

Clustering Does Not Always Imply Latent Geometry

Roya Aliakbarisani,^{1, 2,*} Marián Boguñá,^{1, 2, †} and M. Ángeles Serrano^{1, 2, 3, ‡}

¹Departament de Física de la Matèria Condensada,

Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain

²Universitat de Barcelona Institute of Complex Systems (UBICS), Barcelona, Spain

³Institució Catalana de Recerca i Estudis Avançats (ICREA),
Passeig Lluís Companys 23, E-08010 Barcelona, Spain

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 the multiscale self-similarity of complex networks is a crucial factor in implying latent geometry. Using degree-thresholding renormalization, we prove that any random scale-free graph in a d -dimensional homogeneous and isotropic manifold is self-similar when interactions are pairwise. Hence, both clustering and self-similarity are required to imply geometricity. Our findings highlight that correlated links can lead to finite clustering without self-similarity, and therefore without inherent latent geometry. The implications are significant for network mapping and ensemble equivalence between graphs and continuous spaces.

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 clustering under a single mechanism.

The key property that connects the hyperbolic geometry of a network and its topology is clustering, the tendency of nodes to share neighbors [10], with triangles in the network structure induced as a reflection of the triangle inequality in the metric space. This raises the intriguing question of whether clustering implies geometricity. Previous work [11] suggested that clustering implies geometry by showing that random graphs with homogeneous degrees and clustering are equivalent to random geometric graphs. However, this holds only under certain conditions. Furthermore, network models without any inherent geometry, like the configuration model, can still contain large amounts of triangles. The limitations of using triangle counts to detect geometry led to the development of a weighting scheme designed to determine reliably whether 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. We used the degree-thresholding renormalization (DTR) method [13], which generates a nested hierarchy of subgraphs by progressively filtering out nodes with degree below a threshold, to reveal multiscale self-similarity. DTR self-similarity has been observed in many real-world networks and has been explained by the renormalizability of the geometric S^1 model [13]. Self-similarity of complex networks has also been explored from other perspectives [14–18].

Below, we prove that any geometric random graph in a homogeneous and isotropic manifold with any curvature and any dimension is self-similar under DTR, provided nodes have a scale-free degree distribution and pairwise interactions. Therefore, non-self-similar networks cannot be geometric. We found that some models with finite clustering are non-self-similar under DTR, thus non-geometric. Hence, while geometricity always implies clustering and self-similarity, the reverse is not true; however, the absence of clustering or self-similarity implies non-geometricity, see Fig. (1). Moreover, real networks can be classified according to their self-similarity and clustering levels, providing insights into their potential geometricity. Thus, clustering, geometry, and self-similarity interplay in a non-trivial way in complex networks.

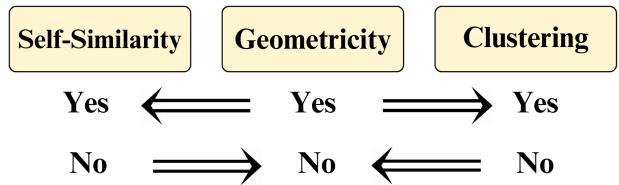


FIG. 1: Logic consequence relationships between geometricity, clustering, and self-similarity for geometric random graphs in homogeneous and isotropic spaces.

We start our proof by considering the thermodynamic

* roya.aliakbarisani@ub.edu

† marian.boguna@ub.edu

‡ marian.serrano@ub.edu

limit of a homogeneous Poisson point process with density δ in a Euclidean space of dimension d . The same is obtained for homogeneous and isotropic hyperbolic and spherical manifolds; the proofs for negative and positive curvature are provided in Appendices 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 hyperbolic formulation the \mathbb{H}^{d+1} model [19, 20], where homogeneous nodes are distributed heterogeneously in the $(d+1)$ -dimensional hyperbolic space. In contrast, the proof for hyperbolic manifolds refers to heterogeneous nodes distributed homogeneously over all the d -dimensional hyperbolic space.

Such a space conceptualizes a similarity space, in which closeness reflects how similar nodes are and determines their likelihood to form connections. Furthermore, we assume that each node has a popularity attribute, independent of its position in the 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

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

We do not impose any particular form to the function $f(\cdot)$ other than being integrable in \mathbb{R}^d . As we show below, this implies that the model generates sparse networks without the need to introduce a size dependence on the connection probability. 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,

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

where we have already integrated the angular part and S_{d-1} is the volume of the $(d-1)$ -sphere. A simple change of variables leads to $\bar{k}(h_i) = \delta I_d \langle h^d \rangle h_i^d$, where $I_d \equiv S_{d-1} \int_0^\infty f(x) x^{d-1} dx$, so that the average degree is

$$\langle k \rangle = \delta I_d \langle h^d \rangle^2. \quad (3)$$

As discussed earlier, the average degree is constant – and so the network is sparse – as long as the integral I_d is bounded. Hereafter, we constraint the distribu-

tion of hidden variables h to be scale-free with $\rho(h) = (\tilde{\gamma} - 1) h_0^{(\tilde{\gamma}-1)} h^{-\tilde{\gamma}}$, $h \geq h_0$ and $\tilde{\gamma} > d$.

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 $G \in \mathcal{G}$, 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 . Hence, $\delta_T/\delta = \lim_{N \rightarrow \infty} N_T/N$, where N and N_T are the sizes of the original network and the subgraph. Since the popularity dimension encoded in the hidden variable h is uncorrelated with the position in the metric space, the ensemble \mathcal{G}_T is the same as \mathcal{G} in the thermodynamic limit, with the only difference that the density of the Poisson point process δ_T and the average $\langle h^d \rangle_T$ are rescaled. Hence, the ensemble \mathcal{G} is statistically invariant under DTR and it 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

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

where we have used

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

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 .

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 . Notice 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, independently of the dimension.

Hence, we have demonstrated that a geometric random graph with pairwise interactions and a scale-free 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.

We analyzed the DTR self-similarity of various network models with different clustering profiles to understand their geometricity. The models considered are the Configuration Model (CM) [21, 22], the geometric soft configuration model or \mathbb{S}^1 model [13], the clique-

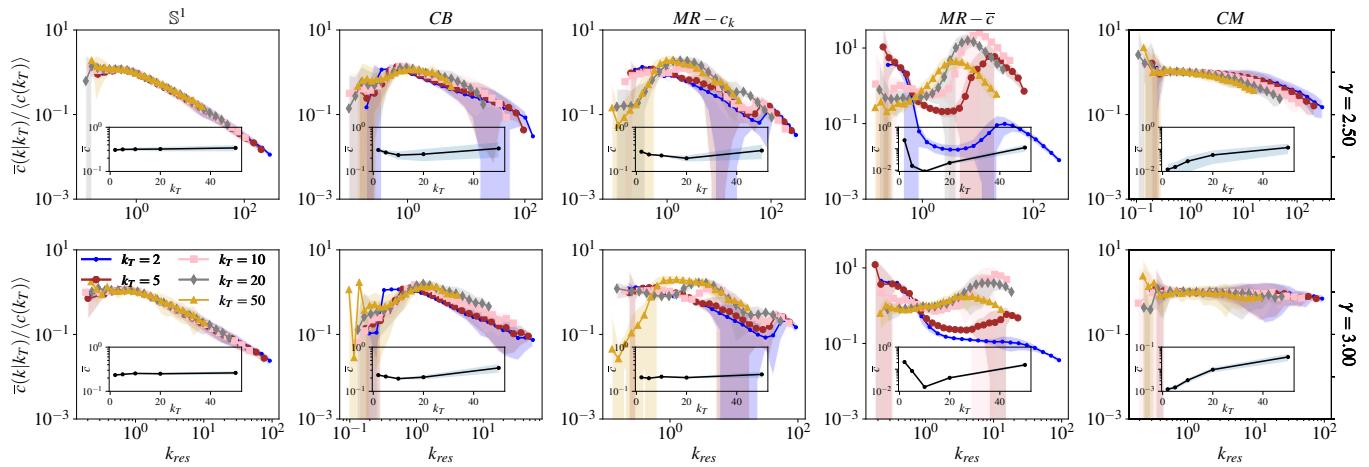


FIG. 2: Rescaled clustering coefficient as a function of rescaled degree in networks with $\beta = 1.5$, $\langle k \rangle = 10$ and $N = 50,000$. The insets illustrate the average clustering coefficient as a function of k_T . The shaded areas represent two- σ intervals around the mean, calculated from 10 realizations of each model. The values of ϵ_c^2 for each model are: 0.08 for S^1 , 0.87 for the CB model, 6.05 for $MR - c_k$, 5.76 for $MR - \bar{c}$, and 0.16 for the CM model.

based model (CB) [23], and two versions of the maximal randomization model (MR) [24], 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 large networks. The other models also have a fixed degree distribution but produce graphs with finite clustering in the thermodynamic limit. The S^1 model maximizes entropy while maintaining a given degree distribution and clustering level [25]. The CB achieves clustering by producing disjoint cliques connected by extra links. 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 S^1 are pairwise, meaning the connection likelihood between two nodes is independent of other nodes, while the CB, $MR - c_k$, and $MR - \bar{c}$ are not. See Table I, and more details on the models in Appendix C.

We generated synthetic networks with the described models and studied the change of their structural properties in the DTR flow. For each network, we applied the degree thresholding method to produce a nested hierarchy of subgraphs. Since all networks had the same degree distributions, the threshold values k_T were fixed to predefined levels, $\{2, 5, 10, 20, 50\}$, ensuring consistency in the analysis across different network models. For each subgraph, we measured the degree distribution, the average nearest-neighbors degree, and the clustering coefficient as functions of the degrees rescaled by the average degree of the corresponding subgraph $k_{res} = k / \langle k(k_T) \rangle$. Multi-scale self-similarity is denoted by overlapping curves for the different subgraphs in the DTR flow, while spread curves denote scale-dependent behavior.

Fig. 2 shows the behavior of the clustering coefficient. The S^1 and CM models display self-similarity, both MR models and the CM are clearly non-self-similar, and the

CB model shows an intermediate behavior that could be named quasi-self-similar. According to our proof, the lack of self-similarity of $MR - c_k$ and $MR - \bar{c}$ implies that they lack a latent geometry, 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 [26]. A detailed analysis of synthetic networks for additional parameter values for β and γ are in Appendix D, (Figs. 8, 9), where we also report other structural properties, including the complementary cumulative degree distribution (Figs. 4, 5) and the average nearest-neighbor degree function (Figs 6, 7) which display congruent behaviors. Our conclusions about the self-similarity and geometricity of the models are summarized in Table I.

TABLE I: Properties of network models.

Model	Pairwise	Clustering	Self-Similarity	Geometric
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

We also analyzed real-world networks using the same methodology. Their description and structural properties are reported in Appendix F and Table III. Due to 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 taking the maximum value of the average degree in the DTR flow (see Fig. 10 in Appendix E). In Fig. 3-(a), we display the DTR flow of the structural properties of two real networks. The curves of Fb-Friends (top row) overlap, suggesting similar structure of the renormalized networks. In contrast,

for PPI-rat (bottom row), the structural properties vary across different subgraphs. In Appendix G, we show the structural properties of all the real-world networks analyzed here, classified as self-similar (Figs. 11, 12, 13), quasi-self-similar (Fig. 14) and non-self-similar (Fig. 15).

