively slow, logarithmic decay in the clustering spectrum, and is clearly visible in the ten data sets.
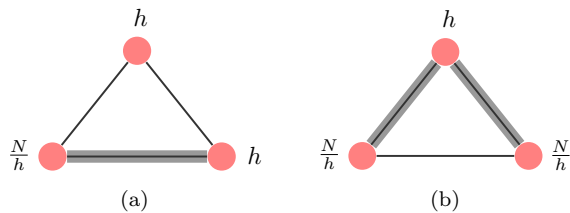
Figure 4. Orders of magnitude of the major contributions in the different h-ranges. The highlighted edges are present with asymptotically positive probability. (a) $h < \sqrt{N}$ (b) $h > \sqrt{N}$.

Range III considers hidden variables above the structural cutoff, when the restrictive effect of degree-degree correlations becomes more evident. In this range we find that

$$c(h) \propto \frac{1}{N} \left(\frac{h}{N} \right)^{-2(3-\tau)}, \quad h \geq \sqrt{N}, \qquad (7)$$

hence power-law decay with a power-law exponent $\alpha = 2(3 - \tau)$. Such power-law decay has been observed in many real-world networks [4, 10, 33–36], where most networks were found to have the power-law exponent close to one. The asymptotic relation (7) shows that the exponent α decreases with τ and takes values in the entire range $(0, 2)$. Table I contains estimated values of α for the ten data sets.

IV. ENERGY MINIMIZATION

We now explain why the clustering spectrum splits into three ranges, using an argument that minimizes the energy needed to create triangles among nodes with specific hidden variables.

In all three ranges for h, there is one type of 'most likely' triangle, as shown in Fig. 4. This means that most triangles containing a vertex v with hidden variable h are triangles with two other vertices v' and v'' with hidden variables h' and h'' of specific sizes, depending on h. The probability that a triangle is present between v, v' and v'' can be written as

$$\min \left(1, \frac{hh'}{N\langle h \rangle} \right) \min \left(1, \frac{hh''}{N\langle h \rangle} \right) \min \left(1, \frac{h'h''}{N\langle h \rangle} \right). \quad (8)$$

While the probability that such a triangle exists among the three nodes thus increases with h' and h'', the number of such nodes decreases with h' and h'' because vertices with higher h-values are rarer. Therefore, the maximum contribution to $c(h)$ results from a trade-off between large enough h', h'' for likeliness of occurrence of the triangle, and h', h'' small enough to have enough copies. Thus, having $h' > N\langle h \rangle/h$ is not optimal, since then the probability that an edge exists between v and v' no longer increases with h'. This results in the bound

$$h', h'' \leq \frac{N\langle h \rangle}{h}. \qquad (9)$$

Similarly, $h'h'' > N\langle h \rangle$ is also suboptimal, since then further increasing h' and h'' does not increase the probability of an edge between v' and v''. This gives as a second bound

$$h'h'' \leq N\langle h \rangle. \qquad (10)$$

In Ranges I and II, $h < \sqrt{N\langle h \rangle}$, so that $N\langle h \rangle/h > \sqrt{N\langle h \rangle}$. In this situation we reach bound (10) before we reach bound (9). Therefore, the maximum contribution to $c(h)$ comes from $h'h'' \approx N$, where also $h', h'' < N\langle h \rangle/h$ because of the bound (9). Here the probability that the edge between v' and v'' exists is large, while the other two edges have a small probability to be present, as shown in Fig. 4a. Note that for h in Range I, the bound (9) is superfluous, since in this regime $N\langle h \rangle/h > h_c$, while the network does not contain vertices with hidden variables larger than h_c. This bound indicates the minimal values of h' such that an h-vertex is guaranteed to be connected to an h'-vertex. Thus, vertices in Range I are not even guaranteed to have connections to the highest degree vertices, hence they are not affected by the single-edge constraints. Therefore the value of $c(h)$ in Range I is independent of h.

In Range III, $h > \sqrt{N\langle h \rangle}$, so that $N\langle h \rangle/h < \sqrt{N\langle h \rangle}$. Therefore, we reach bound (9) before we reach bound (10). Thus, we maximize the contribution to the number of triangles by choosing $h', h'' \approx N\langle h \rangle/h$. Then the probability that the edge from v to v' and from v to v'' is present is large, while the probability that the edge between v' and v'' exists is small, as illustrated in Fig. 4b.

V. CONVERGENCE RATE

We next ask how large networks should be, or become, before they reveal the features of the universal clustering spectrum. In other words, while the results in this paper are shown for the large-N limit, for what finite N-values can we expect to see the different ranges and clustering decay? To bring networks of different sizes N on a comparable footing, we consider

$$\sigma_N(t) = \frac{\ln \left(c(h)/c(h_c) \right)}{\ln(N\langle h \rangle)}, \quad h = (N\langle h \rangle)^t, \qquad (11)$$

for $0 \leq t \leq \frac{1}{\tau - 1}$. The slope of $\sigma_N(t)$ can be interpreted as a measure of the decay of $c(h)$ at $h = (N\langle h \rangle)^t$, and all curves share the same right end of the spectrum; see Appendix B for more details. Figure 5 shows this rescaled clustering spectrum for synthetic networks generated with the hidden variable model with $\tau = 2.25$. Already 10^4 vertices reveal the essential features of the spectrum: the decay and the three ranges. Increasing the network size further to 10^5 and 10^6 nodes shows that the spectrum settles on the limiting curve. Here we note that the real-world networks reported in Figs. 1 and 2 are also of order 10^5-10^6 nodes, see Table I.
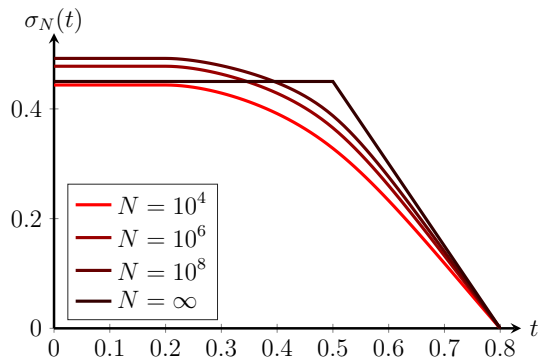
Figure 5. $\sigma_N(t)$ for $N = 10^4, 10^6$ and 10^8 together with the limiting function, using $\tau = 2.25$, for which $\frac{1}{\tau-1} = 0.8$.

	N	τ	g.o.f.	α
Hudong	1.984.484	2,30	0.00	0,85
Baidu	2.141.300	2,29	0.00	0,80
Wordnet	146.005	2,47	0.00	1,01
Google web	875.713	2,73	0.00	1,03
AS-Skitter	1.696.415	2,35	0.06	1,12
TREC-WT10g	1.601.787	2,23	0.00	0,99
Wiki-talk	2.394.385	2,46	0.00	1,54
Catster/Dogster	623.766	2,13	0.00	1,20
Gowalla	196.591	2,65	0.80	1,24
Youtube	1.134.890	2,22	0.00	1,05

Table I. Data sets. N denotes the number of vertices, τ the exponent of the degree distribution estimated by the method proposed in [27] together with the goodness of fit criterion proposed in [27] (when the goodness of fit is at least 0.10, a power-law distribution cannot be rejected), and α denotes the exponent of $c(k)$.

Figure 5 also brings to bear a potential pitfall when the goal is to obtain statistically accurate estimates for the slope of $c(h)$. Observe the extremely slow convergence to the limiting curve for $N = \infty$; a well documented property of certain clustering measures [31, 32, 37, 38]. In Appendix B we again use the integral expression (4) to characterize the limiting curve for $N = \infty$ and the rate of convergence as function of N, and indeed extreme N-values are required for statistically reliable slope estimates for e.g. t-values of $\frac{1}{2}$ and $\frac{1}{\tau-1}$; this is also apparent from visual inspection of Fig. 5. Therefore, the estimates in Table I only serve as indicative values of α. Finally, observe that Range II disappears in the limiting curve, due to the rescaling in (11), but again only for extreme N-values. Because this paper is about structure rather than statistical estimation, the slow convergence in fact provides additional support for the persistence of Range II in Figs. 1 and 2.

Table I also shows that the relation $\alpha = -2(3 - \tau)$ is inaccurate for the real-world data sets, in turn affecting the theoretical boundaries of the three regimes indicated in Fig. 2. One explanation for this inaccuracy is that the real-world networks might not follow pure power-law distributions, as measured by the goodness of fit criterion in Table I. Furthermore, real-world networks are usually highly clustered and contain community structures, whereas the hidden variable model is locally tree-like. These modular structure may explain, for example, why the power-law decay of the hidden variable model is less pronounced in the three social networks of Fig. 2. It is remarkable that despite these differences between hidden variable models and real-world networks, the global shape of the $c(k)$ curve of the hidden variable model is still visible in these heavy-tailed real-world networks.

