Clustering Does Not Always Imply Latent Geometry

Roya Aliakbarisani,[1,2,*] Marián Boguñá,[1,2,†] and M. Ángeles Serrano[1,2,3,‡]
[1]Departament de Física de la Matèria Condensada, Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain
[2]Universitat de Barcelona Institute of Complex Systems (UBICS), Barcelona, Spain
[3]Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, E-08010 Barcelona, Spain

(Received 3 March 2025; accepted 9 October 2025; published 6 November 2025)

The latent space approach to complex networks has revealed fundamental principles and symmetries, enabling geometric methods. However, the conditions under which network topology implies geometricity remain unclear. We provide a mathematical proof and empirical evidence showing that multiscale self-similarity in complex networks is a crucial factor of latent geometry. Using degree-thresholding renormalization, we prove that a general class of ensembles of random scale-free networks in Riemannian manifolds of constant curvature and of any dimension are self-similar when interactions are pairwise. Hence, both nonvanishing local clustering in the thermodynamic limit and self-similarity are required to imply geometricity. Our findings highlight that correlated links can lead to a finite clustering coefficient without self-similarity, and therefore without inherent latent geometry. The implications are significant for network mapping and ensemble equivalence between graphs and continuous spaces.

DOI: 10.1103/ycwq-92ms

Network geometry [1] has emerged as a key paradigm in network science to model real-world networks at both local and global scales. Moreover, it provides a robust framework to conceptualize problems in physics related to ensemble equivalence between graphs of discrete units and emerging continuous space or spacetimes, including approaches to quantum gravity [2] and the study of causal sets [3–5]. In particular, the latent geometry approach [6]—where nodes connect with a likelihood that decreases with distance in a hidden metric space—explains essential network features [7]. Notably, this approach has revealed that real networks exhibit an effective hyperbolic geometry [8,9], unifying the small-world property with heterogeneous degree distributions and high local clustering under a single mechanism.

The key property that connects the geometry of a network and its topology is the existence of triangles in the network structure [10], as a reflection of the triangle inequality in the metric space. This raises the intriguing question of whether local clustering, measuring the likelihood that two neighbors of a node are also connected, implies geometricity. Previous work [11] suggested that local clustering implies geometry by showing that random graphs with homogeneous degrees and local clustering are equivalent to random geometric graphs. However, this holds only under certain conditions. Furthermore, using triangle counts to detect geometry has clear limitations, as manifested by network models without inherent geometry, such as the configuration model, which can still contain many triangles. This led to the development of a weighting scheme designed to determine reliably whether local clustering is induced by hyperbolic spaces and persists in the thermodynamic limit, which requires system size scaling [12].

In this work, we unveil the crucial role of network symmetries to elucidate the conditions under which network topology implies geometricity. By geometricity, we mean networks that emerge from ensembles of random geometric graphs in Riemannian manifolds of any dimension and with constant curvature, where nodes, in addition to having intrinsic attributes, are assigned positions in a latent metric space. Independent pairwise links are formed with probabilities depending on both the intrinsic node attributes and distance, while ensuring local separability (LS). By LS, we mean that the expected degree of a node is only a function of its intrinsic attribute. This definition captures a broad class of random geometric graph models under minimal assumptions.

To reveal the fundamental symmetry of a given network, we used the degree-thresholding renormalization (DTR) method [13]. This method generates a nested hierarchy of subgraphs by progressively filtering out nodes with degree below a threshold. Multiscale self-similarity appears when, in the thermodynamic limit, such subgraphs and the original network belong to the same statistical ensemble. DTR self-similarity has been observed in many real-world networks and has been explained by the renormalizability of the geometric $\mathbb{S}^1$ model [13]. Self-similarity of complex networks has also been explored from other perspectives [14–18].

*Contact author: 85aliakbari@gmail.com
†Contact author: marian.boguna@ub.edu
‡Contact author: marian.serrano@ub.edu

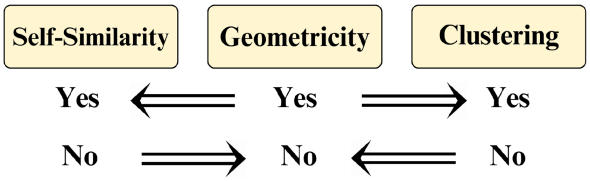

FIG. 1. Logic relationships between geometricity, clustering, and self-similarity for geometric random graphs in Riemannian manifolds of any dimension and with arbitrary constant curvature.

Below, we prove that any undirected geometric random graph in a Riemannian manifold of any dimension and with arbitrary constant curvature is self-similar under DTR, provided nodes have a power-law degree distribution and pairwise interactions that satisfy LS. Therefore, non-self-similar networks cannot be geometric in the sense defined here. We found that some models with finite local clustering are non-self-similar under DTR, thus nongeometric. Hence, while geometricity always implies local clustering and self-similarity, the reverse is not true; however, the absence of local clustering or self-similarity implies nongeometricity (see Fig. 1). Moreover, the DTR self-similarity of real networks can be quantified and, together with local clustering, used to indicate their potential geometricity. Thus, local clustering (hereafter referred to in this Letter as clustering), geometry, and self-similarity interplay in a nontrivial way in complex networks.