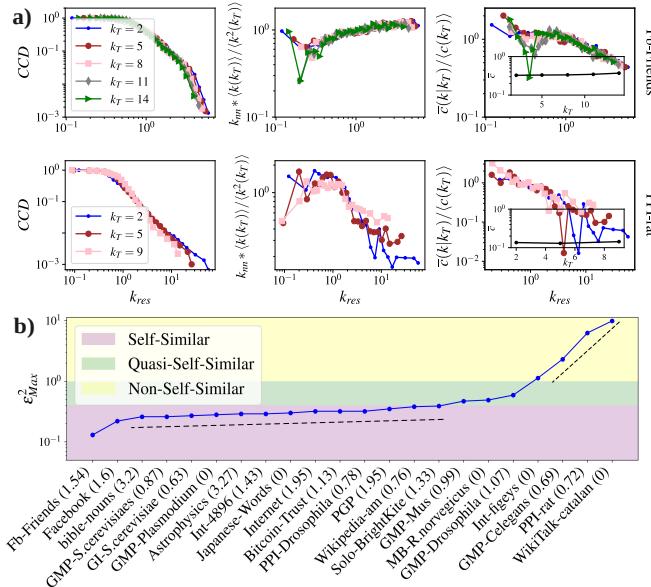


FIG. 3: Geometricity of real networks. (a) Complementary cumulative degree distribution, average nearest-neighbors' degree, and clustering spectrum as functions of the rescaled degree for a self-similar network (top) and a non-self-similar (bottom) network. (b)

Functions show the average clustering vs. k_T . (b) Distribution of ϵ_{Max}^2 values for real networks, sorted in ascending order. The three classification regions are: self-similar (purple, $\epsilon_{\text{Max}}^2 < 0.4$), quasi-self-similar (green, $0.4 \leq \epsilon_{\text{Max}}^2 < 1$), and non-self-similar (yellow, $\epsilon_{\text{Max}}^2 \geq 1$). Black dashed lines highlight trends. The value of β of every network is displayed in parenthesis.

Fig. 3-(b) displays the classification of real networks as self-similar, quasi-self-similar, or non-self-similar, see also the Sankey diagram in Fig. 16 of Appendix G. This classification is based on an ϵ^2 -Test, which calculates the maximum difference ϵ_{Max}^2 between the curves of the original network and its subgraphs in the range of rescaled degrees common to both, computed as $\epsilon_{\text{Max}}^2 = \text{Max}(\epsilon_{CCD}^2, \epsilon_{knn}^2, \epsilon_c^2)$. The values for the three structural properties were obtained by averaging the squared relative differences between the corresponding curves of the original network and the subgraphs. For more details, refer to the algorithm in Appendix E. In Fig. 3-(b), networks are arranged in ascending order along the x-axis based by increasing ϵ_{Max}^2 . The curve shows three distinct regions: low values represent self-similar networks, which exhibit a gradual increase with a very smooth slope, a moderate increase marks quasi-self-similar networks, and a sharp rise indicates non-self-similar networks. Based on this, networks are categorized as self-

similar ($\epsilon_{\text{Max}}^2 < 0.4$), quasi-self-similar ($0.4 \leq \epsilon_{\text{Max}}^2 < 1$), and non-self-similar ($\epsilon_{\text{Max}}^2 \geq 1$). According to these bounds, Fb-Friends ($\epsilon_{\text{Max}}^2 = 0.12$) is self-similar, while PPI-rat ($\epsilon_{\text{Max}}^2 = 6.27$) is non-self-similar. This scheme is also applicable to synthetic networks, correctly classifying S^1 and CM as self-similar, CB as quasi-self-similar, and the two versions of the MR -Model as non-self-similar, see values reported in the caption of Fig. 2.

For a consistency test, we calculated the level of coupling between the topology of the networks and the hyperbolic two-dimensional latent space of the S^1 model. This coupling is quantified by the parameter β , which controls the level of clustering in the model and can be estimated using an extended version of the Mercator embedding tool capable of handling both strongly and weakly geometric networks [27]. The embedding 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 model.

The β values for real networks, shown in Fig. 3-(b) and Fig. 16, indicate that the absence of DTR self-similarity implies non-geometricity. The results reveal that non-self-similar networks have $\beta < 1$, categorizing them as either non-geometric ($\beta \lesssim 0.5$) or quasi-geometric ($0.5 \lesssim \beta < 1$). Geometric networks ($\beta > 1$) are self-similar, except for GMP-Drosophila, which is quasi-self-similar with β near the critical value $\beta_c = 1$. The average β values of self-similar, quasi-self-similar, and non-self-similar networks are 1.36, 0.69, and 0.35, respectively. Table III in Appendix G shows that geometric networks with $\beta > 1$ have high clustering coefficients. However, as noted in [27], networks in the quasi-geometric and non-geometric regions, where topology couples weakly with the underlying metric space, exhibit clustering coefficients that tends to zero in the thermodynamic limit. Despite this, clustering decay is slow, especially in the quasi-geometric regime, so finite systems may still retain significant clustering. In conclusion, networks that fit a geometric description are self-similar and highly clustered. However, self-similar networks are not always geometric, and non-self-similar networks tend to be weakly geometric or incompatible with a geometric description.

To conclude, we used DTR to elucidate the relation between clustering, geometry, and self-similarity. We demonstrated that any geometric random graph in a homogeneous and isotropic manifold (with zero, positive, or negative curvature) of any dimension is self-similar under DTR, when nodes have a scale-free degree distribution and the interactions are pairwise. Therefore, non-self-similar networks cannot be geometric. While geometricity always implies both clustering and self-similarity, neither alone is sufficient to imply geometricity. We found that some network models with finite clustering are non-self-similar under DTR and must therefore be non-geometric. These findings seem to contradict [11]. However, this contradiction is only apparent, highlighting that clustering implies latent geometry only when

interactions are pairwise independent. Correlated links can result in finite clustering without inherent latent geometricity. Moreover, real networks can be classified by their self-similarity and clustering, which allows us assess their geometricity when both properties are present.

Our work unveils how the existence of multiscale symmetries in networks helps in delimiting network geometry. If the conditions are met, a myriad of downstream tasks based on network maps becomes possible and meaningful for networks. The implications are far-reaching and important for any problem related to ensemble equivalence between graphs and continuous spaces.

ACKNOWLEDGMENTS

We acknowledge support from: Grant 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 A way of making Europe”; Generalitat de Catalunya grant number 2021SGR00856. M. B. acknowledges the ICREA Academia award, funded by the Generalitat de Catalunya.

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

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

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

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

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

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

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 the most 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 . Let us then consider a node at the origin of coordinates with hidden variable h and compute its expected degree $\bar{k}(h)$ as

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

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

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

we obtain

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

We then see that the hidden variable h is proportional to the expected degree of the node, so the degree distribution is determined by the distribution $\rho(h)$. The average degree is then

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

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

$$\langle h \rangle_T = \frac{h_T}{h_0} \langle h \rangle. \quad (\text{A8})$$

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

$$\frac{N_T}{N} = \frac{\delta_T}{\delta} = \left(\frac{h_0}{h_T} \right)^{\gamma-1}. \quad (\text{A9})$$

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

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

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

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

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

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

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

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

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

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

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

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

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 . Let us then consider a node at the origin with hidden variable h and compute its expected

degree $\bar{k}(h)$ as

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

By making the change of variables

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

we obtain

$$\bar{k}(h) = S_{d-1} h \int \rho(h') h' I(h, h'), \quad \text{with } I(h, h') \equiv \int_0^{\frac{\alpha_d \delta R^d}{hh'}} f(t) dt, \quad (\text{B8})$$

and $\alpha_d =$. 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

$$\bar{k}(h) \approx S_{d-1} h I \langle h \rangle, \quad \text{with } I \equiv \int_0^\infty f(t) dt. \quad (\text{B9})$$

Again, we see that h is simply proportional to the expected degree 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

$$\langle k \rangle_T = \frac{\delta_T}{\delta} S_{d-1} I \langle h \rangle_T^2. \quad (\text{B10})$$

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

$$\langle k \rangle_T = \langle k \rangle \left(\frac{N}{N_T} \right)^{\frac{3-\gamma}{\gamma-1}}, \quad (\text{B11})$$

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

Appendix C: Random Graph Models:

1. Clique-based Model (CB):

The clique-based model (CB) [23], also referred to as the Gleeson model, is a network model designed to replicate both the degree distribution and clustering properties commonly found in real-world networks. The CB model is governed by a joint probability distribution, $\gamma(k, c)$, which indicates the probability that a randomly chosen node has degree k and belongs to a clique of size c , with $k \geq c - 1$. Here, $k - c + 1$ represents the number of external links that are connected to nodes outside the clique. After assigning each node to a single c -clique, nodes within each clique form super-nodes, and their external link stubs are randomly connected together using the configuration model. This approach allows for both local clustering and long-range interactions across the network.

The CB model can closely approximate the structural characteristics of real-world networks. However, as it starts from cliques rather than individual nodes, the number of nodes and their degrees may not be strictly predetermined, leading to small discrepancies between the expected and actual properties in finite networks. Despite these variations, the CB model remains a valuable tool for simulating networks with high clustering and heterogeneous degree distributions.

2. Maximally Random model (MR model):

The Maximally Random model (MR) [24] is a network model designed to realistically capture clustering while imposing minimal constraints. The MR model preserves the degree distribution and clustering coefficient for each degree level $MR - c_k$ or the average clustering coefficient of the entire network $MR - \bar{c}$ by employing an optimization approach based on a specific Hamiltonian. This Hamiltonian function measures the difference between the model's clustering and a target clustering coefficient, either at each degree level or across the whole network, allowing the model to control clustering levels while maintaining network randomness. A Metropolis-Hastings rewiring process and simulated annealing are used to minimize clustering deviations while keeping the degree sequence fixed. Consequently, the MR model generates clustered networks that closely resemble real-world networks in terms of clustering levels without enforcing an artificially ordered or modular structure.

3. S^1 model:

The S^1 model [7, 13] is a geometric model in a latent similarity metric space that generates synthetic networks with realistic properties, including high clustering [8, 25, 28–30], heterogeneous degree distributions [8, 13, 28], self-similarity [13], small-worldness [31–33], among others. In the S^1 model, each node is characterized by two hidden variables, (κ, θ) . The parameter κ indicates the expected degree of the node, reflecting its popularity, while θ represents its angular position on a one-dimensional sphere, which abstracts the similarity space, with the radius proportional to the total number of nodes. The connection probability between two nodes follows a gravity-like law: nodes that are farther apart in similarity space are less likely to connect, whereas nodes with higher popularity are more likely to form connections. The parameter β , which is the inverse of temperature, controls the level of clustering. At the critical value $\beta_c = 1$, the system transitions from a high-clustering regime at high β to a low-clustering regime at low β [13]. It also regulates the coupling strength between the similarity space and the topology through the triangle inequality.