VI. DISCUSSION

The hidden variable model gives rise to *single-edge* networks in which pairs of vertices can only be connected once. Hierarchical modularity and the decaying clustering spectrum have been contributed to this restriction that no two vertices have more than one edge connecting them [9, 39–42]. The physical intuition is that the single-edge constraint leads to far fewer connections between high-degree vertices than anticipated based on randomly assigned edges. We have indeed confirmed this intuition, not only through analytically revealing the universal clustering curve, but also by providing an alternative derivation of the three ranges based on energy minimization and structural correlations.

We now show that the clustering spectrum revealed using the hidden variable model, also appears for a second widely studied null model. This second model cannot be the Configuration Model (CM), which preserves the degree distribution by making connections between vertices in the most random way possible [6, 43]. Indeed, because of the random edge assignment, the CM has no degree correlations, leading in the case of scale-free networks with diverging second moment to uncorrelated networks with non-negligible fractions of self-loops (a vertex joined to itself) and multiple connections (two vertices connected by more than one edge). This picture changes dramatically when self-loops and multiple edges are avoided, a restriction mostly felt by the high-degree nodes, who can no longer establish multiple edges among each other.

We therefore consider the Erased Configuration Model (ECM) that takes a sample from the CM and then erases all the self-loops and multiple edges. While this removes some of the edges in the graph, thus violating the hard constraint, only a small proportion of the edges is removed, so that the degree of vertex j in ECM is still close to D_j [44, Chapter 7]. In the ECM, the probability that a vertex with degree D_i is connected to a vertex with degree D_j can be approximated by $1 - e^{-D_i D_j/\langle D\rangle N}$ [45, Eq.(4.9)]. Therefore, we expect the ECM and the hidden variable model to have similar properties (see e.g. [31]) when we choose

$$p(h, h') = 1 - e^{-hh'/N\langle h\rangle} \approx \frac{hh'}{N\langle h\rangle}. \tag{12}$$
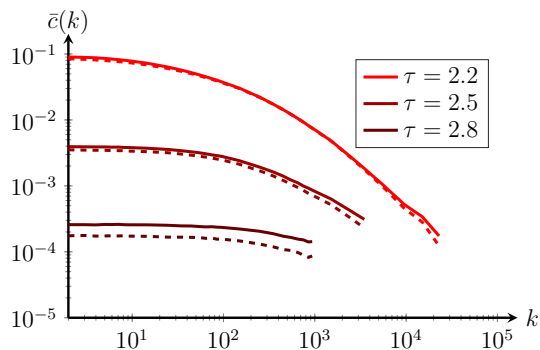
Figure 6. $\bar{c}(k)$ for a hidden variable model with connection probabilities (12) (solid line) and an erased configuration model (dashed line). The presented values of $\bar{c}(k)$ are averages over 10^4 realizations of networks of size $N = 10^5$.

Figure 6 illustrates how both null models generate highly similar spectra, which provides additional support for the claim that the clustering spectrum is a universal property of simple scale-free networks. The ECM is more difficult to deal with compared to hidden variable models, since edges in ECM are not independent. In particular, we expect that these dependencies vanish for the $k \mapsto \bar{c}(k)$ curve. Establishing the universality of the $k \mapsto \bar{c}(k)$ curve for other random graph null models such as ECM, networks with an underlying geometric space [46] or hierarchical configuration models [47] is a major research direction. The ECM and the hidden variable model are both null models with soft constraints on the degrees. Putting hard constraints on the degrees with the CM, has the nice property that simple graphs generated using this null model are uniform samples of all simple graphs with the same degree sequence. Dealing with such uniform samples is notoriously hard when the second moment

of the degrees is diverging, for example since the CM will yield many edges between high-degree vertices. This makes sampling uniform graphs difficult [48–50]. Thus, the joint requirement of hard degree and single-edge constraints, as in the CM, presents formidable technical challenges. Whether our results for the $k \mapsto \bar{c}(k)$ curve for soft-constraint models also carry over to these uniform simple graphs is a challenging open problem.

In this paper we have investigated the presence of triangles in the hidden variable model. We have shown that by first conditioning on the node degree, there arises a unique 'most likely' triangle with two other vertices of specific degrees. We have not only explained this insight heuristically, but it is also reflected in the elaborate analysis of the double integral for $c(h)$ in Appendix A. As such, we have introduced an intuitive and tractable mathematical method for asymptotic triangle counting. It is likely that the method carries over to counting other motifs, such as squares, or complete graphs of larger sizes. For any given motif, and first conditioning on the node degree, we again expect to find specific configuration that are most likely. Further mathematical challenges need to be overcome, though, because we expect that the 'most likely' configurations critically depend on the precise motif topologies and the associated energy minimization problems.

ACKNOWLEDGMENTS

[1] R. Albert, H. Jeong, and A.-L. Barabási, Nature **401**, 130 (1999).

[2] M. Faloutsos, P. Faloutsos, and C. Faloutsos, in *ACM SIGCOMM Computer Communication Review*, Vol. 29 (ACM, 1999) pp. 251–262.

[3] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, Nature **407**, 651 (2000).

[4] A. Vázquez, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. E **65**, 066130 (2002).

[5] R. van der Hofstad, G. Hooghiemstra, and D. Znamenski, Electron. J. Probab. **12**, 703 (2007).

[6] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Phys. Rev. E **64**, 026118 (2001).

[7] S. Janson, Electron. J. Probab. **14**, 86 (2009).

[8] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).

[9] R. Pastor-Satorras, A. Vázquez, and A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001).

[10] E. Ravasz and A.-L. Barabási, Phys. Rev. E **67**, 026112 (2003).

[11] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford large network dataset collection," `http://snap.stanford.edu/data` (2014), date of access: 14/03/2017.

[12] X. Niu, X. Sun, H. Wang, S. Rong, G. Qi, and Y. Yu, in *The Semantic Web – ISWC 2011* (Springer Nature, 2011) pp. 205–220.

[13] G. Miller and C. Fellbaum, "Wordnet: An electronic lexical database," (1998).

[14] P. Bailey, N. Craswell, and D. Hawking, Information Processing & Management **39**, 853 (2003).

[15] J. Kunegis, in *Proceedings of the 22Nd International Conference on World Wide Web*, WWW '13 Companion (ACM, New York, NY, USA, 2013) pp. 1343–1350.

[16] P. Colomer-de Simón, M. A. Serrano, M. G. Beiró, J. I. Alvarez-Hamelin, and M. Boguñá, Sci. Rep. **3**, 2517 (2013).

[17] M. A. Serrano and M. Boguñá, Phys. Rev. Lett. **97**, 088701 (2006).

[18] M. Boguñá and R. Pastor-Satorras, Phys. Rev. E **68**, 036112 (2003).

[19] M. A. Serrano, M. Boguñá, R. Pastor-Satorras, and A. Vespignani, Large scale structure and dynamics of complex networks: From information technology to finance and natural sciences , 35 (2007).

[20] S. N. Dorogovtsev, Phys. Rev. E **69**, 027104 (2004).

[21] A. Krot and L. Ostroumova Prokhorenkova, "Local clustering coefficient in generalized preferential attachment models," in *Algorithms and Models for the Web Graph: 12th International Workshop, WAW 2015, Eindhoven, The Netherlands, December 10-11, 2015, Proceedings*, edited by D. F. Gleich, J. Komjáthy, and N. Litvak (Springer International Publishing, Cham, 2015) pp. 15–28.

[22] G. Szabó, M. Alava, and J. Kertész, Phys. Rev. E **67**, 056102 (2003).

[23] J. Park and M. E. J. Newman, Phys. Rev. E **70**, 066117 (2004).

[24] B. Bollobás, S. Janson, and O. Riordan, Random Structures & Algorithms **31**, 3 (2007).

[25] T. Britton, M. Deijfen, and A. Martin-Löf, J. Stat. Phys. **124**, 1377 (2006).

[26] I. Norros and H. Reittu, Adv. Appl. Probab. **38**, 59 (2006).

[27] A. Clauset, C. R. Shalizi, and M. E. J. Newman, SIAM Rev. **51**, 661 (2009).

[28] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[29] S. Dhara, R. van der Hofstad, and J. S. H. van Leeuwaarden, arXiv:1605.02868 (2016).

[30] F. Chung and L. Lu, Proc. Natl. Acad. Sci. USA **99**, 15879 (2002).