We start our proof by considering the thermodynamic limit of a homogeneous Poisson point process (PPP) with density δ in a Euclidean space of dimension d defining the positions of nodes with intrinsic attributes. The same is obtained for hyperbolic and spherical manifolds; the proofs for negative and positive curvature are in the End Matter Appendixes A and B, respectively. Note that the proof for spherical manifolds applies, in particular, to the $\mathbb{S}^d$ geometric network model [13] and, thus, to its isomorphic purely geometric hyperbolic formulation the $\mathbb{H}^{d+1}$ model [19,20], where nodes without intrinsic attributes are distributed within a finite disk in the $(d+1)$-dimensional hyperbolic space. In contrast, the proof for hyperbolic manifolds in End Matter Appendix A refers to a new model where nodes with intrinsic attributes are distributed homogeneously over the entire d-dimensional hyperbolic space. The three schemes correspond to families of models that generate networks that are simultaneously sparse, highly clustered, and self-similar, and exhibit arbitrary degree distributions, but differ in the curvature of the underlying space.

The latent space conceptualizes a similarity space, in which closeness reflects how similar nodes are and determines their likelihood to form connections. Furthermore, we assume each node has a popularity attribute, independent of its position in space, ultimately accounting for its number of neighbors in the network. We represent popularity by a hidden variable h (a positive real number), distributed as $\rho(h)$. Suppose now that interactions are pairwise, such that the connection probability between a pair of nodes i and j with hidden variables h_i and h_j and separated by a distance x_{ij} takes the form

$$p_{ij} = f\left(\frac{x_{ij}}{h_i h_j}\right). \tag{1}$$

We do not impose any particular form to the function $f(\cdot)$ other than being integrable in $\mathbb{R}^d$, that is, $\int_0^\infty t^{d-1} f(t) dt < \infty$. As we show below, this implies that the model generates sparse networks without the need to introduce a size dependence on the connection probability [21]. Since clustering is defined as the probability that two nodes that share a common neighbor are connected, it is a function of p_{ij} and must be finite in the thermodynamic limit. This defines our ensemble of random geometric graphs $\mathcal{G}$.

In this setting, the expected degree of node i is simply given by $\bar{k}(h_i) = \sum_{j \neq i} p_{ij}$. However, since the space is homogeneous and isotropic, and the hidden variable h is independent of space, we can place node i at the origin of coordinates so that x_{ij} becomes the radial distance between the nodes in spherical coordinates. Then, taking the continuum limit,

$$\bar{k}(h_i) = S_{d-1}\delta \int dh_j \rho(h_j) \int dx_{ij} x_{ij}^{d-1} f\left(\frac{x_{ij}}{h_i h_j}\right), \tag{2}$$

where we have already integrated the angular part and S_{d-1} is the volume of the $(d-1)$-dimensional unit sphere. The change of variables $t = (x_{ij}/h_i h_j)$ leads to $\bar{k}(h_i) = \delta I_d \langle h^d \rangle h_i^d$, where $I_d \equiv S_{d-1} \int_0^\infty f(t) t^{d-1} dt$, so that the expected average degree of the network is

$$\langle k \rangle = \delta I_d \langle h^d \rangle^2. \tag{3}$$

As long as the integral I_d is bounded, the average degree is constant and so the network is sparse. Notice that the particular form of the argument of the connection probability in Eq. (1) [and similarly in Eqs. (A3) and (B5)] ensures that the model satisfies LS. Hereafter, we constrain the distribution of hidden variables h to be power law with $\rho(h) = (\tilde{\gamma}-1) h_0^{(\tilde{\gamma}-1)} h^{-\tilde{\gamma}}$, $h \geq h_0$, and $\tilde{\gamma} > d+1$.

We are now interested in the properties of subgraphs obtained after applying the DTR procedure to networks generated by the ensemble $\mathcal{G}$. Given a graph $G \in \mathcal{G}$ and its hidden variables, we obtain a subgraph G_T by removing all nodes with hidden variable $h_0 \leq h \leq h_T$ from G, defining the ensemble $\mathcal{G}_T$. Since we are truncating a power-law distribution from below, the hidden variables of nodes in G_T are also power-law distributed with the same exponent $\tilde{\gamma}$

but starting at h_T instead of h_0. Besides, since the popularity dimension encoded in the hidden variable h is uncorrelated with the position in the metric space, the original PPP in $\mathcal{G}$ defines a homogeneous PPP in $\mathcal{G}_T$ but at density δ_T. Thus, the ensemble $\mathcal{G}_T$ is the same as $\mathcal{G}$ in the thermodynamic limit, with the only difference that the density δ_T of the PPP and the average $\langle h^d \rangle_T$ are rescaled. Hence, ensemble $\mathcal{G}$ is statistically invariant under DTR and contains an infinite hierarchy of nested subgraphs belonging to the same ensemble. The only relevant observable is the average degree, which may change in the DTR flow. Its renormalized value is obtained from Eq. (3), and reads

$$\langle k \rangle_T = \langle k \rangle \left(\frac{\delta}{\delta_T} \right)^{\frac{3-\gamma}{\gamma-1}}, \tag{4}$$

where we have used

$$\delta_T = \delta \left(\frac{h_0}{h_T} \right)^{(\tilde{\gamma}-1)}, \qquad \langle h^d \rangle_T = \langle h^d \rangle \left(\frac{h_T}{h_0} \right)^d, \tag{5}$$

and the definition of the expected degree of node i, which is proportional to h^d as $\kappa_i \equiv \langle k \rangle h_i^d / \langle h^d \rangle$. This gives $\langle \kappa \rangle = \langle k \rangle$, and κ power-law distributed with an exponent $\gamma = 1 + (\tilde{\gamma} - 1)/d$. Since κ is a monotonically increasing function of h, thresholding by h_T is equivalent to thresholding by κ_T. Note that the expression for δ_T in Eq. (5) comes from integrating the power-law degree distribution in the range (h_T, ∞), combined with the definition of density [22].