Appendix D: Experimental Analysis of synthetic networks

In the following, we analyze the degree-thresholding renormalization (DTR) self-similarity of network models to assess their geometricity. Synthetic networks were initially generated using the S^1 model with a total of $N = 50000$ nodes and varying parameter values for β and γ , while keeping the average degree fixed at $\langle k \rangle = 10$. These networks were subsequently randomized using random graph models, including the *CB*, *MR*, and *CM* models.

Figs. 4 and 5 show the complementary cumulative degree distribution as a function of the rescaled degree k_{res} in DTR flows for synthetic networks generated by the aforementioned models, with β set to 1.5 and 3, respectively. Figs. 6 and 7 highlight the average nearest-neighbor degree as a function of rescaled degree. Finally, Figs. 8 and 9 present the clustering spectrum as a function of rescaled degree, with the insets showing the average clustering coefficient as a function of k_T . It is evident from these figures, particularly Figs. 8 and 9, that the S^1 and *CM* models exhibit self-similarity, the *CB* model demonstrates quasi-self-similarity, and both the $MR - c_k$ and $MR - \bar{c}$ models are non-self-similar.

Appendix E: ϵ^2 -test

In order to classify networks as self-similar, quasi-self-similar, or non-self-similar, we measure variations in their structural properties in the DTR flows. In DTR, nodes with degree $k \leq k_T$ are removed from the given network to generate a hierarchy of its subgraphs. We then quantify the variations in the fundamental structural properties in the DTR flows, including the complementary cumulative degree distribution (CCD), the average nearest-neighbor degree function (k_{nn}), and the clustering spectrum ($\bar{c}(k)$).

Algorithm 1 outlines the steps of the ϵ^2 -test, which computes ϵ^2 values as the average squared relative difference between the curves representing structural properties of the original and renormalized networks. Based on the maximum value of ϵ^2 (ϵ_{Max}^2), the test categorizes networks into three types: networks with smaller ϵ_{Max}^2 are classified as self-similar, moderate values indicate quasi-self-similarity, and larger values suggest non-self-similarity. The threshold values for this classification networks are determined by analyzing ϵ_{Max}^2 values computed for a variety of real-world networks, as shown in Fig. 3-(b). These values are sorted in ascending order and plotted, forming a curve that reveals three distinct regions: low values represent self-similar networks, which exhibit a gradual increase with a very smooth slope, a moderate increase marks quasi-self-similar networks, and a sharp rise indicates non-self-similar networks.

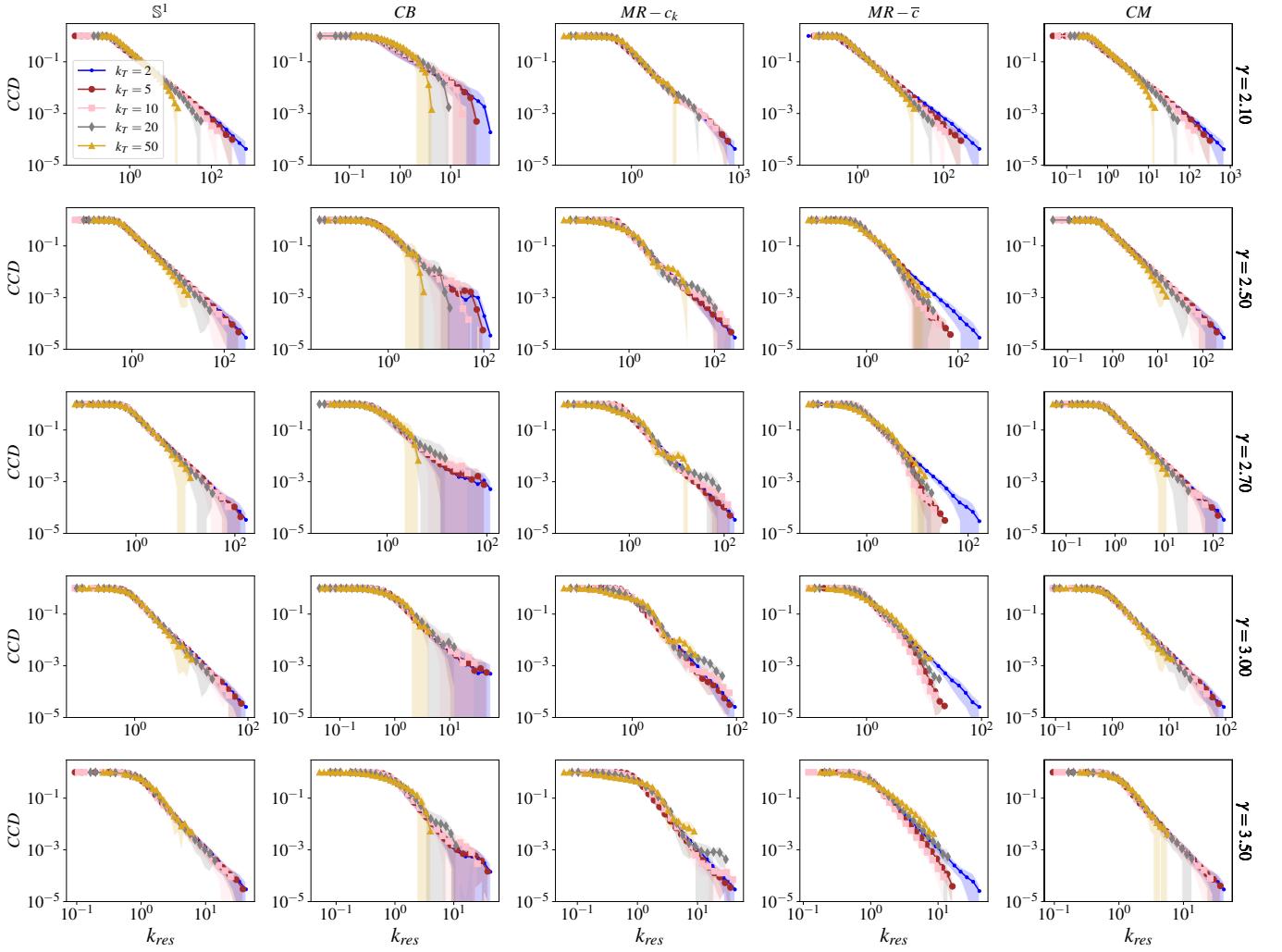


FIG. 4: The complementary cumulative degree distribution of the synthetic networks in the DTR flow as a function of the degree rescaled by the average degree in the corresponding subgraph, $k_{\text{res}} = k/\langle k(k_T) \rangle$. Each column corresponds to synthetic networks generated by a model with $N = 50000$ nodes, $\langle k \rangle = 10$, $\beta = 1.5$ and varying values of γ . The shaded areas represent two- σ intervals around the mean, calculated from 10 realizations of each model.

Based on this, we select $\epsilon_{Max}^2 < 0.4$ for self-similar networks, $0.4 \leq \epsilon_{Max}^2 < 1$ for quasi-self-similar networks, and $\epsilon_{Max}^2 \geq 1$ for non-self-similar networks.

Here, we consider the subgraph corresponding to $k_T = 2$ as the original network. Since low-degree nodes often introduce unwanted fluctuations in structural properties, removing them allows the analysis to focus on the core topological characteristics of the network.

Fig. 10 illustrates the normalized average degree of subgraphs, $\langle k(k_T) \rangle / \langle k \rangle$, as a function of k_T in the DTR flows for real networks. The peak of this function is used to determine the k_T^{max} value in real networks, accounting for finite-size effects and diverse degree distributions. Moreover, Table II highlights ϵ^2 values computed for different structural properties in real networks.

In the experiments with real networks, the maximum number of intervals in $[2, k_T^{max}]$ is set to $n^{\text{max}} = 5$, the minimum number to $n^{\text{min}} = 2$ and the minimum gap between k_T values is $\Delta k^{\text{min}} = 3$. The number of bins for the exponential binning is $n_{\text{bins}} = 20$ for both synthetic and real networks.

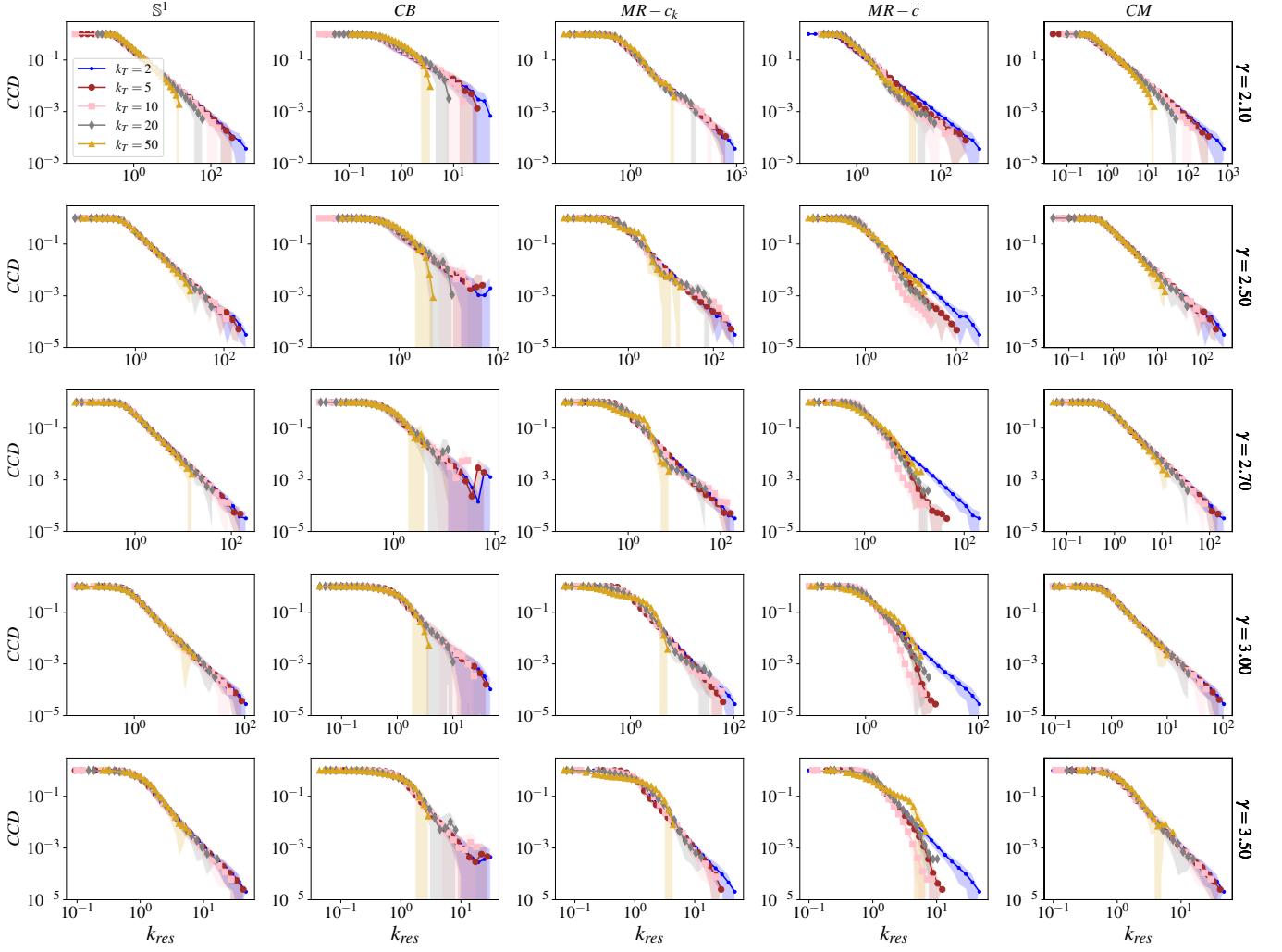


FIG. 5: The complementary cumulative degree distribution of the synthetic networks in the DTR flow as a function of the degree rescaled by the average degree in the corresponding subgraph, $k_{\text{res}} = k/\langle k(k_T) \rangle$. Each column corresponds to synthetic networks generated by a model with $N = 50000$ nodes, $\langle k \rangle = 10$, $\beta = 3$ and varying values of γ . The shaded areas represent two- σ intervals around the mean, calculated from 10 realizations of each model.