[31] R. van der Hofstad, A. J. E. M. Janssen, J. S. H. van Leeuwaarden, and C. Stegehuis, Phys. Rev. E **95**, 022307 (2017).

[32] P. Colomer-de Simon and M. Boguñá, Phys. Rev. E **86**, 026120 (2012).

[33] M. A. Serrano and M. Boguñá, Phys. Rev. E **74**, 056114 (2006).

[34] M. Catanzaro, G. Caldarelli, and L. Pietronero, Phys. Rev. E **70**, 037101 (2004).

[35] J. Leskovec, *Dynamics of large networks* (ProQuest, 2008).

[36] D. Krioukov, M. Kitsak, R. S. Sinkovits, D. Rideout, D. Meyer, and M. Boguná, Sci. Rep. **2**, 793 (2012).

[37] M. Boguñá, C. Castellano, and R. Pastor-Satorras, Phys. Rev. E **79**, 036110 (2009).

[38] A. J. E. M. Janssen and J. S. H. van Leeuwaarden, EPL (Europhysics Letters) **112**, 68001 (2015).

[39] S. Maslov, K. Sneppen, and A. Zaliznyak, Phys. A **333**, 529 (2004).

[40] J. Park and M. E. J. Newman, Phys. Rev. E **68**, 026112 (2003).

[41] M. E. J. Newman, Phys. Rev. Lett. **89**, 208701 (2002).

[42] M. E. J. Newman, Phys. Rev. E **67**, 026126 (2003).

[43] B. Bollobás, European J. Combin. **1**, 311 (1980).

[44] R. van der Hofstad, *Random Graphs and Complex Networks Vol. 1* (Cambridge University Press, 2017).

[45] R. van der Hofstad, G. Hooghiemstra, and P. Van Mieghem, Random Structures & Algorithms **27**, 76 (2005).

[46] M. Á. Serrano, D. Krioukov, and M. Boguñá, Phys. Rev. Lett. **100**, 078701 (2008).

[47] C. Stegehuis, R. van der Hofstad, and J. S. H. van Leeuwaarden, Phys. Rev. E **94**, 012302 (2016).

[48] R. Milo, N. Kashtan, S. Itzkovitz, M. E. J. Newman, and U. Alon, arXiv preprint cond-mat/0312028 (2003).

[49] F. Viger and M. Latapy, in *Lecture Notes in Computer Science* (Springer Berlin Heidelberg, 2005) pp. 440–449.

[50] C. I. D. Genio, H. Kim, Z. Toroczkai, and K. E. Bassler, PLoS ONE **5**, e10012 (2010).

Appendix A: Derivation for the three ranges

In this appendix, we compute $c(h)$ in (4), and we show that $c(h)$ can be approximated by (5), (6), or (7), depending on the value of h. Throughout the appendix, we assume that $p(h,h') = \min(1, hh'/h_s^2)$ and $\rho(h) = Ch^{-\tau}$. Then, the derivation of $c(h)$ in [16] yields

$$
c(h) = \frac{\int_1^{h_c} \int_1^{h_c} \rho(h')p(h,h')\rho(h'')p(h,h'')p(h',h'')\mathrm{d}h''\mathrm{d}h'}{\left[\int_1^{h_c} \rho(h')p(h,h')\mathrm{d}h'\right]^2}
$$
$$
= \frac{\int_1^{h_c} \int_1^{h_c} (h'h'')^{-\tau} \min(\frac{hh'}{h_s^2},1)\min(\frac{hh''}{h_s^2},1)\min(\frac{h'h''}{h_s^2},1)\mathrm{d}h''\mathrm{d}h'}{\left[\int_1^{h_c}(h')^{-\tau}\min(\frac{hh'}{h_s^2},1)\mathrm{d}h'\right]^2}.
$$
(A1)

Computing $c(h)$ will also allow us to compute

$$
\sigma_N(t) = \frac{\ln(c(h)/c(h_{\mathrm{ref}}))}{\ln(N\langle h\rangle)}, \quad h = (N\langle h\rangle)^t, \qquad \text{(A2)}
$$

for $0 \le t \le \frac{1}{\tau-1}$, where $h_{\mathrm{ref}} \in [0,h_c]$ is fixed. We are interested in computing the value of $\sigma_N(t)$ for large values of N.

Adopting the standard choices [31]

$$
h_s = \sqrt{N\langle h\rangle}, \quad h_c = (N\langle h\rangle)^{1/(\tau-1)}, \qquad \text{(A3)}
$$

and setting $h_{\min} = 1$ gives

$$
\langle h\rangle = \frac{\tau-1}{\tau-2}\frac{1-N^{2-\tau}}{1-N^{1-\tau}}. \qquad \text{(A4)}
$$

For ease of notation in the proofs below, we will use

$$
a = h_s^{-1} = (N\langle h\rangle)^{-1/2}, \quad b = \frac{h_c}{h_s} = (N\langle h\rangle)^{\frac{3-\tau}{2(\tau-1)}}, \quad \text{(A5)}
$$

and

$$
r(u) = \min(u,1). \qquad \text{(A6)}
$$

In this notation, (A1) can be succinctly written as

$$
c(h) = \frac{\int_a^b \int_a^b (xy)^{-\tau} r(ahx)r(ahy)r(xy)\mathrm{d}x\mathrm{d}y}{\left[\int_a^b x^{-\tau}r(ahx)\mathrm{d}x\right]^2}. \qquad \text{(A7)}
$$

Because of the four min operators in the expression (A1), we have to consider various h-ranges. We compute the value of $c(h)$ in these three ranges one by one.

Range I: $h < h_s^2/h_c$. We now show that in this range

$$c(h) \approx \frac{\tau - 2}{3 - \tau} h_s^{4-2\tau} \ln\left(\frac{h_c^2}{h_s^2}\right) \propto N^{2-\tau} \ln N, \qquad (A8)$$

which proves (5).

This range corresponds to $h < 1/(ab)$ with a and b as in (A5). In this range, $r(ahx) = ahx$ and $r(ahy) = ahy$ for all $x \in [a, b]$. This yields for $c(h)$

$$c(h) = \frac{\int_a^b \int_a^b (xy)^{1-\tau} r(xy) \mathrm{d}x\mathrm{d}y}{\left[\int_a^b x^{1-\tau}\mathrm{d}x\right]^2}. \qquad (A9)$$

For the denominator we compute

$$\int_a^b x^{1-\tau}\mathrm{d}x = \frac{a^{2-\tau} - b^{2-\tau}}{\tau - 2}. \qquad (A10)$$

Since $a \ll b$, this can be approximated as

$$\frac{a^{2-\tau} - b^{2-\tau}}{\tau - 2} \approx \frac{a^{2-\tau}}{\tau - 2}. \qquad (A11)$$

We can compute the numerator of (A9) as

$$
\int_a^b \int_a^b (xy)^{1-\tau} r(xy)\mathrm{d}x\mathrm{d}y
$$
$$
= \int_a^{1/b} \int_a^b (xy)^{2-\tau}\mathrm{d}x\mathrm{d}y + \int_{1/b}^b \int_a^{1/x} (xy)^{2-\tau}\mathrm{d}x\mathrm{d}y
$$
$$
+ \int_{1/b}^b \int_{1/x}^b (xy)^{1-\tau}\mathrm{d}x\mathrm{d}y
$$
$$
= \frac{\left(b^{\tau-3} - a^{3-\tau}\right)\left(b^{3-\tau} - a^{3-\tau}\right)}{(3-\tau)^2}
$$
$$
+ \frac{1}{3-\tau}\left(\ln\left(b^2\right) - \frac{a^{3-\tau}\left(b^{3-\tau} - b^{\tau-3}\right)}{3-\tau}\right)
$$
$$
+ \frac{1}{2-\tau}\left(\frac{b^{2-\tau}(b^{2-\tau} - b^{\tau-2})}{2-\tau} - \ln\left(b^2\right)\right)
$$
$$
= \frac{\ln\left(b^2\right)}{(3-\tau)(\tau-2)} - \frac{1 - b^{4-2\tau}}{(\tau-2)^2} + \frac{1 - 2(ab)^{3-\tau} + a^{6-2\tau}}{(3-\tau)^2}. \qquad (A12)
$$

The first of these three terms dominates when

$$\frac{3-\tau}{\tau-1} \frac{\ln(N\langle h\rangle)}{(3-\tau)(\tau-2)} \gg \frac{1}{(\tau-2)^2} \qquad (A13)$$

and

$$\frac{3-\tau}{\tau-1} \frac{\ln(N\langle h\rangle)}{(3-\tau)(\tau-2)} \gg \frac{1}{(3-\tau)^2}, \qquad (A14)$$

where we have used that $b^2 = (N\langle h\rangle)^{(3-\tau)/(\tau-1)}$. Thus, when $\ln(N\langle h\rangle)$ is large compared to $(\tau-1)/(\tau-2)$ and $(\tau-1)(\tau-2)/(\tau-3)^2$, we obtain

$$c(h) \approx \frac{\tau-2}{3-\tau} a^{2\tau-4} \ln\left(b^2\right) \propto N^{2-\tau} \ln(N), \qquad (A15)$$

which proves (A8).