Hence, we have demonstrated that a random geometric graph with pairwise interactions and a power-law degree distribution defined in a Euclidean metric space of any dimension is necessarily self-similar under DTR and the only property that may change in the flow is the average degree. Therefore, if a network does not exhibit self-similarity across the DTR hierarchy of subgraphs, it cannot be geometric in the sense stated above.

Besides, from Eq. (4) we can analyze the DTR flow of the average degree as a function of the relative size of G_T with respect to the size of G. Note that it is independent of the dimension. For scale-free networks with $2 < \gamma < 3$, the average degree of subgraphs G_T grows as one goes deeper into the hierarchy of subgraphs, while for $\gamma = 3$ it remains stable, and for $\gamma > 3$ it decreases. Although other geometric models, such as growing spatial preferential attachment models [23,24], are self-similar under node birth time (i.e., selecting subgraphs of older nodes; since age and degree are directly related in these models, this is equivalent to DTR) their subgraphs have a constant average degree, which conflicts with empirical observations showing an increasing trend in the subgraphs flow (Fig. S1 of Sec. I in Supplemental Material [25]). In contrast, our general class of random geometric graphs for heterogeneous networks reproduces the empirical observations.

TABLE I. Properties of network models. Here, the $\mathbb{S}^1$ model refers to its geometric region $\beta > 1$.

Model	Pairwise	Clustering	Self-similar	Geometric
$\mathbb{S}^1$	Yes	Yes	Yes	Yes
CB	No	Yes	No	No
$MR - c_k$	No	Yes	No	No
$MR - \bar{c}$	No	Yes	No	No
CM	Yes	No	Yes	No

In what follows, we analyzed numerically the DTR self-similarity of various network models with different clustering profiles to understand their geometricity. The models considered are the configuration model (CM) [48,49], the geometric soft configuration model or $\mathbb{S}^1$ model [13] (to which the analytical proof above readily applies), the clique-based (CB) model [50], and two versions of the maximal randomization (MR) model [51], named MR-c_k and MR-$\bar{c}$, applied to a seed network with a given degree distribution. The CM generates graphs with a fixed degree distribution and no clustering in the thermodynamic limit. The other models also have a fixed degree distribution but produce graphs with finite clustering in the thermodynamic limit. The $\mathbb{S}^1$ model maximizes entropy while maintaining a given degree distribution and finite clustering, controlled through parameter β [52], which vanishes slowly for $\beta \leq 1$ and remains constant and increasing with β for $\beta > 1$ [53]. The CB model achieves clustering by producing disjoint cliques connected by extra links among those cliques. The MR-c_k and the MR-$\bar{c}$ tune the clustering spectrum and average clustering coefficient, respectively, by rewiring network connections to achieve target clustering levels. The CM and $\mathbb{S}^1$ model are pairwise models where links are formed following independent Bernoulli trials, while the CB, MR-c_k, and MR-$\bar{c}$ models are not. See Table I and Sec. II in Supplemental Material [25] for more details on the models.

We generated synthetic networks with the described models and studied the change of the structural properties along the DTR flow of subgraphs (see Fig. 2). The same degree distribution was used across models so that, for consistency, the threshold values k_T could be set to predefined levels, $\{2, 5, 10, 20, 50\}$. For each subgraph, we measured the degree distribution, the average nearest-neighbors degree, and the local clustering coefficient as functions of the degrees rescaled by the average degree of the corresponding subgraph $k_{\text{res}} \equiv k / \langle k(k_T) \rangle$. Multiscale self-similarity is denoted by overlapping curves for the different subgraphs in the DTR flow, while spread curves denote scale-dependent behavior.

In Fig. 2(a), the flow of the clustering spectrum (local clustering coefficient averaged over nodes with equal degree) is shown for each model. The $\mathbb{S}^1$ model and CM display self-similarity, both MR models are clearly