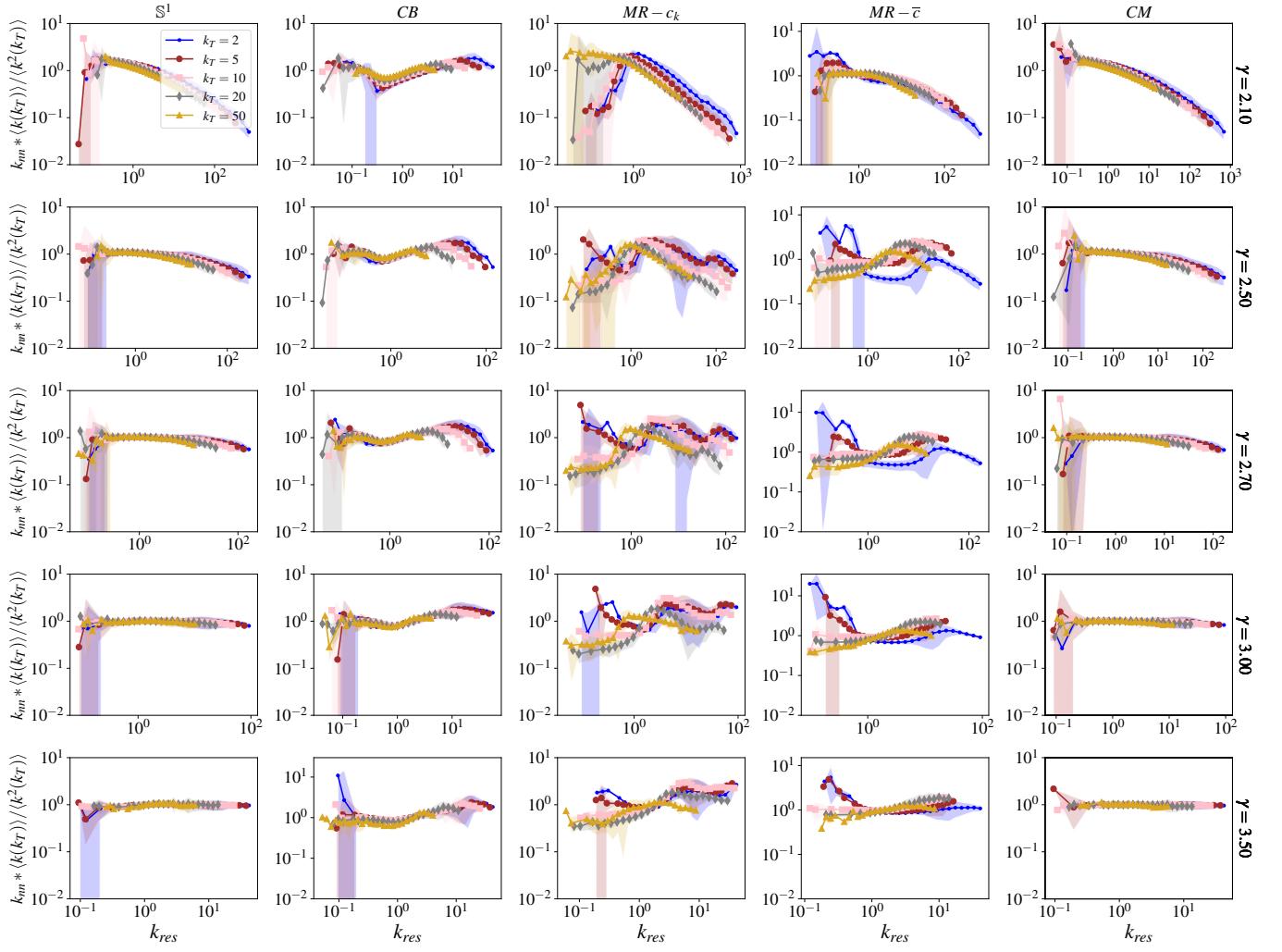


FIG. 6: Average nearest-neighbor degree of the synthetic networks in the DTR flow as a function of the degree rescaled by the average degree in the corresponding subgraph, $k_{\text{res}} = k / \langle k(k_T) \rangle$. Each column corresponds to synthetic networks generated by a model with $N = 50000$ nodes, $\langle k \rangle = 10$, $\beta = 1.5$ and varying values of γ . The shaded areas represent two- σ intervals around the mean, calculated from 10 realizations of each model.

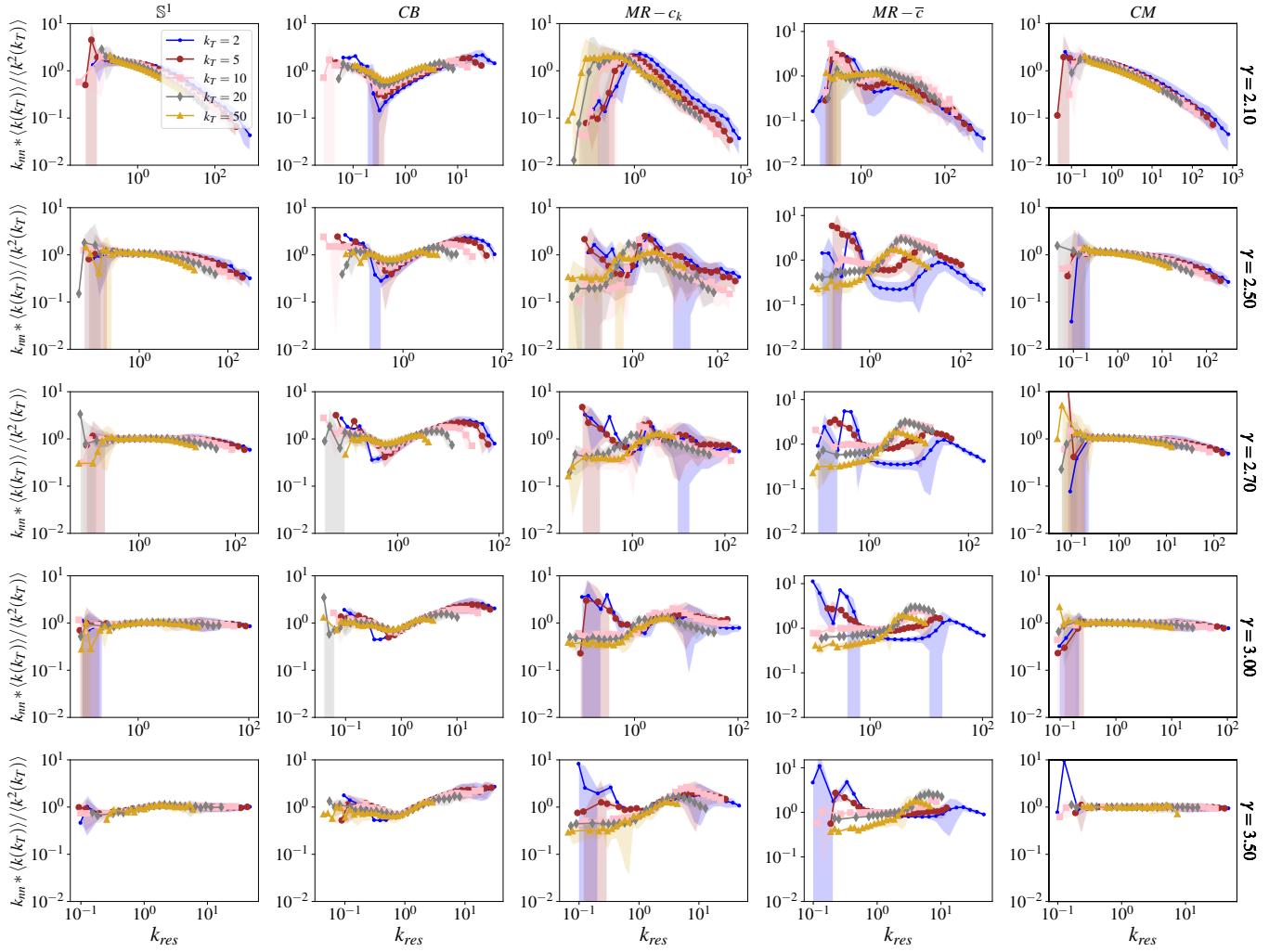


FIG. 7: Average nearest-neighbor degree of the synthetic networks in the DTR flow as a function of the degree rescaled by the average degree in the corresponding subgraph, $k_{\text{res}} = k / \langle k(k_T) \rangle$. Each column corresponds to synthetic networks generated by a model with $N = 50000$, $\langle k \rangle = 10$, $\beta = 3$ and varying values of γ . The shaded areas represent two- σ intervals around the mean, calculated from 10 realizations of each model.