Range II: $h_s^2/h_c < h < h_s$ In this range, we show that

$$c(h) \approx h_s^{4-2\tau} \frac{\ln\left(\frac{h_s^2}{h^2}\right) + M}{(\tau-2)(3-\tau)} \propto N^{2-\tau}\left(\ln\left(N/h^2\right) + M\right), \qquad (A16)$$

for some positive constant M, which proves (6).

This range corresponds to $(ab)^{-1} < h < a^{-1}$. For these values of h, we have $ahx, ahy = 1$ for $x, y = (ah)^{-1} \in (1, b)$ and $xy = 1$ for $y = 1/x \in [a, b]$ when $b^{-1} < x < b$. Then for the denominator of (A7) we compute

$$
\int_a^{1/(ah)} ahx^{1-\tau}\mathrm{d}x + \int_{1/(ah)}^b x^{-\tau}\mathrm{d}x
$$
$$
= \frac{1}{\tau-2}\left(a^{3-\tau}h - (ah)^{\tau-1}\right)
$$
$$
+ \frac{1}{\tau-1}\left((ah)^{\tau-1} - b^{1-\tau}\right)
$$
$$
= ah\left(\frac{a^{2-\tau}}{\tau-2} - \frac{(ah)^{\tau-2}}{(\tau-1)(\tau-2)} - \frac{b^{1-\tau}/(ah)}{\tau-1}\right). \qquad (A17)
$$

Splitting up the integral in the numerator results in

$$
\mathrm{Num}(h) = \int_a^b \int_a^b (xy)^{-\tau} r(ahx)r(ahy)r(xy)\mathrm{d}x\mathrm{d}y
$$
$$
= \int_{1/(ah)}^b \int_{1/(ah)}^b (xy)^{-\tau}\mathrm{d}y\mathrm{d}x
$$
$$
+ 2ah \int_{1/(ah)}^b \int_{1/x}^{1/(ah)} (xy)^{-\tau} y\mathrm{d}y\mathrm{d}x
$$
$$
+ 2ah \int_{1/(ah)}^b \int_a^{1/x} (xy)^{1-\tau} y\mathrm{d}y\mathrm{d}x
$$
$$
+ a^2h^2 \int_{ah}^{1/(ah)} \int_a^{1/x} (xy)^{2-\tau}\mathrm{d}y\mathrm{d}x
$$
$$
+ a^2h^2 \int_{ah}^{1/(ah)} \int_{1/x}^{1/(ah)} (xy)^{1-\tau}\mathrm{d}y\mathrm{d}x
$$
$$
+ a^2h^2 \int_a^{ah} \int_a^{1/(ah)} (xy)^{2-\tau}\mathrm{d}y\mathrm{d}x
$$
$$
=: I_1 + I_2 + I_3 + I_4 + I_5 + I_6, \qquad (A18)
$$

where the factors 2 arise by symmetry of the integrand in x and y. Computing these integrals yields

$$I_1 = a^2h^2 \left(\frac{(ah)^{\tau-2} - a^{-1}b^{1-\tau}h^{-1}}{\tau-1}\right)^2, \qquad (A19)$$

$$I_2 = 2a^2h^2\left(\frac{1 - 1/(abh)}{\tau-2} - \frac{(ah)^{2\tau-4}}{(\tau-1)(\tau-2)}\right. \qquad (A20)$$

$$\left. \times \left(1 - (abh)^{1-\tau}\right)\right), \qquad (A21)$$

$$I_3 = 2a^2h^2\left(\frac{1 - 1/(abh)}{3-\tau} - \frac{h^{\tau-3}\left(1 - (abh)^{2-\tau}\right)}{(3-\tau)(\tau-2)}\right), \qquad (A22)$$

$$I_4 = a^2h^2\left(\frac{\ln((ah)^{-2})}{3-\tau} + \frac{(a^2h)^{3-\tau} - h^{\tau-3}}{(3-\tau)^2}\right), \qquad (A23)$$

$$I_5 = a^2h^2\left(\frac{\ln((ah)^{-2})}{\tau-2} - \frac{1-(ah)^{2\tau-4}}{(\tau-2)^2}\right), \qquad \text{(A24)}$$

$$I_6 = a^2h^2\left(\frac{1-h^{\tau-3}+a^{6-2\tau}-(a^2h)^{3-\tau}}{(3-\tau)^2}\right). \qquad \text{(A25)}$$

We have $ah < 1 < ahb$ and so the leading behavior of Num(h) is determined by the terms involving $\ln((ah)^{-2})$ in I_3 and I_4, all other terms being bounded. Retaining only these dominant terms, we get

$$\text{Num}(h) = a^2h^2\frac{\ln((ah)^{-2})}{(\tau-2)(3-\tau)}(1+o(1)), \qquad \text{(A26)}$$

provided that $ah \to 0$ as $N \to \infty$. In terms of the variable t in $h = (N\langle h\rangle)^t$, see (11) and (A2), this condition holds when we restrict to $t \in [(\tau-2)/(\tau-1), \frac{1}{2}-\varepsilon]$ for any $\varepsilon > 0$. Furthermore, from (A17),

$$\left(\int_a^b x^{-\tau}r(ahx)\mathrm{d}x\right)^2 = a^2h^2\left(\frac{a^{2-\tau}}{\tau-2}\right)^2(1+o(1)). \qquad \text{(A27)}$$

Hence, when $ah \to 0$, we have

$$c(h) = \frac{\tau-2}{3-\tau}a^{2\tau-4}\ln\left((ah)^{-2}\right)(1+o(1)) \propto N^{2-\tau}\ln\left(N/h^2\right). \qquad \text{(A28)}$$

We compute $c(h = 1/a)$ asymptotically by retaining only all constant terms between brackets in (A19)-(A25) since all other terms vanish or tend to 0 as $N \to \infty$. This gives

$$\text{Num}(h = 1/a) = a^2h^2\Big(\frac{1}{(\tau-1)^2} + \frac{2}{\tau-2}$$
$$- \frac{2}{(\tau-1)(\tau-2)} + \frac{2}{3-\tau} + \frac{1}{(3-\tau)^2}\Big)(1+o(1))$$
$$= Pa^2h^2(1+o(1)), \qquad \text{(A29)}$$

where $P = \frac{1}{(\tau-1)^2} + \frac{1}{(3-\tau)^2} + \frac{2}{\tau-1} + \frac{2}{3-\tau}$. Together with (A27), we find

$$c(h = 1/a) = P(\tau-2)^2a^{2\tau-4}(1+o(1)) \propto N^{2-\tau}. \qquad \text{(A30)}$$

In [31], it has been shown that $c(h)$ decreases in h, and then (A16) follows from (A28) and (A30).

Range III: $h_s < h < h_c$. We now show that when $h_s < h < h_c$, then

$$c(h) \approx \frac{1}{(3-\tau)^2}(h_s/h)^{6-2\tau}h_s^{4-2\tau} \propto N^{5-2\tau}h^{2\tau-6}, \qquad \text{(A31)}$$

which proves (7).