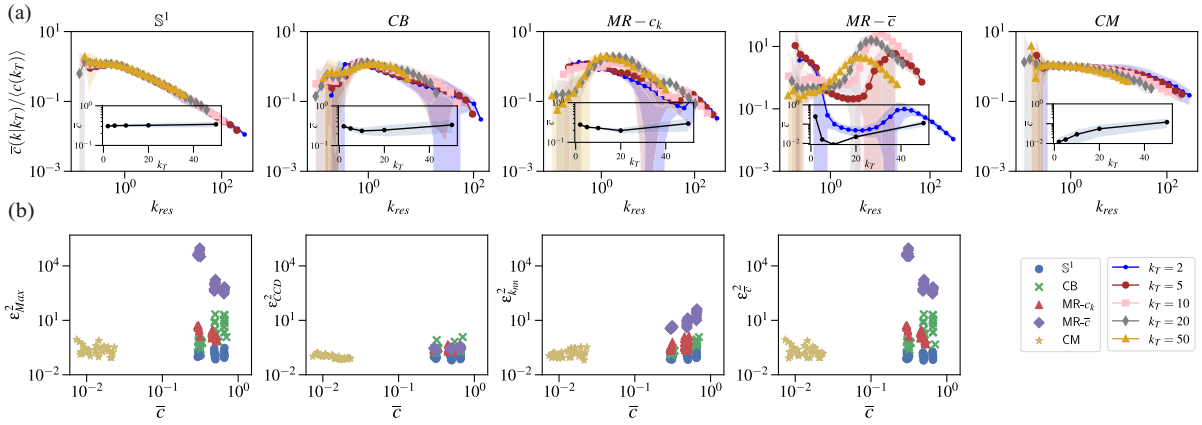

FIG. 2. Geometricity of network models. (a) Clustering spectrum of DTR subgraphs, $\bar{c}(k|k_T)$, rescaled by the average local clustering coefficient of the corresponding subgraph as a function of rescaled degree for original networks with $\langle k\rangle = 10$, $N = 50,000$. We used $\beta = 1.5$ and $\gamma = 2.5$ for the $\mathbb{S}^1$ model and used the generated synthetic $\mathbb{S}^1$ networks as seeds for the other models. The insets illustrate the average clustering coefficient as a function of k_T. The shaded areas represent two-σ intervals around the mean, calculated from 10 realizations of each model. (b) Collapse-quality metrics of structural properties as a function of the average clustering coefficient of the synthetic networks with $\gamma = 2.5$ and $\beta = \{1.5, 2, 3\}$.

non-self-similar, and the CB model shows an intermediate behavior. According to our proof, the lack of self-similarity of MR-c_k and MR-$\bar{c}$ models implies that they lack a latent geometry in the terms assumed in our framework, even if they generate high clustering. Interestingly, the peculiar behavior of MR-$\bar{c}$ is indicative of a core-periphery structure and double-percolation phenomena explored in [54]. An extended report of the flow of the topological properties of synthetic networks with additional values for the parameters β and γ are in Figs. S2–S7 of Sec. III in Supplemental Material [25].

To assess quantitatively the level of self-similarity we employed the ϵ^2 metric, a collapse-quality metric computed by averaging the squared relative differences between the curves of the original network and the ones in the corresponding subgraphs, in the range of rescaled degrees common to both. For more details, see the algorithm of Sec. IV in Supplemental Material [25]. The metric was calculated for each of the three topological properties analyzed in this Letter, and the maximum $\epsilon^2_{\mathrm{Max}} = Max(\epsilon^2_{CCD}, \epsilon^2_{knn}, \epsilon^2_{\bar{c}})$ was used as a measure of self-similarity.

Results are shown in Fig. 2(b), where we plot the ϵ^2 of each network against its average clustering coefficient (see also Fig. S8 in Supplemental Material [25]). By definition, ϵ^2 is a discriminative indicator for assessing the level of network self-similarity. As expected, the $\mathbb{S}^1$ model exhibits very small ϵ^2. The CM has a similar ϵ^2 value, and thus self-similarity level, but a low clustering coefficient from finite-size effects. The CB and MR-c_k models show appreciably higher values of ϵ^2. However, as clustering increases, $\epsilon^2_{\bar{c}}$ tends to increase for the CB model and decrease for the MR-c_k model. Finally, the MR-$\bar{c}$ model shows very high ϵ^2 values, incompatible with self-similarity. Our conclusions about the self-similarity and geometricity of the models are summarized in Table I.

We also analyzed real-world networks using the same methodology. Their description and structural properties are reported in Sec. VI and Table SI in Supplemental Material [25]. Because of finite-size effects and varying degree distributions, values for k_T were chosen in $[2, k_T^{\max}]$ for every network in enough number to ensure adequate resolution, where the maximum $k_T^{\max}$ was determined by the maximum value of the average degree in the DTR flow (Fig. S1 in Supplemental Material [25]). In Fig. 3(a), we display the DTR flow of the structural properties of two real networks (see Sec. I in Supplemental Material [25] for the rest). While the curves of Fb-Friends (top row) clearly overlap, suggesting self-similarity of the renormalized networks, PPI-rat (bottom row) displays more variation across the different subgraphs. A scatter plot of all the real networks in the $\epsilon^2_{\mathrm{Max}} - \bar{c}$ space is shown in Fig. 3(b). In general, highly clustered real networks display low values of the collapse-quality metric, indicating a substantial level of self-similarity, with more variability of self-similarity profiles in the low clustering region.

For consistency, we calculated the parameter β of the real networks to quantify the coupling between their topology and the hyperbolic two-dimensional latent space of the $\mathbb{S}^1$ model. For the inference, we used an extended version of the embedding tool Mercator [53], which adjusts the coordinates of nodes in the latent space and infers the parameters such that the observed topology of the network is best reproduced by synthetic networks of the $\mathbb{S}^1$ model. The measures (Table SI in Supplemental Material [25]) reveal a negative correlation between the value of β and $\epsilon^2_{\mathrm{Max}}$, consistent with results in Fig. 3(b). Networks with $\beta > 1$ (in the geometric regime of the $\mathbb{S}^1$ model) have low