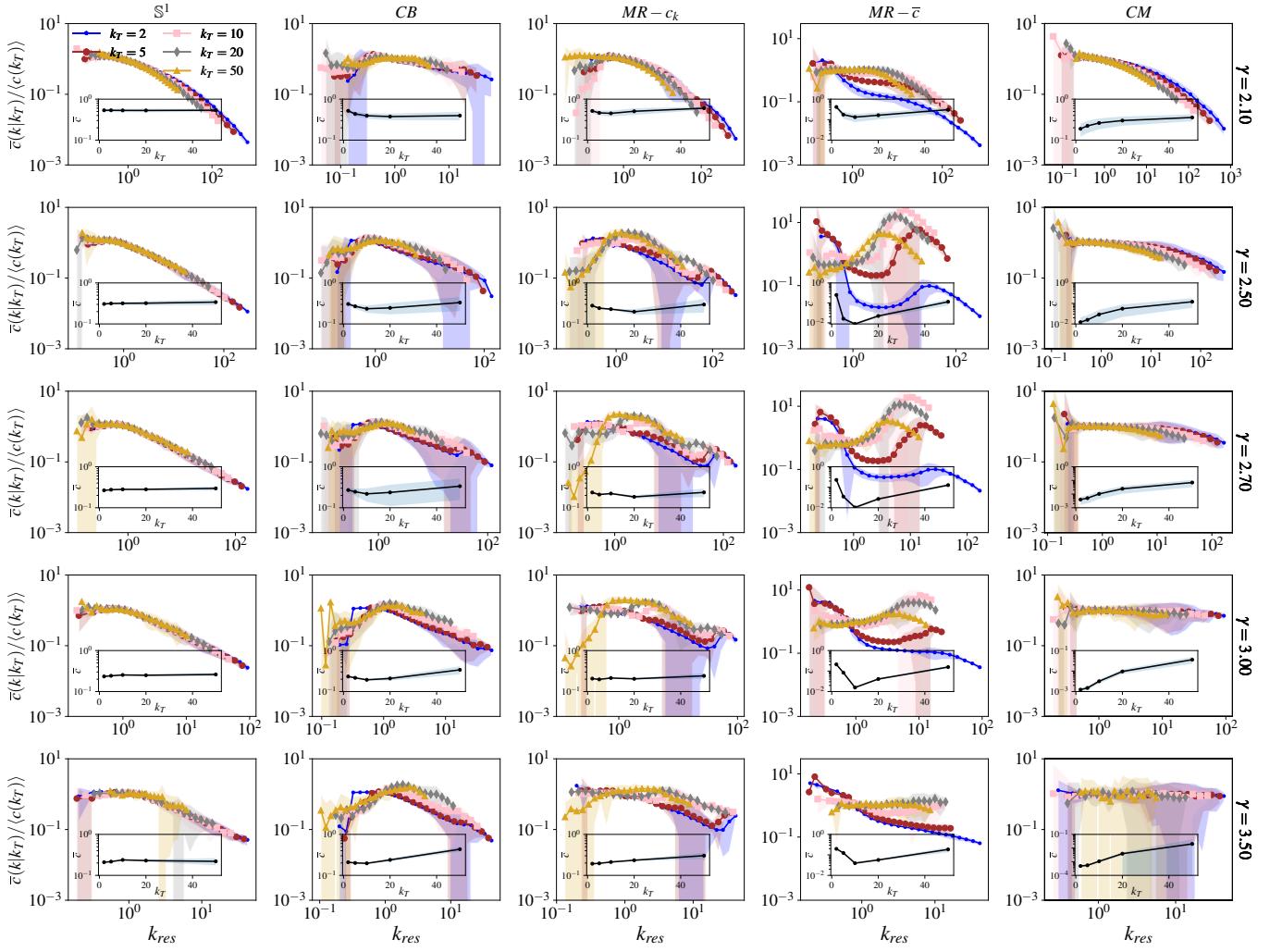


FIG. 8: Degree-dependent clustering coefficient of the synthetic networks in the DTR flow as a function of the degree rescaled by the average degree in the corresponding subgraph, $k_{\text{res}} = k/\langle k(k_T) \rangle$. Each column corresponds to synthetic networks generated by a model with $N = 50000$, $\langle k \rangle = 10$, $\beta = 1.5$ and varying values of γ . The inset highlights 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.

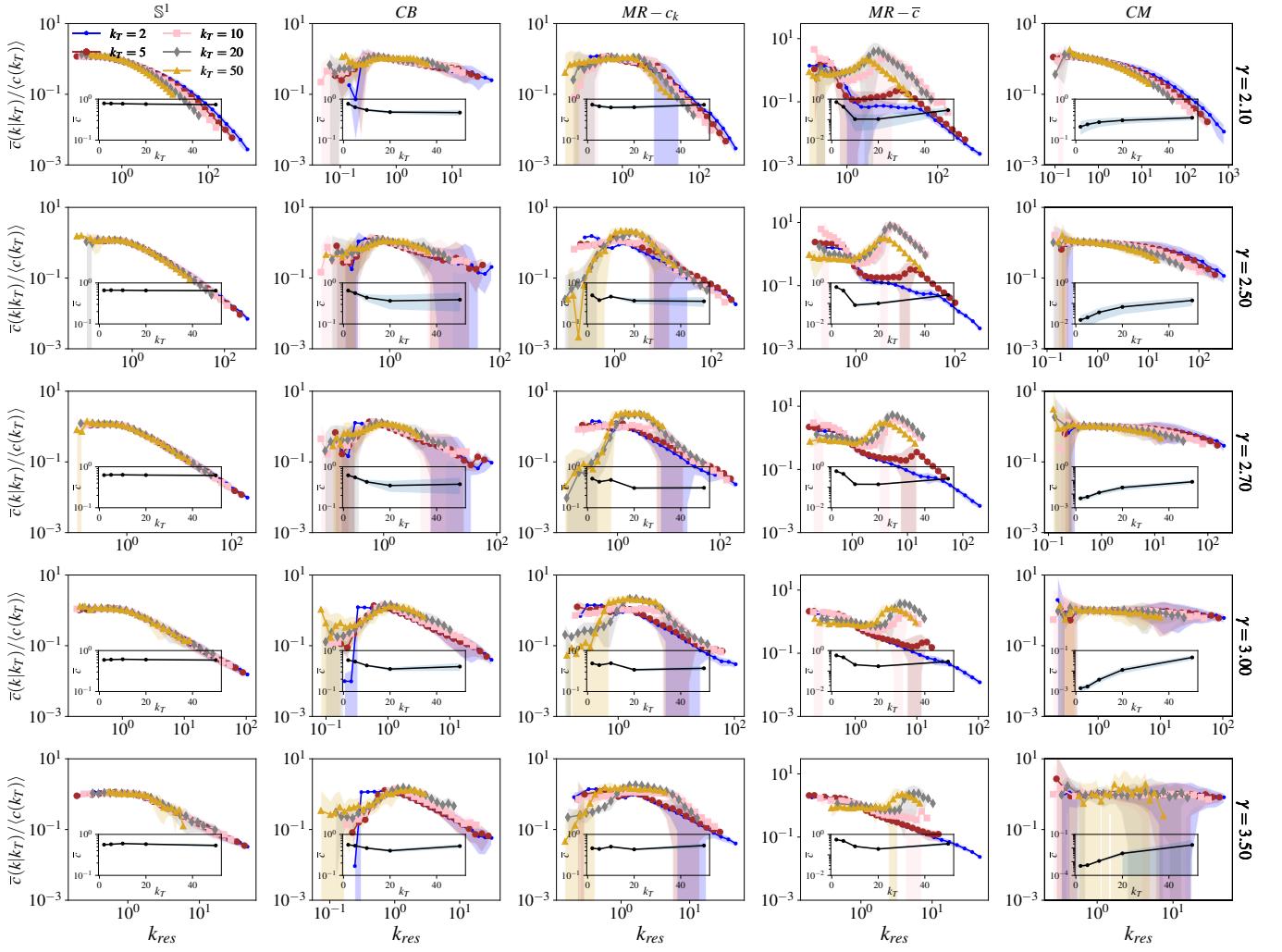


FIG. 9: Degree-dependent clustering coefficient of the synthetic networks in the DTR flow as a function of the degree rescaled by the average degree in the corresponding subgraph, $k_{\text{res}} = k/\langle k(k_T) \rangle$. Each column corresponds to synthetic networks generated by a model with $N = 50000$, $\langle k \rangle = 10$, $\beta = 3$ and varying values of γ . The inset highlights average clustering coefficient as a function of k_T . The shaded areas represent two-\$\sigma\$ intervals around the mean, calculated from 10 realizations of each model.

Algorithm 1 Algorithm for classifying networks as self-similar, quasi-self-similar, or non-self-similar

- 1: Calculate the maximum value of the degree-thresholding parameter, k_T^{\max} , corresponding to the peak of the curve $\langle k(k_T) \rangle / \langle k \rangle$ as a function of k_T .
- 2: Divide $[2, k_T^{\max}]$ into n evenly spaced intervals (except for the last one) to determine $n + 1$ values for k_T . Set $n = \max(n^{\min}, \min(n^{\max}, \lfloor (k_T^{\max} - 2)/\Delta k \rfloor))$, with $\Delta k = \max(\Delta k^{\min}, \lfloor (k_T^{\max} - 2)/n^{\max} \rfloor)$.
- 3: For each value of k_T , apply DTR on the original network
- 4: Plot CCD as a function of rescaled degree, $k_{res} = k / \langle k(k_T) \rangle$, in the DTR flow
- 5: Plot $k_{nn} \times \langle k(k_T) \rangle / \langle k^2(k_T) \rangle$ as a function of rescaled degree in the DTR flow
- 6: Plot $\bar{c}(k|k_T) / \langle c(k_T) \rangle$ as a function of rescaled degree in the DTR flow
- 7: Compute the average of the squared relative differences between each curve in the DTR flow and the one for original network across all three plots (k_{nn} , CCD , and $\bar{c}(k)$), using exponential binning with n_{bins} number of bins within the shared range of the x-axis for the two curves.

$$\epsilon^2 = \frac{1}{n+1} \frac{1}{n_{bins}} \sum_{k_T} \sum_{i=1}^{n_{bins}} \left(\frac{\text{bin value}_i^{org} - \text{bin value}_i^{k_T}}{\text{bin value}_i^{org}} \right)^2$$

- 8: Select $\epsilon_{Max}^2 = \text{Max}(\epsilon_{CCD}^2, \epsilon_{knn}^2, \epsilon_{\bar{c}}^2)$
- 9: **if** $\epsilon_{Max}^2 < 0.4$ **then**
- 10: Classify the network as self-similar
- 11: **else if** $0.4 \leq \epsilon_{Max}^2 < 1$ **then**
- 12: Classify the network as Quasi-self-similar
- 13: **else**
- 14: Classify the network as Non-self-similar
- 15: **end if**

TABLE II: ϵ^2 values computed for different structural properties including complementary cumulative degree distribution ϵ_{CCD}^2 , average nearest-neighbor degree function ϵ_{knn}^2 , and clustering spectrum $\epsilon_{\bar{c}}^2$ in real-world network. The bolded values represent the maximum ϵ^2 for each network, denoted as ϵ_{Max}^2 and used to classify the networks into self-similar, quasi-self-similar, or non-self-similar.