This range corresponds to $1/a < h < b/a$. The denominator of (A7) remains the same as in the previous range and is given by (A17). Splitting up the integral in

the numerator of (A7) now results in

$$\text{Num}(h) = \int_a^b\int_a^b(xy)^{-\tau}r(ahx)r(ahy)r(xy)\mathrm{d}x\mathrm{d}y$$
$$= \int_{1/(ah)}^{ah}\int_{1/x}^b(xy)^{-\tau}\mathrm{d}y\mathrm{d}x + \int_{ah}^b\int_{1/(ah)}^b(xy)^{-\tau}\mathrm{d}y\mathrm{d}x$$
$$+ \int_{1/(ah)}^{ah}\int_{1/(ah)}^{1/x}(xy)^{1-\tau}\mathrm{d}y\mathrm{d}x$$
$$+ 2ah\int_{ah}^b\int_{1/x}^{1/(ah)}(xy)^{-\tau}y\mathrm{d}y\mathrm{d}x$$
$$+ 2ah\int_{1/(ah)}^{ah}\int_a^{1/(ah)}(xy)^{1-\tau}y\mathrm{d}y\mathrm{d}x$$
$$+ 2ah\int_{ah}^b\int_a^{1/x}(xy)^{1-\tau}y\mathrm{d}y\mathrm{d}x$$
$$+ a^2h^2\int_a^{1/(ah)}\int_a^{1/(ah)}(xy)^{2-\tau}\mathrm{d}y\mathrm{d}x$$
$$=: I_1 + I_2 + I_3 + I_4 + I_5 + I_6 + I_7. \qquad \text{(A32)}$$

Computing these integrals yields

$$I_1 = a^2h^2\Big((ah)^{-2}\frac{\ln(a^2h^2)}{\tau-1} \qquad \text{(A33)}$$
$$+ \frac{b^{1-\tau}\left((ah)^{-\tau-1}-(ah)^{\tau-3}\right)}{(\tau-1)^2}\Big), \qquad \text{(A34)}$$

$$I_2 = a^2h^2\Big(\frac{(ah)^{-2}+b^{2-2\tau}(ah)^{-2}}{(\tau-1)^2} \qquad \text{(A35)}$$
$$- \frac{b^{1-\tau}\left((ah)^{\tau-3}+(ah)^{-\tau-1}\right)}{(\tau-1)^2}\Big), \qquad \text{(A36)}$$

$$I_3 = a^2h^2\left(-(ah)^{-2}\frac{\ln(a^2h^2)}{\tau-2} + \frac{(ah)^{2\tau-6}-(ah)^{-2}}{(\tau-2)^2}\right), \qquad \text{(A37)}$$

$$I_4 = 2a^2h^{-2}\left(-\frac{(abh)^{-1}}{\tau-2} + \frac{(ah)^{-2}}{\tau-1} + \frac{b^{1-\tau}(ah)^{\tau-3}}{(\tau-1)(\tau-2)}\right), \qquad \text{(A38)}$$

$$I_5 = 2a^2h^2\left(\frac{(ah)^{2\tau-6}+h^{1-\tau}a^{4-2\tau}-h^{\tau-3}-(ah)^{-2}}{(3-\tau)(\tau-2)}\right), \qquad \text{(A39)}$$

$$I_6 = 2a^2h^2\Big(\frac{(ab)^{2-\tau}h^{-1}-h^{1-\tau}a^{4-2\tau}}{(3-\tau)(\tau-2)} \qquad \text{(A40)}$$
$$- \frac{(abh)^{-1}-(ah)^{-2}}{3-\tau}\Big), \qquad \text{(A41)}$$

$$I_7 = a^2h^2\left(\frac{a^{6-2\tau}-2h^{\tau-3}+(ah)^{2\tau-6}}{\tau-3}\right). \qquad \text{(A42)}$$

A careful inspection of the terms between brackets in (A34) and (A42) shows that the terms involving $(ah)^{2\tau-6}$ are dominant when $ah \to \infty$. In terms of the variable t in $h = (N\langle h\rangle)^t$, see (11) and (A2), we have that $ah \to \infty$ when we restrict to $t \in [\frac{1}{2}+\varepsilon, 1/(\tau-1)]$ for

any $\varepsilon > 0$. When we retain only these dominant terms, we have, when $ah \to \infty$,

$$\text{Num}(h) = a^2 h^2 (ah)^{2\tau-6}$$
$$\times \left(\frac{1}{(\tau-2)^2} + \frac{2}{(3-\tau)(\tau-2)} + \frac{1}{(3-\tau)^2} \right) (1+o(1))$$
$$= a^2 h^2 \frac{(ah)^{2\tau-6}}{(\tau-2)^2(3-\tau)^2} (1+o(1)).$$
(A43)

Using (A27) again, we get, when $ah \to \infty$,

$$c(h) = \frac{1}{(3-\tau)^2} (ah)^{2\tau-6} a^{2\tau-4} (1+o(1)) \propto N^{5-2\tau} h^{2\tau-6}.$$
(A44)

Furthermore, $c(1/a)$ is given by (A30), while $c(h)$ decreases in h. This gives (A31).

Other connection probabilities In [31] we have presented a class of functions $r(u) = uf(u)$, $u \geq 0$, so that

$$p(h, h') = r(u) \quad \text{with} \quad u = \frac{hh'}{h_s^2}. \quad (\text{A45})$$

has appropriate monotonicity properties. The maximal member $r(u) = \min(u, 1)$ of this class yields p in (3) and is quite representative of the whole class, while allowing explicit computation and asymptotic analysis of $c(h)$ as in [31] and this paper. Figure 7 shows that other asymptotically equivalent choices such as $r(u) = u/(1+u)$ and $r(u) = 1 - e^{-u}$ have comparable clustering spectra. A minor difference is that the choice $r(u) = \min(1, u)$ for p in (3) forces $c(h)$ to be constant on the range $h \leq N^{\beta(\tau)}$, while the other two choices show a gentle decrease.
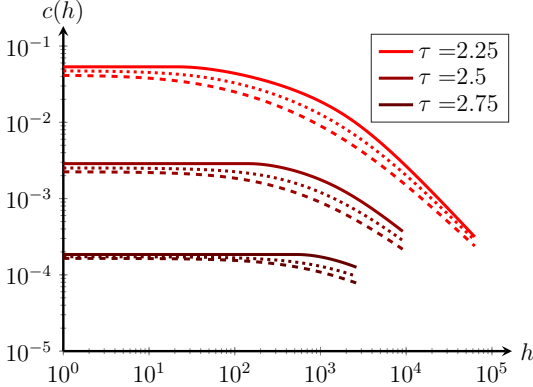

Figure 7. $c(h)$ for $r(u) = \min(u, 1)$ (line), $r(u) = u/(1+u)$ (dashed) and $r(u) = 1 - e^{-u}$ (dotted), obtained by calculating (A7) numerically.

Limiting form of $\sigma_N(t)$ and finite-size effects We consider $\sigma_N(t)$ as in (A2) with $h_{\text{ref}} = 0$. Using (A8), (A16) and (A31), it is readily seen that

$$\lim_{N \to \infty} \sigma_N(t) = \begin{cases} 0, & 0 \leq t \leq \frac{1}{2}, \\ (3-\tau)(1-2t), & \frac{1}{2} \leq t \leq \frac{1}{\tau-1}. \end{cases}$$
(A46)

Hence, some of the detailed information that is present in (A8), (A16) and (A31), disappears when taking the limit as in (A46). This is in particular so for the $\ln N$-factor in (A8) and the logarithmic decaying factor $\ln(N^2/h)$ in Region II.

Consider $\sigma_N(t)$ of (A2) with $h_{\text{ref}} = h_c$ as is done in Fig. 5. It follows from the detailed form of (A8) and (A31), that

$$\sigma_N(0) = \frac{\ln(c(0)/c(h_c))}{\ln(N\langle h \rangle)} = \gamma + \frac{\ln(\beta y)}{y}, \quad (\text{A47})$$

where

$$\gamma = \frac{(3-\tau)^2}{\tau-1}, \quad \beta = (\tau-2)\gamma, \quad y = \ln(N\langle h \rangle). \quad (\text{A48})$$

We have that $\sigma_N(0) \to \gamma$ as $N \to \infty$, and the right-hand side of (A47) exceeds this limit γ from $y = 1/\beta$ onwards with a maximum excess β/e for $N\langle h \rangle$ as large as $\exp(e/\beta)$. This explains why the excess of $\sigma_N(0)$ over its limit value in Fig. 5 with $e^{e/\beta} = 3 \times 10^{10}$ when $\tau = 9/4$ persists.

Appendix B: Exact and asymptotic result for decay rate of $c(h)$ at $h = h_c$ and $h = h_s$

We let $h_c = (N\langle h \rangle)^{1/(\tau-1)}$, where we assume that N is so large that $h_c \leq N$. This requires N to be of the order $(1/\varepsilon)^{1/\varepsilon}$, where $\varepsilon = \tau - 2$. We again consider the function $\sigma_N(t)$ of (11),

$$\sigma_N(t) = \frac{\ln(c(h)/c(h_{\text{ref}}))}{\ln(N\langle h \rangle)}, \quad h = (N\langle h \rangle)^t, \quad (\text{B1})$$

for $0 \leq t \leq \frac{1}{\tau-1}$ and h_{ref} is fixed, so that

$$c(h) = c(h_{\text{ref}})(N\langle h \rangle)^{\sigma_N(t)}, \quad h = (N\langle h \rangle)^t. \quad (\text{B2})$$

When we fix a t_0 and linearize $\sigma_N(t)$ around t_0, we get

$$c(h) \approx c(h_{\text{ref}})(N\langle h \rangle)^{\sigma_N(t_0)+(t-t_0)\sigma'_N(t_0)}$$
$$= c(h_0) \left(\frac{h}{h_0} \right)^{\sigma'_N(t_0)} \quad (\text{B3})$$

so that $\sigma'_N(t) = \frac{d}{dt}\sigma_N(t)$ is a measure for the decay rate of $c(h)$ at $h = h_0 = (N\langle h \rangle)^{t_0}$.