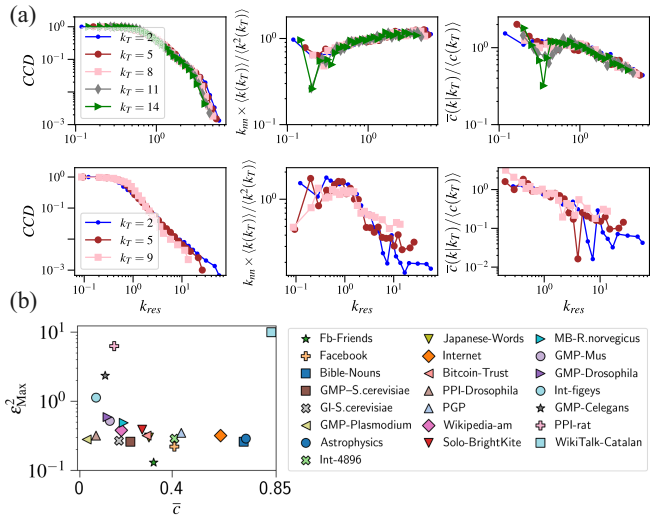

FIG. 3. Geometricity of real networks. (a) Complementary cumulative degree distribution, average nearest-neighbors degree, and clustering spectrum as functions of the rescaled degree for the DTR flow of a more self-similar network at the top, Fb-Friends, and a less self-similar network at the bottom, PPI-rat. Insets show the average clustering vs k_T. (b) Scatter plot of the collapse-quality metric against the average clustering coefficient where each dot represents a real network.

values of $\epsilon^2_{\mathrm{Max}}$ and higher clustering. Yet, in the region $\beta < 1$ (corresponding to the quasi-geometric and nongeometric regimes of the model with vanishing clustering in the thermodynamic limit) there are some networks with low values of $\epsilon^2_{\mathrm{Max}}$ and some with very high values.

In conclusion, we used DTR to elucidate the relation between local clustering, geometry, and self-similarity. We demonstrated that any geometric random graph in a Riemannian manifold with constant (zero, positive, or negative) curvature and of any dimension is self-similar under DTR, when nodes have a power-law degree distribution and the interactions are pairwise. Therefore, in this framework non-self-similar networks cannot be geometric. While geometricity always implies both local clustering and self-similarity, neither alone nor together are sufficient to imply geometricity.

We found that some network models with finite local clustering are non-self-similar under DTR and must, therefore, be nongeometric. These findings seem to contradict [11]. However, this contradiction is only apparent, highlighting that local clustering implies latent geometry only when interactions are pairwise independent. Correlated links can result in finite local clustering without inherent latent geometricity. Moreover, the DTR self-similarity of real networks can be quantified and, together with local clustering, used to indicate their potential geometricity.

Our work clarifies the necessary conditions, particularly multiscale self-similarity, under which ensemble equivalence between graphs and continuous spaces holds. When these conditions are satisfied, downstream tasks based on network maps become meaningful and tractable. The implications are broad and cross-disciplinary. A geometric framework is natural for systems where connection likelihood depends on similarity and is common in real-world networks ranging from social systems to critical infrastructures. It also enables the adaptation of methodologies from the study of continuous physical systems, as in statistical physics, and strengthens the connection between network science and fields where space is fundamental, including quantum gravity and spacetime modeling, and machine learning through graph embedding and manifold learning. These applications require a reliable network geometry framework that extends the combinatorial foundations of network science by integrating geometric concepts that capture similarity and hierarchy. Our results take a step in this direction, helping establish the theoretical basis for a solid geometric theory of complex networks.

Acknowledgments—We acknowledge support from: Grant No. TED2021-129791B-I00 funded by MCIN/AEI/10.13039/501100011033 and by "European Union NextGenerationEU/PRTR"; Grant PID2022-137505NB-C22 funded by MCIN/AEI/10.13039/501100011033 and by "ERDF/EU"; Generalitat de Catalunya Grant No. 2021SGR00856. M. B. acknowledges the ICREA Academia award, funded by the Generalitat de Catalunya.

Data availability—The data for the networks used in this Letter are publicly available in the cited references. The data supporting the figures are not publicly available but can be provided by the authors upon reasonable request.

[1] M. Boguñá, I. Bonamassa, M. D. Domenico, S. Havlin, and D. Krioukov, and M. Á. Serrano, Nat. Rev. Phys. **3**, 114 (2021).
[2] Z. Wu, G. Menichetti, C. Rahmede, and G. Bianconi, Sci. Rep. **5**, 10073 (2015).
[3] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin, Phys. Rev. Lett. **59**, 521 (1987).
[4] S. Surya, Living Rev. Relativity **22**, 5 (2019).
[5] D. Krioukov, M. Kitsak, R. S. Sinkovits, D. Rideout, D. Meyer, and M. Boguñá, Sci. Rep. **2**, 793 (2012).
[6] M. Penrose, *Random Geometric Graphs* (Oxford University Press, New York, 2003).
[7] M. A. Serrano and M. Boguñá, *The Shortest Path to Network Geometry: A Practical Guide to Basic Models and Applications*, Elements in Structure and Dynamics of Complex Networks (Cambridge University Press, Cambridge, England, 2022).
[8] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguñá, Phys. Rev. E **82**, 036106 (2010).
[9] E. Candellero and N. Fountoulakis, Internet Math. **12**, 2 (2016).
[10] M. E. Newman, Phys. Rev. E **68**, 026121 (2003).
[11] D. Krioukov, Phys. Rev. Lett. **116**, 208302 (2016).