Networks	ϵ_{CCD}^2	ϵ_{knn}^2	$\epsilon_{\bar{c}}^2$
Astrophysics	0.29	0.06	0.13
Bible-Nouns	0.25	0.16	0.26
PGP	0.35	0.16	0.24
Internet	0.32	0.21	0.22
Facebook	0.22	0.07	0.13
Fb-Friends	0.12	0.05	0.13
Int-4896	0.29	0.22	0.24
Solo-BrightKite	0.25	0.07	0.39
Bitcoin-Trust	0.32	0.18	0.26
GMP-Drosophila	0.26	0.11	0.59
GMP-Mus	0.27	0.23	0.47
GMP-S.cerevisiae	0.26	0.08	0.08
PPI-Drosophila	0.22	0.12	0.32
Wikipedia-am	0.38	0.37	0.33
PPI-rat	0.31	0.55	6.27
GMP-Celegans	0.24	0.28	2.31
GI-S.cerevisiae	0.27	0.05	0.05
Japanese-Words	0.3	0.21	0.23
GMP-Plasmodium	0.28	0.27	0.26
MB-R.norvegicus	0.49	0.45	0.48
Int-figeys	0.24	0.33	1.13
WikiTalk-Catalan	9.89	0.53	0.29

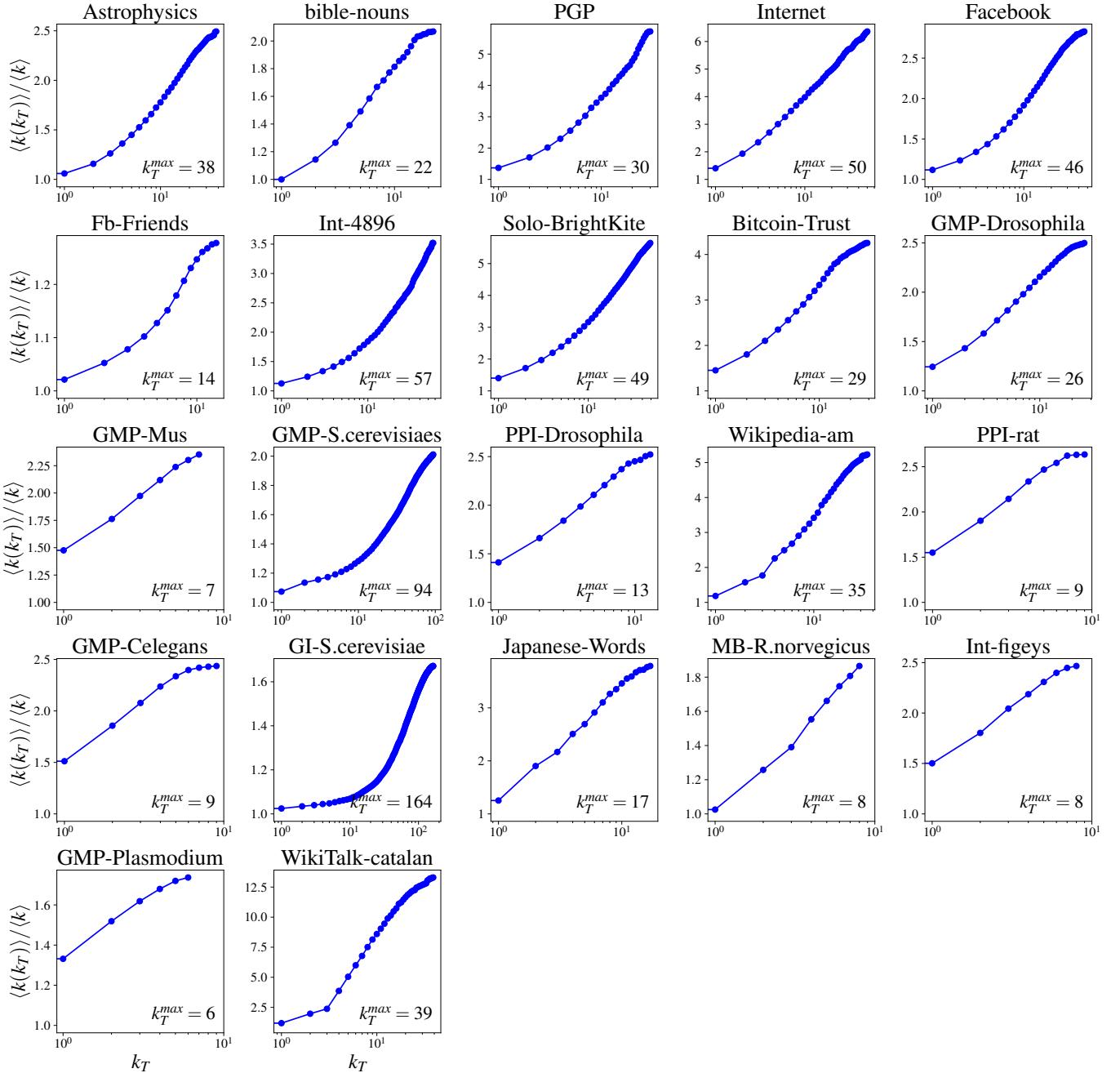


FIG. 10: Average degree of subgraphs, normalized by the average degree of the original network as a function of k_T in DTR flows. The maximum value for k_T is the one at the peak of the curve.

Appendix F: Dataset description

Astrophysics [34]: Nodes represent authors from the astrophysics section of arXiv (astro-p), and edges indicate collaborations.

Bible-Nouns [35]: Nodes are noun phrases (places and names) from the King James Bible, with edges denoting co-occurrence in verses.

PGP [36]: Nodes are users of the Pretty Good Privacy algorithm, and edges represent trust relationships through key signing.

Internet [37, 38]: Nodes are Autonomous Systems (ASs), and edges represent the connections between them in the network.

Facebook [39]: Nodes represent Facebook pages, and edges denote mutual likes between them.

Fb-Friends [40]: Nodes are university students from the Copenhagen Networks Study, and edges represent Facebook friendships.

Int-4896 [41]: Nodes represent the proteins of the *Schizosaccharomyces pombe* species, and edges denote experimentally verified physical interactions, including protein–protein, protein–DNA, metabolic pathway, and kinase–substrate interactions.

Solo-BrightKite [42]: Nodes represent users of the Brightkite location-based social networking service, and edges indicate mutual friendships between users.

Bitcoin-Trust [43, 44]: Nodes represent users of the Bitcoin OTC platform, and directed edges indicate trust ratings, where each edge includes a weight assigned by one user to another.

GMP-Drosophila, GMP-Mus, GMP-S.cerevisiae, GMP-Celegans, GMP-Plasmodium [45]: Multiplex networks representing different types of genetic interactions in various organisms. Nodes represent genes, and edges capture different genetic interactions across different layers, including physical, association, co-localization, direct, suppressive, and additive/synthetic interactions.

PPI-Drosophila [46]: Nodes represent proteins in the *Drosophila melanogaster* species, and edges represent binary physical protein–protein interactions identified using yeast two-hybrid (Y2H) analysis.

Wikipedia-am [47]: Nodes represent articles in the Amharic language edition of Wikipedia, and directed edges represent hyperlinks between articles.

PPI-rat [48]: Nodes represent proteins in the rat species, and edges represent physical protein–protein interactions between them.

GI-S.cerevisiae [48]: Nodes represent genes in *Saccharomyces cerevisiae* (baker’s yeast), and edges indicate genetic interactions where mutations in one gene affect or are modified by mutations in another gene.

Japanese-Words [49]: Nodes represent words in Japanese texts, and edges indicate adjacency, where one word directly follows another. The original network is directed.

MB-R.norvegicus [50]: Nodes represent substances involved in enzymatic reactions in *Rattus norvegicus* (rat), and edges indicate reactant-product relationships in these reactions.

Int-figeys [51]: Nodes represent human proteins, and edges denote binding interactions between proteins.

WikiTalk-Catalan [47]: Nodes represent registered editors in the Catalan Wikipedia, and edges indicate interactions where one user writes a message on another user’s talk page.

In this work, we ignore weights in the weighted networks and construct undirected versions of all directed networks. In the case of multiplex networks, we create a monolayer network by treating all interaction types equally and removing repeated links.

Appendix G: Experimental Analysis of Real-World Networks

Here, we analyze the DTR self-similarity in real networks. Table III summarizes the key properties of these networks. Figs. 11, 12, and 13 showcase how the fundamental structural properties of self-similar networks varies under DTR. As expected for self-similar networks and in agreement with the ϵ^2 -Test values reported in Table III, the degree distribution, average nearest-neighbor degree, and clustering spectrum exhibit minimal variation across different values of k_T . Similarly, Figs. 14 and 15 depict the corresponding variations for quasi-self-similar and non-self-similar networks, respectively. These variations are more pronounced for quasi-self-similar and non-self-similar networks, highlighting their weaker structural invariance under DTR.

Moreover, Fig. 16 presents a Sankey diagram for the real-world networks sorted from the top in descending order of their estimated β , which quantifies the coupling between network topology and the underlying hyperbolic space of the S^1 model. On the left side, they are categorizing as self-similar, quasi-self-similar, or non-self-similar based on their ϵ^2 -Test values displayed in parenthesis and the thresholds from Fig. 3-(b). The results highlight that all non-self-similar networks have $\beta < 1$, classifying them as either non-geometric or quasi-geometric. Additionally, nearly all geometric networks are self-similar, with the exception of GMP-Drosophila, which has a β value near the critical threshold $\beta_c = 1$ and is thus classified as quasi-self-similar. These findings reinforce the connection between self-similarity and geometricity in real networks.

TABLE III: Real network Properties. N is the number of nodes, E is the number of edges, $\langle k \rangle$ is average degree, k_{\max} is maximum degree and \bar{c} is the average clustering in the giant connected component of the real networks. β indicates the level of clustering and computed using Mercator embedding tool. Type S refers to self-similar networks, Type NS indicates non-self-similar networks, and Type QS denotes quasi-self-similar networks. The following abbreviations are used: (MB) for metabolic, (GI) for genetic interactions, (GMP) for genetic multiplex, (PPI) for protein-protein interactions, (Solo) for social locations, and (Int) for interactome.