In this appendix, we compute an exact expression for $\sigma'_N(t)$ at $t = \frac{1}{\tau-1}$, we compute its limit as $N \to \infty$ and discuss convergence speed, and we show that this limit is a lower bound for $\sigma'_N(t)$.

More precisely, we show the following result:

Proposition 1. *Let a and b be as in* (A5). *Then,*

$$\sigma'_N \left(\frac{1}{\tau-1} \right) = -2 \left(\frac{A + \frac{3-\tau}{\tau-2}C}{A + \frac{4-\tau}{\tau-2}C} - \frac{D}{E+D} \right), \quad (\text{B4})$$

where

$$A = \frac{1}{b^2}\left(\frac{-\ln(b^2)}{(\tau-1)(\tau-2)} - \frac{1-b^{2(1-\tau)}}{(\tau-1)^2} + \frac{b^{2(\tau-2)}-1}{(\tau-2)^2}\right), \tag{B5}$$

$$C = \left(\frac{b^{\tau-3}-a^{3-\tau}}{3-\tau}\right)^2, \tag{B6}$$

$$D = \frac{1}{b}\frac{b^{\tau-1}-b^{1-\tau}}{\tau-1}, \tag{B7}$$

$$E = \frac{a^{2-\tau}-b^{\tau-2}}{\tau-2}. \tag{B8}$$

Furthermore,

$$\sigma'_N(\tfrac{1}{\tau-1}) > \lim_{M\to\infty}\sigma'_M\left(\tfrac{1}{\tau-1}\right) = -2(3-\tau) \tag{B9}$$

for all N.

The limiting value in (B9) is consistent with the limiting value of $\sigma_N(t)$ that has been found in (A46). We assess this convergence result with plots. While these indicate that the limits are only reached for very large N, especially when τ is close to 2, it can also be seen that the limiting shape of $\sigma_N(t)$ already shows up for considerably smaller N.

To start the proof of Proposition 1, note that in the a, b notation of (A5),

$$c(h) = \frac{K(h)}{J(h)}, \quad 0 \le h \le h_c, \tag{B10}$$

where

$$K(h) = \int_a^b \int_a^b (xy)^{2-\tau} f(ahx)f(ahy)f(xy)\mathrm{d}x\mathrm{d}y, \tag{B11}$$

$$J(h) = \left(\int_a^b x^{1-\tau} f(ahx)\mathrm{d}x\right)^2, \tag{B12}$$

with $f(u) = \min(1, u^{-1})$. Note that $r(u) = uf(u)$, see (A6). We compute

$$\begin{aligned}
\sigma'_N(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\ln\left(c((N\langle h\rangle)^t)/c(h_{\mathrm{ref}})\right)}{\ln(N\langle h\rangle)}\right) \\
&= (N\langle h\rangle)^t \ln(N\langle h\rangle)\frac{c'((N\langle h\rangle)^t)}{c((N\langle h\rangle)^t)\ln(N\langle h\rangle)} \\
&= h\frac{c'(h)}{c(h)}, \quad h = (N\langle h\rangle)^t, \tag{B13}
\end{aligned}$$

where the prime on c indicates differentiation with respect to h. With (B10) we get

$$\frac{c'(h)}{c(h)} = \frac{K'(h)}{K(h)} - \frac{J'(h)}{J(h)}, \tag{B14}$$

and we have to evaluate $K(h), K'(h), J(h)$ and $J'(h)$ at

$$h = h_c = b/a. \tag{B15}$$

Lemma 1.

$$K(h_c) = A + \frac{4-\tau}{2-\tau}C, \quad K'(h_c) = \frac{-2a}{b}\left(A + \frac{3-\tau}{\tau-2}C\right), \tag{B16}$$

$$J(h_c) = (D+E)^2, \quad J'(h_c) = -\frac{2a}{b}(D+E)D, \tag{B17}$$

with A, C, D, E as in (B5)–(B8).

From Lemma 1, (B13) and (B15) we get (B4) in Proposition 1.

Proof of Lemma 1. Since $h_c = b/a$,

$$K(h_c) = \int_a^b \int_a^b (xy)^{2-\tau} f(bx)f(by)f(xy)\mathrm{d}x\mathrm{d}y. \tag{B18}$$

With $f(u) = \min(1, u^{-1})$ we split up the integration range $[a,b]\times[a,b]$ into the four regions $[a,1/b]\times[a,1/b]$, $[1/b,b]\times[1/b,b]$, $[1/b,b]\times[a,1/b]$ and $[a,1/b]\times[1/b,b]$, where we observe that $a \le 1/b \le 1 \le b$. We first get

$$\begin{aligned}
&\int_a^{1/b}\int_a^{1/b}(xy)^{2-\tau}f(bx)f(by)f(xy)\mathrm{d}x\mathrm{d}y \\
&= \int_a^{1/b}\int_a^{1/b}(xy)^{2-\tau}\cdot 1\cdot 1\cdot 1\mathrm{d}x\mathrm{d}y \\
&= \left(\frac{b^{\tau-3}-a^{3-\tau}}{3-\tau}\right)^2 = C. \tag{B19}
\end{aligned}$$

Next,

$$\begin{aligned}
&\int_{1/b}^b\int_{1/b}^b(xy)^{2-\tau}f(bx)f(by)f(xy)\mathrm{d}x\mathrm{d}y \\
&= \int_{1/b}^b\int_{1/b}^b(xy)^{2-\tau}\frac{1}{bx}\frac{1}{by}f(xy)\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{b^2}\int_{1/b}^b\int_{1/b}^b(xy)^{1-\tau}f(xy)\mathrm{d}x\mathrm{d}y. \tag{B20}
\end{aligned}$$

The remaining double integral with $\tau+1$ instead of τ has been evaluated in [31, Appendix C, (C3)] as

$$-\frac{\ln(b^2)}{(\tau-1)(\tau-2)} - \frac{1-b^{2(1-\tau)}}{(\tau-1)^2} + \frac{b^{2(\tau-2)-1}}{(\tau-2)^2} = b^2 A. \tag{B21}$$

Finally, the two double integrals over $[1/b,b]\times[a,1/b]$ and $[a,1/b]\times[1/b,b]$ are by symmetry both equal to

$$\begin{aligned}
&\int_{1/b}^b\int_a^{1/b}(xy)^{2-\tau}f(bx)f(by)f(xy)\mathrm{d}x\mathrm{d}y \\
&= \int_{1/b}^b\int_a^{1/b}(xy)^{2-\tau}\frac{1}{bx}\cdot 1\cdot 1\mathrm{d}x\mathrm{d}y \\
&= \frac{1}{b}\frac{b^{\tau-2}-b^{2-\tau}}{\tau-2}\frac{b^{\tau-3}-a^{3-\tau}}{3-\tau} = \frac{(b^{\tau-3}-a^{3-\tau})^2}{(\tau-2)(3-\tau)} \\
&= \frac{3-\tau}{\tau-2}C. \tag{B22}
\end{aligned}$$

Here we have used that, see (A5),

$$b^{1-\tau} = a^{3-\tau}. \tag{B23}$$

Now the expression in (B16) for $K(h_c)$ follows.

To evaluate $K'(h_c)$, we observe by symmetry that

$$K'(h) = 2\int_a^b \int_a^b (xy)^{2-\tau} axf'(ahx)f(ahy)f(xy)\mathrm{d}x\mathrm{d}y. \tag{B24}$$

At $h = h_c$, we have $ah = b$, and so

$$K'(h_c) = 2\frac{a}{b}\int_a^b \int_a^b (xy)^{2-\tau} bxf'(bx)f(by)f(xy)\mathrm{d}x\mathrm{d}y. \tag{B25}$$

Now $uf'(u) = 0$ for $0 \le u \le 1$ and $uf'(u) = -f(u)$ for $u \ge 1$. Hence, splitting up the integration range into the four regions as earlier, we see that those over $[a, 1/b] \times [a, 1/b]$ and $[a, 1/b] \times [1/b, b]$ vanish while those over $[1/b, b] \times [1/b, b]$ and $[1/b, b] \times [a, 1/b]$ give rise to the same double integrals as in (B20) and (B22) respectively. This yields the expression in (B16) for $K'(h_c)$.