[12] R. Michielan, N. Litvak, and C. Stegehuis, Phys. Rev. E **106**, 054303 (2022).
[13] M. A. Serrano, D. Krioukov, and M. Boguñá, Phys. Rev. Lett. **100**, 078701 (2008).
[14] C. Song, S. Havlin, and H. A. Makse, Nature (London) **433**, 392 (2005).
[15] C. Song, S. Havlin, and H. A. Makse, Nat. Phys. **2**, 275 (2006).
[16] J. S. Kim, K.-I. Goh, B. Kahng, and D. Kim, New J. Phys. **9**, 177 (2007).
[17] J. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani, Networks Heterogen. Media **3**, 371 (2008).
[18] G. García-Pérez, M. Boguñá, and M. Á. Serrano, Nat. Phys. **14**, 583 (2018).
[19] R. Jankowski, A. Allard, M. Boguñá, and M. Á. Serrano, Nat. Commun. **14**, 7585 (2023).
[20] G. Budel, M. Kitsak, R. Aldecoa, K. Zuev, and D. Krioukov, Phys. Rev. E **109**, 054131 (2024).
[21] Because of the compact nature of the spherical geometry, a very weak dependence on the system size in the connection probability is needed to obtain sparse networks. Notice that this dependence goes away in the thermodynamic limit, in which case the model becomes equivalent to the Euclidean one.
[22] Notice that $\delta_T/\delta = \lim_{N\to\infty} N_T/N$, where N and N_T are the sizes of the original network and the subgraph. Thus, for real networks it is possible to evaluate the ratio δ_T/δ as N_T/N.
[23] F. Papadopoulos, M. Kitsak, M. Á. Serrano, M. Boguñá, and D. Krioukov, Nature (London) **489**, 537 (2012).
[24] E. Jacob and P. Mörters, Ann. Appl. Probab. **25**, 632 (2015).
[25] See Supplemental Material at http://link.aps.org/supplemental/10.1103/ycwq-92ms, which includes network dataset descriptions, supplemental information, figures, tables, and Refs. [26–47].
[26] L. Gugelmann, K. Panagiotou, and U. Peter, in *Automata, Languages, and Programming. ICALP 2012*, Lecture Notes in Computer Science Vol. 7392 (Springer, Berlin, Heidelberg, 2012).
[27] N. Fountoulakis, P. van der Hoorn, T. Müller, and M. Schepers, Electron. J. Probab. **26**, 1 (2021).
[28] M. A. Abdullah, N. Fountoulakis, and M. Bode, Internet Math. **1** (2017), https://www.internetmathematicsjournal.com/article/1916-modularity-of-complex-networks-models.
[29] T. Friedrich and A. Krohmer, SIAM J. Discrete Math. **32**, 1314 (2018).
[30] T. Müller and M. Staps, Adv. Appl. Probab. **51**, 358 (2019).
[31] P. Sapiezynski, A. Stopczynski, D. D. Lassen, and S. Lehmann, Sci. Data **6**, 315 (2019).
[32] C. Römhild and C. Harrison, Bible cross-references, https://www.chrisharrison.net/index.php/Visualizations/BibleViz, accessed: 2024-10-02.
[33] M. De Domenico, M. A. Porter, and A. Arenas, J. Complex Netw. **3**, 159 (2014).
[34] Y. Hu, A. Vinayagam, A. Nand, A. Comjean, V. Chung, T. Hao, S. E. Mohr, and N. Perrimon, Nucleic Acids Res. **46**, D567 (2017).
[35] M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001).
[36] M. Zitnik, R. Sosič, M. W. Feldman, and J. Leskovec, Proc. Natl. Acad. Sci. U.S.A. **116**, 4426 (2019).
[37] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon, Science **303**, 1538 (2004).
[38] K. Claffy, Y. Hyun, K. Keys, M. Fomenkov, and D. Krioukov, in *Proceedings of 2009 Cybersecurity Applications and Technology Conference for Homeland Security* (IEEE Computer Society, Los Alamitos, CA, USA, 2009), pp. 205–211.
[39] M. Boguñá, F. Papadopoulos, and D. Krioukov, Nat. Commun. **1**, 62 (2010).
[40] S. Kumar, F. Spezzano, V. Subrahmanian, and C. Faloutsos, in *Proceedings of 2016 IEEE 16th International Conference on Data Mining (ICDM)* (IEEE, New York, 2016), pp. 221–230.
[41] S. Kumar, B. Hooi, D. Makhija, M. Kumar, C. Faloutsos, and V. Subrahmanian, in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining* (ACM, New York, 2018), pp. 333–341.
[42] H.-W. Tang *et al.*, Nat. Commun. **14**, 2162 (2023).
[43] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, Phys. Rev. E **70**, 056122 (2004).
[44] J. Kunegis, in *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13 Companion (Association for Computing Machinery, New York, 2013), pp. 1343–1350.
[45] E. Cho, S. A. Myers, and J. Leskovec, in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11 (Association for Computing Machinery, New York, 2011), pp. 1082–1090.
[46] P. H. M. Huss, IET Syst. Biol. **1**, 280 (2007).
[47] R. M. Ewing *et al.*, Mol. Syst. Biol. **3**, 89 (2007).
[48] E. A. Bender and E. Canfield, J. Comb. Theory Ser. A **24**, 296 (1978).
[49] M. Molloy and B. Reed, Random Struct. Algorithms **6**, 161 (1995).
[50] J. P. Gleeson, Phys. Rev. E **80**, 036107 (2009).
[51] P. Colomer-de Simón, M. Á. Serrano, M. G. Beiró, J. I. Alvarez-Hamelin, and M. Boguñá, Sci. Rep. **3**, 2517 (2013).
[52] M. Boguñá, D. Krioukov, P. Almagro, and M. Á. Serrano, Phys. Rev. Res. **2**, 023040 (2020).
[53] J. van der Kolk, M. A. Serrano, and M. Boguñá, Commun. Phys. **5**, 245 (2022).
[54] P. Colomer-de Simón and M. Boguñá, Phys. Rev. X **4**, 041020 (2014).