Networks	N	E	$\langle k \rangle$	k_{\max}	\bar{c}	β	ϵ_{Max}^2	Type
Astrophysics	14845	119652	16.12	360	0.72	3.27	0.29	S
Bible-Nouns	1707	9059	10.61	364	0.71	3.2	0.26	S
PGP	10680	24316	4.55	205	0.44	1.95	0.35	S
Internet	23748	58414	4.92	2778	0.61	1.95	0.32	S
Facebook	22470	170823	15.2	709	0.41	1.6	0.22	S
Fb-Friends	800	6418	16.05	101	0.32	1.54	0.12	S
Int-4896	4086	47961	23.48	448	0.41	1.43	0.29	S
Solo-BrightKite	56739	212945	7.51	1134	0.27	1.33	0.39	S
Bitcoin-Trust	5875	21489	7.31	795	0.29	1.13	0.32	S
GMP-Drosophila	8114	38909	9.6	179	0.12	1.07	0.59	QS
GMP-Mus	7402	16858	4.56	368	0.13	0.99	0.47	QS
GMP-S.cerevisiae	6567	223539	68.08	3254	0.22	0.87	0.26	S
PPI-Drosophila	2705	8458	6.25	129	0.07	0.78	0.32	S
Wikipedia-am	20883	94022	9	3911	0.18	0.76	0.38	S
PPI-rat	6803	14636	4.3	836	0.15	0.72	6.27	NS
GMP-Celegans	3692	7650	4.14	526	0.11	0.69	2.31	NS
GI-S.cerevisiae	5933	441991	148.99	2244	0.17	0.63	0.27	S
Japanese-Words	2698	7995	5.93	725	0.3	$\beta \approx 0$	0.3	S
GMP-Plasmodium	1158	2402	4.15	83	0.03	$\beta \approx 0$	0.28	S
MB-R.norvegicus	1590	4666	5.87	498	0.19	$\beta \approx 0$	0.49	QS
Int-figeys	2217	6418	5.79	314	0.07	$\beta \approx 0$	1.13	NS
WikiTalk-Catalan	79209	181529	4.59	53234	0.83	$\beta \approx 0$	9.89	NS

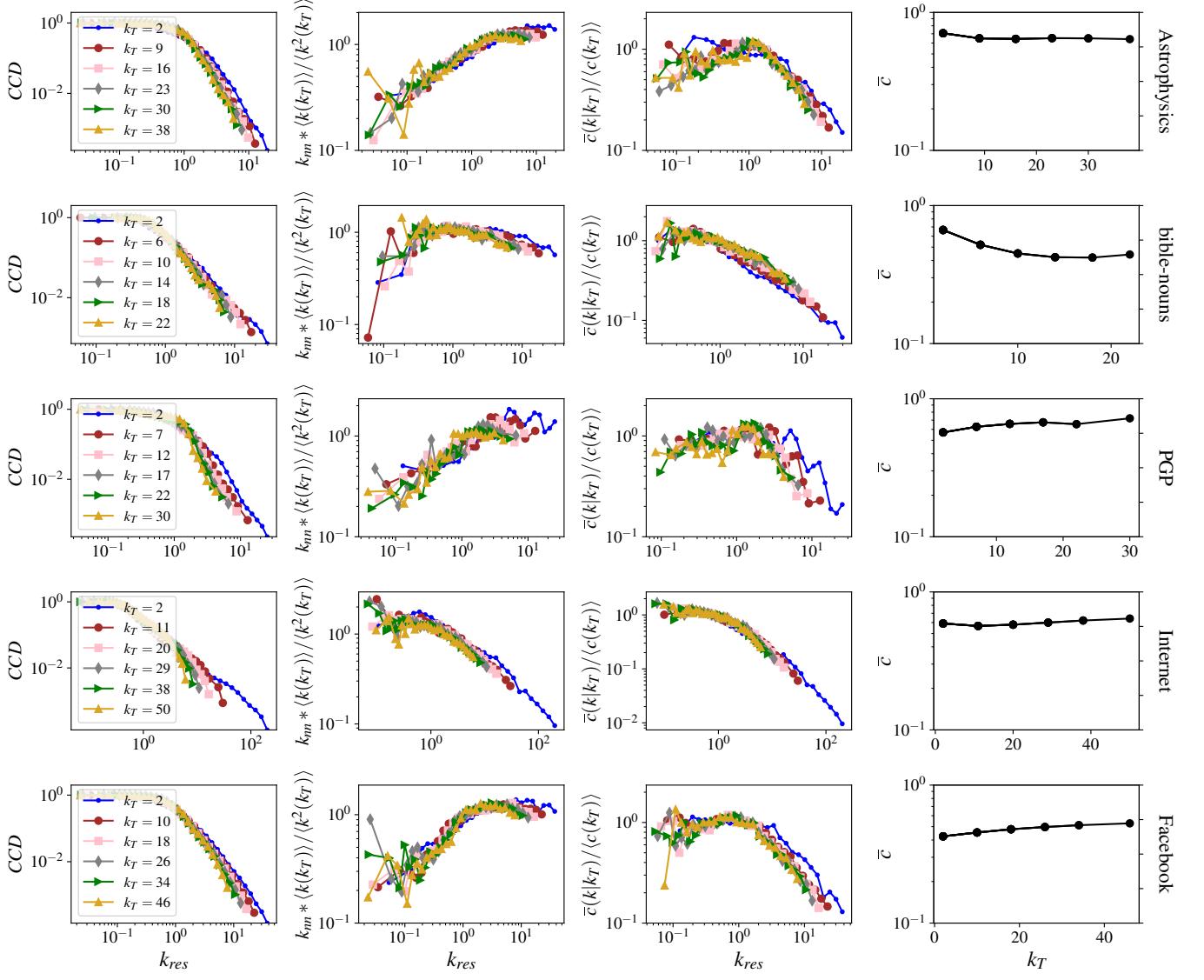


FIG. 11: Topological properties of real-world **self-similar** networks, including the complementary cumulative degree distribution (CCD), the rescaled average nearest-neighbor degree $k_{nn} * \langle k(k_T) \rangle / \langle k^2(k_T) \rangle$, the rescaled clustering spectrum $\bar{c}(k|k_T) / \langle c(k_T) \rangle$ as a function of the rescaled degree $k_{res} = k / \langle k(k_T) \rangle$, and the average clustering coefficient \bar{c} as a function of k_T in the DTR flow.

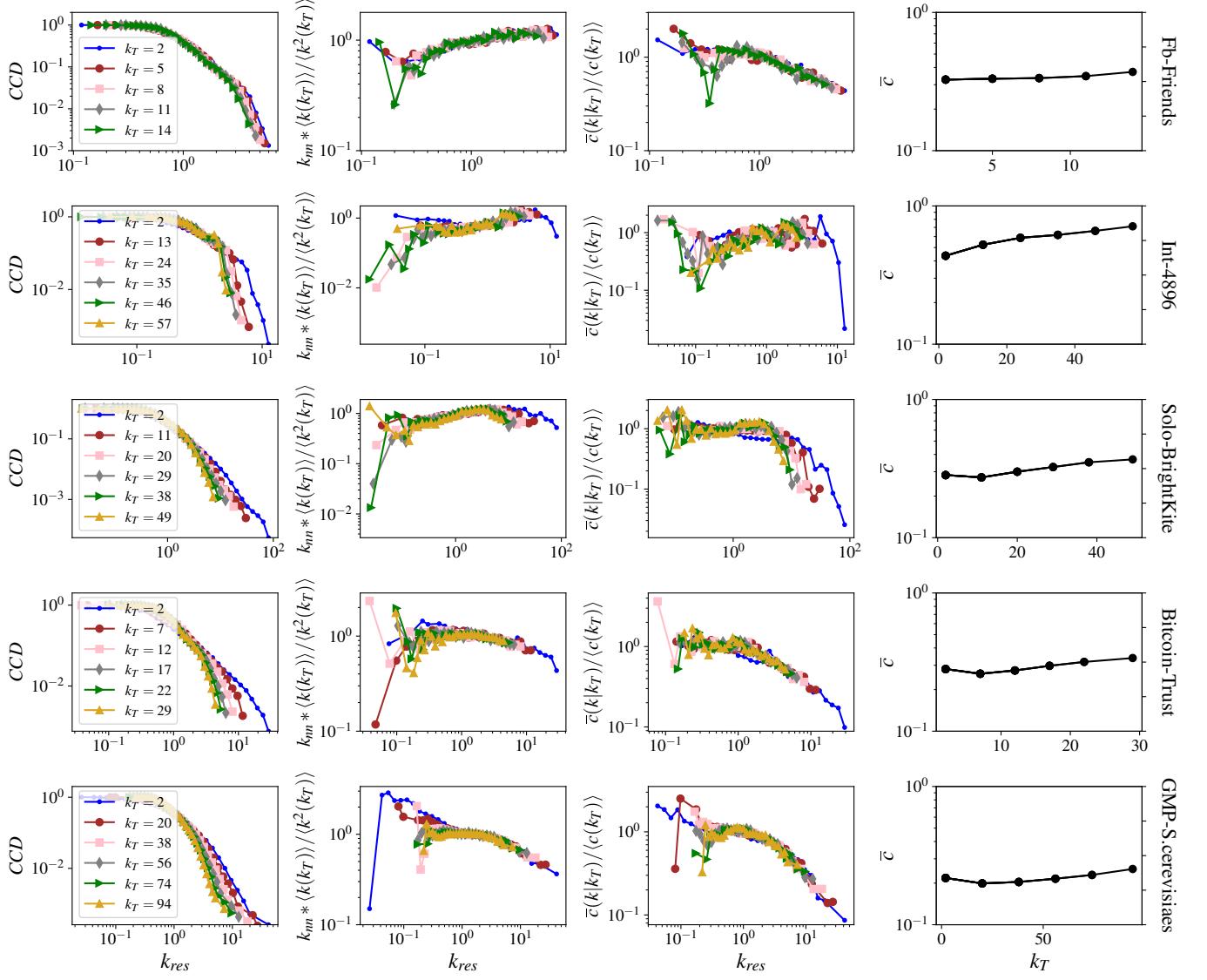


FIG. 12: Topological properties of real-world **self-similar** networks, including the complementary cumulative degree distribution (CCD), the rescaled average nearest-neighbor degree $k_{\text{nn}} * \langle k(k_T) \rangle / \langle k^2(k_T) \rangle$, the rescaled clustering spectrum $\bar{c}(k|k_T) / \langle c(k_T) \rangle$ as a function of the rescaled degree $k_{\text{res}} = k / \langle k(k_T) \rangle$, and the average clustering coefficient \bar{c} as a function of k_T in the DTR flow.