The evaluation of $J(h_c)$ and $J'(h_c)$ is straightforward from (B12) with $ah = b$ and a splitting of the integration range $[a, b]$ into $[a, 1/b]$ and $[1/b, b]$. This yields (B17), and the proof of Lemma 1 is complete.

We now turn to the limiting behavior of $\sigma'_N(\frac{1}{\tau-1})$ as $N \to \infty$. For this we write

$$0 < \frac{D}{D+E} = \frac{1 - b^{2(1-\tau)}}{\frac{\tau-1}{\tau-2}(ab)^{2-\tau} - \frac{1}{\tau-2} - \frac{1}{\tau-1}b^{2(1-\tau)}}, \tag{B26}$$

in which

$$b^{2(1-\tau)} = (N\langle h\rangle)^{\tau-3} \to 0, \tag{B27}$$

$$(ab)^{2-\tau} = (N\langle h\rangle)^{\frac{(\tau-2)^2}{\tau-1}} \to \infty, \tag{B28}$$

as $N \to \infty$. Hence, $D/(D+E) \to 0$ as $N \to \infty$. Furthermore, we write

$$C = \frac{b^{2(\tau-3)}}{(\tau-3)^2}\left(1 - (ab)^{3-\tau}\right)^2, \tag{B29}$$

and

$$A = \frac{b^{2(\tau-3)}}{(\tau-2)^2}(1 - F), \tag{B30}$$

where

$$F = b^{-2(\tau-2)}\left[\frac{\tau-2}{\tau-1}\ln(b^2)\right.$$
$$\left. + \left(\frac{\tau-2}{\tau-1}\right)^2(1 - b^{2(1-\tau)}) + 1\right]$$
$$= \frac{1}{\tau-1}b^{-2(\tau-2)}\ln(b^{2(\tau-2)})\left(1 + O\left(\frac{1}{\ln(b)}\right)\right). \tag{B31}$$

Now, using (B23), we have

$$(ab)^{3-\tau} = b^{-2(\tau-2)} = (N\langle h\rangle)^{\frac{(\tau-2)(3-\tau)}{\tau-1}} \to 0 \tag{B32}$$

as $N \to \infty$. Thus, we get

$$\lim_{N\to\infty} \frac{A + \frac{3-\tau}{2-\tau}C}{A + \frac{4-\tau}{2-\tau}C} = \frac{\frac{1}{(\tau-2)^2} + \frac{3-\tau}{\tau-2}\frac{1}{(3-\tau)^2}}{\frac{1}{(\tau-2)^2} + \frac{4-\tau}{\tau-2}\frac{1}{(3-\tau)^2}} = 3 - \tau, \tag{B33}$$

and this yields (B9).

Note that $D/(D+E)$ approaches 0 much slower than the limit in (B33) is reached when τ is close to 2, compare (B28) and (B33). Thus, we can concentrate on $D/(D+E)$, and the relative deviation of $\sigma'_N(t)$ from $-2(3-\tau)$ is approximately

$$\frac{2D}{D+E}\frac{1}{2(3-\tau)} \approx \frac{\tau-2}{3-\tau}\frac{1}{(ab)^{2-\tau} - 1}$$
$$\approx \frac{\tau-2}{3-\tau}(N\langle h\rangle)^{-\frac{(\tau-2)^2}{\tau-1}}. \tag{B34}$$

We finally turn to the inequality in (B9) in Proposition 1. Obviously, we have

$$\sigma'_N\left(\frac{1}{\tau-1}\right) > -2\frac{A + \frac{3-\tau}{\tau-2}C}{A + \frac{4-\tau}{\tau-2}C}. \tag{B35}$$

We shall show that

$$\frac{A + \frac{3-\tau}{\tau-2}C}{A + \frac{4-\tau}{\tau-2}C} \le \frac{A_{\mathrm{as}} + \frac{3-\tau}{\tau-2}C_{\mathrm{as}}}{A_{\mathrm{as}} + \frac{4-\tau}{\tau-2}C_{\mathrm{as}}} = 3 - \tau, \tag{B36}$$

where

$$A_{\mathrm{as}} = \frac{b^{2(\tau-3)}}{(\tau-2)^2}, \quad C_{\mathrm{as}} = \frac{b^{2(\tau-3)}}{(3-\tau)^2}, \tag{B37}$$

the asymptotic form of A and C as $N \to \infty$ obtained from (B30) and (B29) by deleting F and $(ab)^{3-\tau}$, respectively. The function

$$x \in [0, \infty) \mapsto \frac{1 + \frac{3-\tau}{\tau-2}x}{1 + \frac{4-\tau}{\tau-2}x} \tag{B38}$$

is decreasing in $x \ge 0$, and so it suffices to show that

$$\frac{C_{\mathrm{as}}}{A_{\mathrm{as}}} \le \frac{C}{A}, \text{ i.e., that } \frac{C_{\mathrm{as}}}{C} \le \frac{A_{\mathrm{as}}}{A}. \tag{B39}$$

We have from (B29) that

$$\frac{C_{\mathrm{as}}}{C} = \frac{1}{\left(1 - (ab)^{3-\tau}\right)^2}, \tag{B40}$$

and from (B30) and (B31) that

$$\frac{A}{A_{\mathrm{as}}} = 1 - F = 1 - b^{-2(\tau-2)} - b^{-2(\tau-2)}$$
$$\times \left[\frac{\tau-2}{\tau-1}\ln(b^2) + \left(\frac{\tau-2}{\tau-1}\right)^2(1 - b^{2(1-\tau)})\right]. \tag{B41}$$

Using that $(ab)^{3-\tau} = b^{-2(\tau-2)}$, see (B32), we see that the inequality $C_{\mathrm{as}}/C \le A_{\mathrm{as}}/A$ in (B39) is equivalent to

$$(1 - b^{-2(\tau-2)})^2 \ge 1 - b^{-2(\tau-2)} - b^{-2(\tau-2)}$$
$$\times \left[\frac{\tau-2}{\tau-1}\ln(b^2) + \left(\frac{\tau-2}{\tau-1}\right)^2(1 - b^{2(1-\tau)})\right]. \tag{B42}$$

Using that $(1-u)^2 - (1-u) = -u(1-u)$ and dividing through by $u = b^{-2(\tau-2)}$, we see that (B42) is equivalent to

$$\frac{\tau-2}{\tau-1}\ln(b^2) + \left(\frac{\tau-2}{\tau-1}\right)^2(1-b^{2(1-\tau)}) \geq 1 - b^{-2(\tau-2)}. \tag{B43}$$

With $y = \ln(b^2) \geq 0$, we write (B43) as

$$K(y) := \left(\frac{\tau-2}{\tau-1}\right)^2(1-\mathrm{e}^{(1-\tau)y}) + \frac{\tau-2}{\tau-1}y - (1-\mathrm{e}^{(2-\tau)y}) \geq 0. \tag{B44}$$

Taylor development of $K(y)$ at $y = 0$ yields

$$K(y) = 0 \cdot y^0 + 0 \cdot y^1 + 0 \cdot y^2 + \frac{1}{6}(\tau-2)^2 y^3 + \dots \tag{B45}$$

Furthermore,

$$K''(y) = (\tau-2)^2 \mathrm{e}^{(1-\tau)y}(\mathrm{e}^y - 1) > 0, \quad y > 0. \tag{B46}$$

Therefore, $K(0) = K'(0) = 0$, while $K''(y) > 0$ for $y > 0$. This gives $K(y) > 0$ when $y > 0$, as required.

Similar to Proposition 1, we can derive the following result for $\sigma'_N(\frac{1}{2})$:

Proposition 2.

$$\sigma'_N(\tfrac{1}{2}) = -2\left(\frac{G+H}{\left(1+\left(\frac{\tau-1}{3-\tau}\right)^2\right)G + 2H} - \frac{I}{I+J}\right), \tag{B47}$$

where

$$G = \left(\frac{1-b^{1-\tau}}{\tau-1}\right)^2, \tag{B48}$$

$$I = \frac{1-b^{1-\tau}}{\tau-1}, \tag{B49}$$

$$J = \frac{b^{(\tau-2)(\tau-1)/(3-\tau)} - 1}{\tau-2}, \tag{B50}$$

$$H = \frac{1-1/b-b^{1-\tau}(1-b^{2-\tau})}{(\tau-2)(3-\tau)} - \frac{1-b^{1-\tau}}{(\tau-1)(\tau-2)}. \tag{B51}$$

Furthermore,

$$\sigma'_N(\tfrac{1}{2}) > \lim_{M\to\infty}\sigma'_M(\tfrac{1}{2}) = -1 + \frac{2(\tau-2)}{3-(\tau-2)^2}, \tag{B52}$$

for all N.