End Matter

Appendix A: Proof of the self-similarity of random geometric graphs in hyperbolic geometry—Let us consider the d-dimensional hyperbolic space with radius of curvature R, $\mathbb{H}^d$, where the square of the line element is defined as

$$ds^2 = dx^2 + R^2 \sinh^2\left(\frac{x}{R}\right) d\Omega_{d-1}^2, \tag{A1}$$

and $d\Omega_{d-1}^2$ is the squared line element of a $(d-1)$-sphere of unit radius. Thus, $x \in (0, \infty)$ is the hyperbolic distance between a point at the origin of coordinates and the set of points with fixed angular coordinates on the $(d-1)$-sphere. Therefore, the volume element is

$$dV = R^{d-1} \sinh^{d-1}\left(\frac{x}{R}\right) dx d\Omega_{d-1}. \tag{A2}$$

We now define a Poisson point process on $\mathbb{H}^d$ at density δ. This implies that the expected number of points within the volume element is simply δdV. Using this Poisson point process, we define a random geometric graph in the following way. We identify each point as a node in the graph and endow each node with a hidden variable h drawn from the distribution $\rho(h)$, independently of the position of the node in $\mathbb{H}^d$. Then, a pair of nodes with hidden variables h and h', separated by the hyperbolic distance x, gets connected with probability

$$p = f\left(\frac{1}{hh'} \int_0^{\frac{x}{R}} \sinh^{d-1} z dz\right), \tag{A3}$$

where $0 \leq f(\cdot) \leq 1$ is an arbitrary but integrable function of its argument. Notice that the argument of function f is a monotonically increasing function of the hyperbolic distance. Thus, Eq. (A3) represents a general class of heterogeneous random geometric graphs with pairwise interactions in hyperbolic geometry.

Since $\mathbb{H}^d$ is a homogeneous and isotropic space, all nodes in the graph are geometrically equivalent and are only distinguished by their hidden variables h, so that there exists an isometry that maps its similarity coordinate to the origin. Let us, then, consider a node at the origin of coordinates with hidden variable h and compute its expected degree $\bar{k}(h)$ as

$$\bar{k}(h) = \delta R^{d-1} S_{d-1} \int \rho(h') dh' \int_0^\infty \sinh^{d-1}\left(\frac{x}{R}\right) \times f\left(\frac{1}{hh'} \int_0^{\frac{x}{R}} \sinh^{d-1} z\, dz\right) dx, \tag{A4}$$

where S_{d-1} is the volume of the $(d-1)$-dimensional unit sphere. By making the change of variables

$$t = \frac{1}{hh'} \int_0^{\frac{x}{R}} \sinh^{d-1} z dz, \tag{A5}$$

we obtain

$$\bar{k}(h) = \delta R^d S_{d-1} I \langle h \rangle h, \quad \text{with} \quad I \equiv \int_0^\infty f(t) dt. \tag{A6}$$

We then see that the hidden variable h is proportional to the expected degree of the node—thus, satisfying LS—so the degree distribution is determined by the distribution $\rho(h)$, independently of the functional form of the connection probability. The average degree is then

$$\langle k \rangle = \delta R^d S_{d-1} I \langle h \rangle^2. \tag{A7}$$

We are interested in scale-free networks. Thus, we choose h to be power-law distributed as

$$\rho(h) = (\gamma - 1) h_0^{\gamma-1} h^{-\gamma}, \quad \text{with} \quad h \geq h_0,$$

so that $\langle h \rangle = (\gamma - 1) h_0 / (\gamma - 2)$.

Now, we apply the DTR transformation to the hidden variable and decimate the network by removing all nodes with hidden variables in the range $h_0 < h < h_T$. By doing so, the remaining nodes are also power-law distributed, but with hidden variables starting at h_T instead of h_0, so that their average is

$$\langle h \rangle_T = \frac{h_T}{h_0} \langle h \rangle. \tag{A8}$$

Furthermore, the relative size of the network after decimation compared to the original network is simply given by

$$\lim_{N \to \infty} \frac{N_T}{N} = \frac{\delta_T}{\delta} = \left(\frac{h_0}{h_T}\right)^{\gamma-1}. \tag{A9}$$

Therefore, the subgraph obtained after decimation is a Poisson point process on $\mathbb{H}^d$ at density δ_T and with hidden variables power-law distributed with exponent γ and average $\langle h \rangle_T$. The subgraph is then a realization of the same ensemble as the original network, with the only difference being that the average degree is given by

$$\langle k \rangle_T = \langle k \rangle \left(\frac{N}{N_T}\right)^{\frac{3-\gamma}{\gamma-1}}, \tag{A10}$$

just as in the Euclidean case.