Figure 8 shows the values of $\sigma'_N(\frac{1}{2})$ and $\sigma'_N(\frac{1}{\tau-1})$ for finite-size networks together with its limiting value. For example, when $\tau = 2.25$, Fig. 8a shows that N needs to be of the order 10^{16} for the slope to be 'close' to its limiting value -1.5. When for example $N = 10^6$ we see that the slope is much smaller: approximately -1.1. This makes statistical estimation of the true underling power-law exponent α extremely challenging, especially for the relevant regime τ close to 2, because enormous amounts of data should be available to get sufficient statistical accuracy. Most data sets, even the largest available networks

used in this paper, are simply not large enough to have sufficiently many samples from the large-degree region to get a statistically accurate estimate of the power-law part. This also explains why based on smaller data sets it is common to assume that α is roughly one [4, 10, 33–36]. Comparing Fig. 8a and Fig. 8b shows that the convergence to the limiting value is significantly faster at the point $t = \frac{1}{2}$ than at the point $t = \frac{1}{\tau-1}$.

Appendix C: From hidden variables to degrees

In this paper, we focus on computing $c(h)$, the local clustering coefficient of a randomly chosen vertex with hidden variable h. However, when studying local clustering in real-world data sets, we can only observe $\bar{c}(k)$, the local clustering coefficient of a vertex of degree k. In this appendix, we show that for the hidden variable model, the difference between these two methods of computing the clustering coefficient is small and asymptotically negligible. We consider

$$c(h) = \frac{\int_1^{h_c}\int_1^{h_c}(xy)^{2-\tau}p(h,h')p(h,h'')p(h',h'')\mathrm{d}h'\mathrm{d}h''}{\left(\int_1^{h_c}x^{1-\tau}p(h,h')\mathrm{d}h'\right)^2}. \tag{C1}$$

We define $\bar{c}(k)$ as the average clustering coefficient over all vertices of degree k. By [32], the probability that a vertex with hidden variable h has degree k equals

$$g(k \mid h) = \frac{\mathrm{e}^{-h}h^k}{k!}. \tag{C2}$$

Then, by [32],

$$\bar{c}(k) = \begin{cases} \frac{1}{P(k)}\int_1^{h_c}\rho(h)c(h)g(k \mid h)\mathrm{d}h, & k \geq 2, \\ 0, & k < 2, \end{cases} \tag{C3}$$

where $\bar{c}(k) = 0$ for $k < 2$ because a vertex with degree less than 2 cannot be part of a triangle. Here

$$P(k) = \int_1^{h_c} g(k \mid h)\rho(h)\mathrm{d}h \tag{C4}$$

is the probability that a randomly chosen vertex has degree k.

First we consider the case where $h > N^{\frac{\tau-2}{\tau-1}}$. The Chernoff bound gives for the tails of the Poisson distribution that

$$\mathbb{P}\left(\mathrm{Poi}(\lambda) > x\right) \leq \mathrm{e}^{-\lambda}\left(\frac{\mathrm{e}\lambda}{x}\right)^x, \quad x > \lambda, \tag{C5}$$

$$\mathbb{P}\left(\mathrm{Poi}(\lambda) < x\right) \leq \mathrm{e}^{-\lambda}\left(\frac{\mathrm{e}\lambda}{x}\right)^x, \quad x < \lambda. \tag{C6}$$

Let $k(h)$ be the degree of a node with hidden variable h. Then, for any $M > 1$

$$\sum_{k=Mh}^{\infty} g(k \mid h) \leq \left(\frac{\mathrm{e}^{M-1}}{M^M}\right)^h, \tag{C7}$$
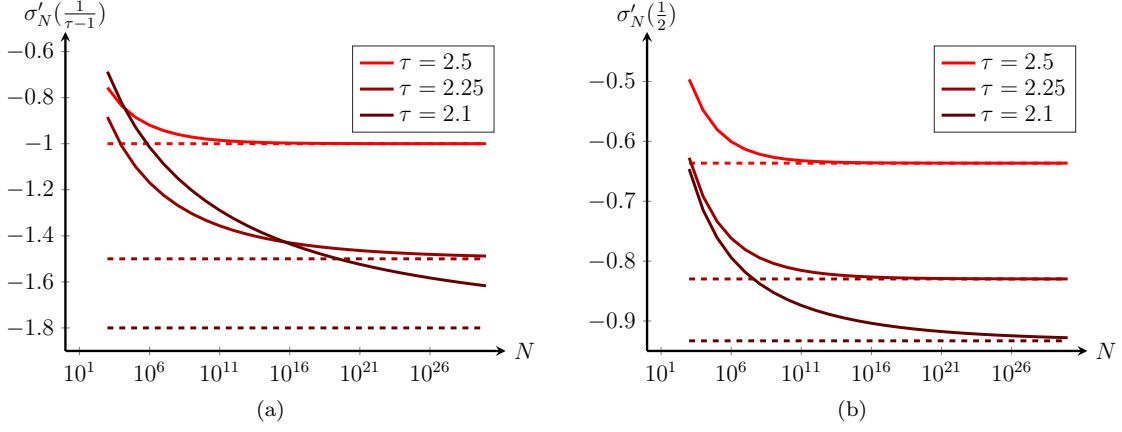
Figure 8. $\sigma'_N(t)$ plotted against N for (a) $t = \frac{1}{\tau-1}$ and (b) $t = \frac{1}{2}$. The dashed line gives the limiting value of $\sigma'_N(t)$ as $N \to \infty$.
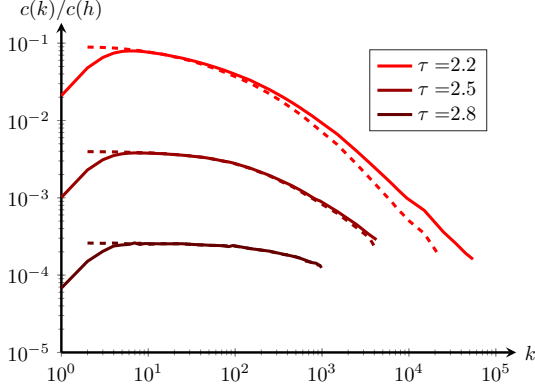

Figure 9. $\bar{c}(k)$ (dashed) and $c(h)$ (line) for $N = 10^5$, averaged over 10^4 realizations.

Furthermore, $c(h) = c(0)$ in this regime of h. This results in

$$\bar{c}(k) \approx \frac{c(0) \int_1^{h_s^2/h_c} \rho(h)g(k \mid h)\mathrm{d}h}{\int_1^{h_s^2/h_c} \rho(h)g(k \mid h)\mathrm{d}h} = c(0). \qquad (C13)$$

Therefore, $\bar{c}(h) \approx c(h)$ also when h is small.

Figure 9 shows that indeed the difference between $\bar{c}(k)$ and $c(k)$ is small. When τ approaches 2, the difference becomes larger. We see that for small values of k, $\bar{c}(k)$ and $c(k)$ are not very close. This is due to the fact that (C1) does not take into account that a vertex with hidden variable h may have less than 2 neighbors, so that its local clustering is zero. In [31] we show how to adjust (A7) to account for this.

and for any $\delta \in (0, 1)$,

$$\sum_{k=1}^{\delta h} g(k \mid h) \le \left(\frac{\mathrm{e}^{\delta-1}}{\delta^\delta}\right)^h. \qquad (C8)$$

Because $\mathrm{e}^{x-1}/x^x < 1$ for $x \ne 1$, (C7) and (C8) tend to zero as $h \to \infty$. Therefore, for h large,

$$k(h) = h(1 + o(1)) \qquad (C9)$$

with high probability. Therefore, when k is large,

$$\bar{c}(k) \approx c(k). \qquad (C10)$$

Thus, $c(h)$ is very similar to $\bar{c}(k)$.

On the other hand, for $h \ll h_s^2/h_c$,

$$\sum_{h_s^2/h_c}^{\infty} g(k \mid h) \le \mathrm{e}^{-h}\left(\frac{\mathrm{e}h}{h_s^2/h_c}\right)^{h_s^2/h_c}, \qquad (C11)$$

which is small by the assumption on h. Thus,

$$P(k) \approx \int_1^{h_s^2/h_c} g(k \mid h)\rho(h)\mathrm{d}h. \qquad (C12)$$