Appendix B: Proof of the self-similarity of random geometric graphs in spherical geometry—We now turn to geometries with constant positive curvature. Let us consider the d-dimensional spherical space with radius of curvature R, $\mathbb{S}^d$, where the square of the line element is

$$ds^2 = dx^2 + R^2 \sin^2\left(\frac{x}{R}\right) d\Omega_{d-1}^2, \tag{B1}$$

and $d\Omega_{d-1}^2$ is the squared line element of a $(d-1)$-sphere of unit radius. Thus, $x \in (0, \pi R)$ is the spherical distance between the origin of coordinates and the set of points with fixed angular coordinates on the $(d-1)$-sphere. Therefore, the volume element is

$$dV = R^{d-1} \sin^{d-1}\left(\frac{x}{R}\right) dx d\Omega_{d-1}. \tag{B2}$$

A key difference between spherical and hyperbolic geometries is that the former are compact manifolds with a finite volume given by

$$V_{\text{tot}} = \frac{2\pi^{\frac{d+1}{2}}}{\Gamma\left(\frac{d+1}{2}\right)} R^d. \tag{B3}$$

Thus, if we define a Poisson point process on $\mathbb{S}^d$ at density δ, the expected number of nodes in the graph is given by

$$N = \frac{2\pi^{\frac{d+1}{2}}}{\Gamma\left(\frac{d+1}{2}\right)} \delta R^d. \tag{B4}$$

This means that, to take the thermodynamic limit, we can either let $R \gg 1$ while keeping δ constant or, alternatively, keep R fixed and let $\delta \gg 1$ or, in fact, any other combination that leads to $\delta R^d \gg 1$. As in the hyperbolic case, we endow each node with a hidden variable h drawn from the distribution $\rho(h)$, independently of the node's position in $\mathbb{S}^d$. Then, a pair of nodes with hidden variables h and h', separated by the spherical distance x, are connected with probability

$$p = f\left(\frac{\delta R^d}{hh'} \int_0^{\frac{x}{R}} \sin^{d-1} z \, dz\right), \tag{B5}$$

where $f(\cdot)$ is an arbitrary but integrable function. Notice that, contrary to the hyperbolic case, we include an explicit dependence in the connection probability on the term δR^d, which is proportional to the system size. This choice is taken so that the model defines an ensemble of sparse networks in the thermodynamic limit. Since function $f(\cdot)$ is any integrable function of its argument, Eq. (B5) represents the most general class of heterogeneous random geometric sparse graphs with pairwise interactions on $\mathbb{S}^d$.

Since $\mathbb{S}^d$ is a homogeneous and isotropic space, all nodes in the graph are geometrically equivalent and differ only by their hidden variables h, so that there exists an isometry that maps its similarity coordinate to the origin. Let us consider a node at the origin with hidden variable h and compute its expected degree $\bar{k}(h)$ as

$$\bar{k}(h) = \delta R^{d-1} S_{d-1} \int \rho(h') dh' \int_0^{\pi R} \sin^{d-1}\left(\frac{x}{R}\right) \times f\left(\frac{\delta R^d}{hh'} \int_0^{\frac{x}{R}} \sin^{d-1} z \, dz\right) dx. \tag{B6}$$

By making the change of variables

$$t = \frac{\delta R^d}{hh'} \int_0^{\frac{x}{R}} \sin^{d-1} z \, dz, \tag{B7}$$

we obtain

$$\bar{k}(h) = S_{d-1} h \int \rho(h') h' I(h, h'), \tag{B8}$$

with

$$I(h, h') \equiv \int_0^{\frac{\alpha_d \delta R^d}{hh'}} f(t) dt, \tag{B9}$$

and $\alpha_d = [\sqrt{\pi}\Gamma(d/2)/\Gamma(d+1/2)]$. In the limit of very large networks, $\delta R^d \gg 1$, and so the integral $I(h, h')$ becomes a constant. Thus, in this limit we can write

$$\bar{k}(h) \approx S_{d-1} h I \langle h \rangle, \quad \text{with} \quad I \equiv \int_0^{\infty} f(t) dt. \tag{B10}$$

Again, we see that h is simply proportional to the expected degree—thus, satisfying LS—and $\langle k \rangle = S_{d-1} I \langle h \rangle^2$. As in the hyperbolic case, we are interested in scale-free networks, so that we consider $h > h_0$ to be power-law distributed with exponent γ.

Now, we apply the DTR transformation to the hidden variable and decimate the network by removing all nodes with hidden variables in the range $h_0 < h < h_T$. By doing so, the remaining nodes also follow a power-law distribution, but with hidden variables starting at h_T instead of h_0, such that their average is given by Eq. (A8). Similarly, the relation between the size and density of the subgraph and the corresponding quantities of the original network is also given by Eq. (A9). The connection probability between

nodes in the subgraph remains the same as in the original network, thus retaining its explicit dependence on δ. However, the density of nodes in the subgraph is δ_T. Therefore, the average degree in the subgraph is given by

$$\langle k\rangle_T = \frac{\delta_T}{\delta} S_{d-1} I \langle h\rangle_T^2. \tag{B11}$$

Combining this result with Eqs. (A8) and (A9), we obtain

$$\langle k\rangle_T = \langle k\rangle \left(\frac{N}{N_T}\right)^{\frac{3-\gamma}{\gamma-1}}, \tag{B12}$$

which is, once again, the same as in the Euclidean case.