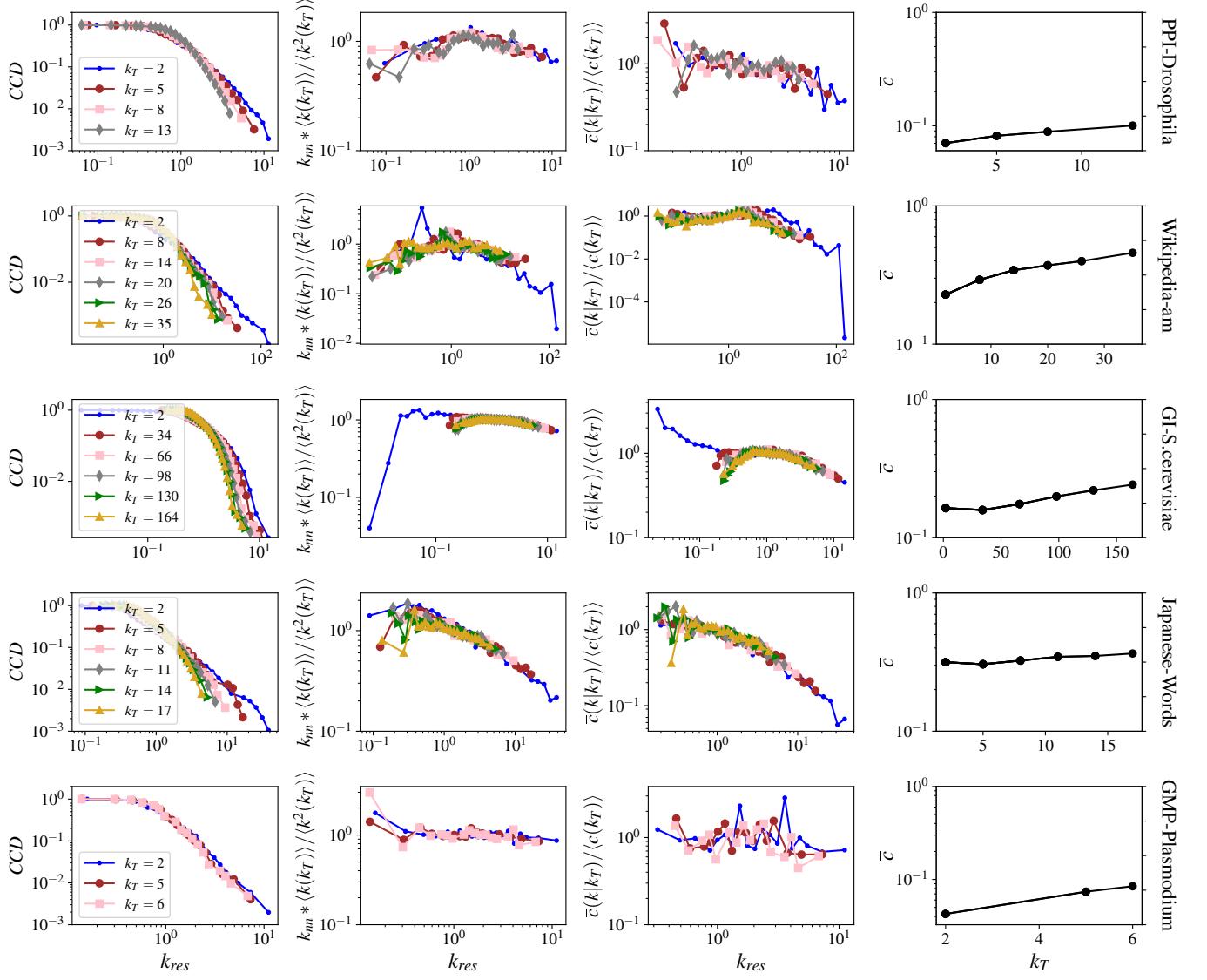


FIG. 13: Topological properties of real-world **self-similar** networks, including the complementary cumulative degree distribution (CCD), the rescaled average nearest-neighbor degree $k_{nn} * \langle k(k_T) \rangle / \langle k^2(k_T) \rangle$, the rescaled clustering spectrum $\bar{c}(k|k_T) / \langle c(k_T) \rangle$ as a function of the rescaled degree $k_{res} = k / \langle k(k_T) \rangle$, and the average clustering coefficient \bar{c} as a function of k_T in the DTR flow.

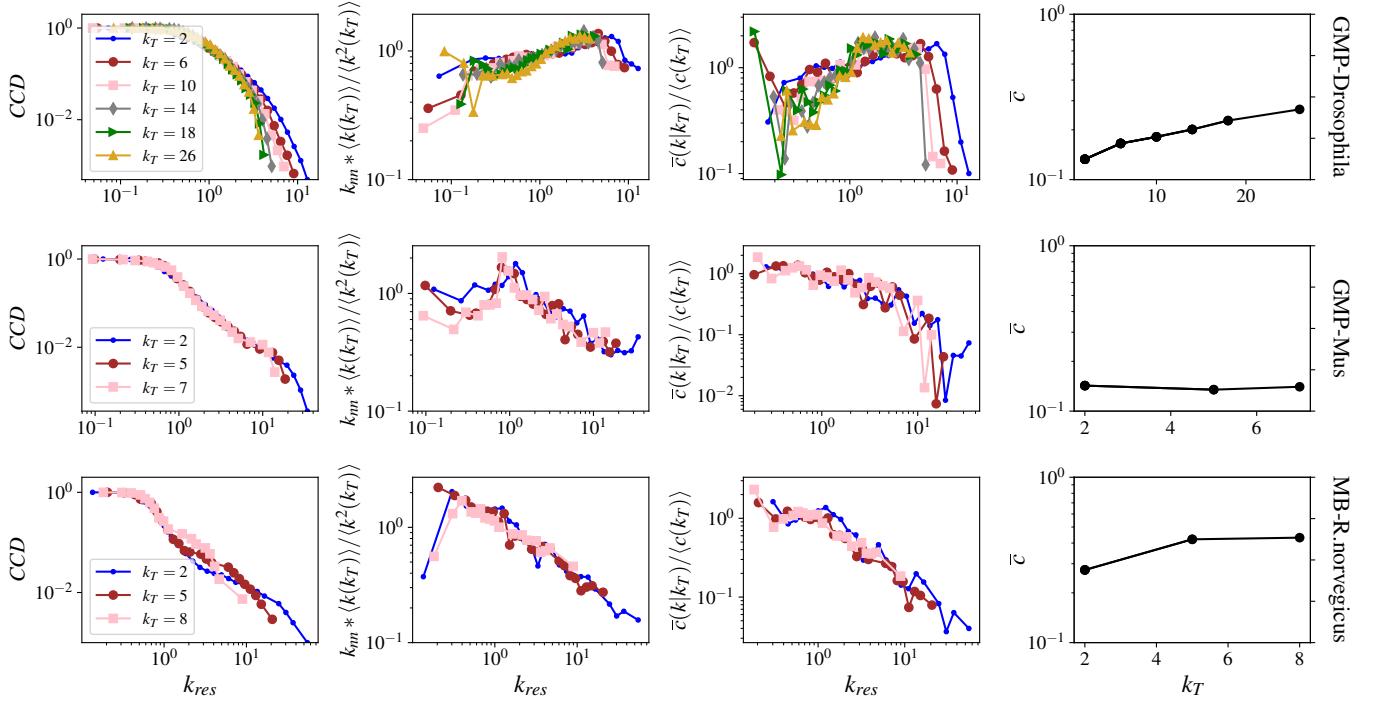


FIG. 14: Topological properties of real-world **quasi-self-similar** networks, including the complementary cumulative degree distribution (CCD), the rescaled average nearest-neighbor degree $k_{nn} * \langle k(k_T) \rangle / \langle k^2(k_T) \rangle$, the rescaled clustering spectrum $\bar{c}(k|k_T) / \langle c(k_T) \rangle$ as a function of the rescaled degree $k_{res} = k / \langle k(k_T) \rangle$, and the average clustering coefficient \bar{c} as a function of k_T in the DTR flow.

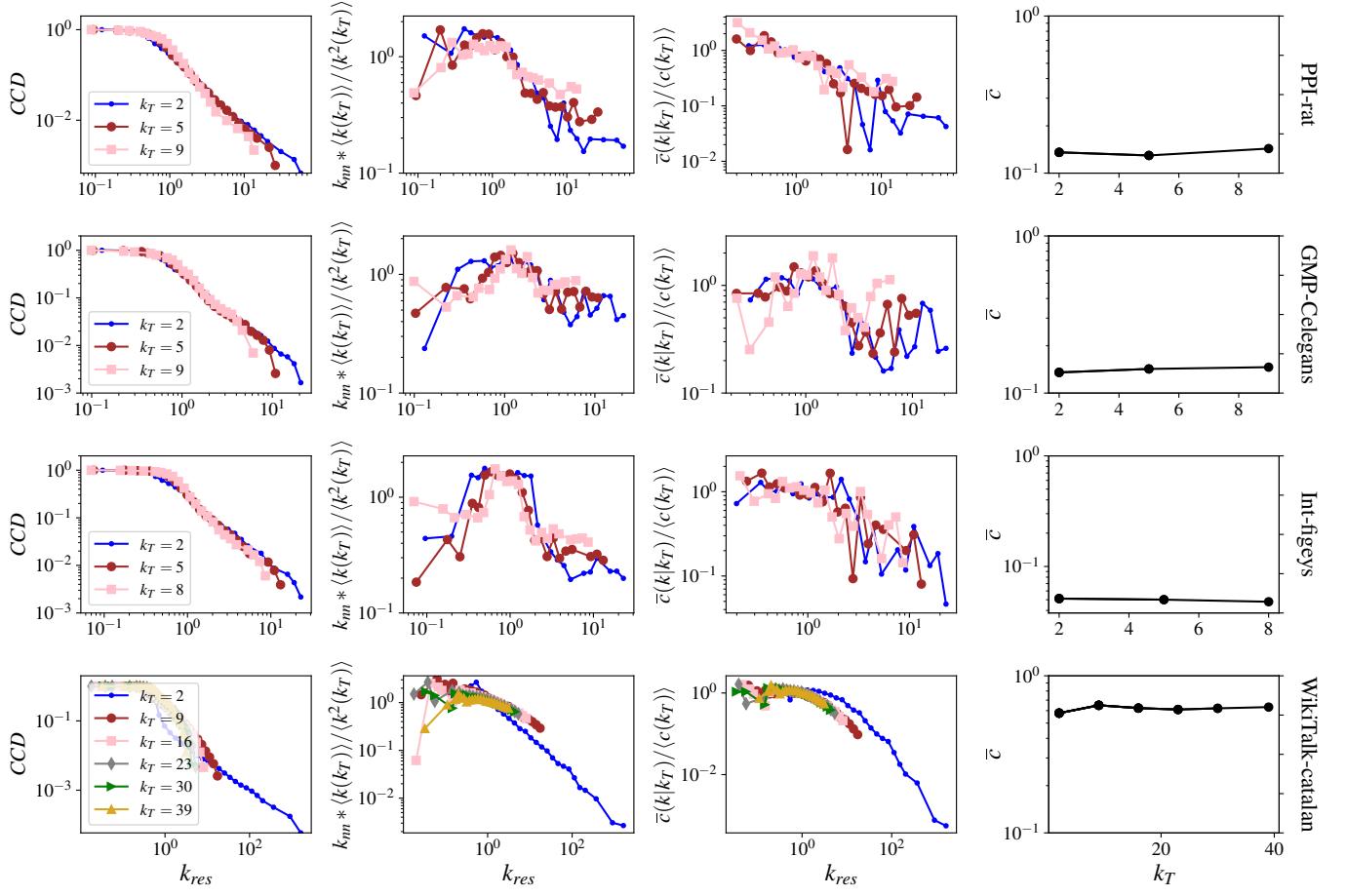


FIG. 15: Topological properties of real-world **non-self-similar** networks, including the complementary cumulative degree distribution (CCD), the rescaled average nearest-neighbor degree $k_{nn} * \langle k(k_T) \rangle / \langle k^2(k_T) \rangle$, the rescaled clustering spectrum $\bar{c}(k|k_T) / \langle c(k_T) \rangle$ as a function of the rescaled degree $k_{res} = k / \langle k(k_T) \rangle$, and the average clustering coefficient \bar{c} as a function of k_T in the DTR flow.

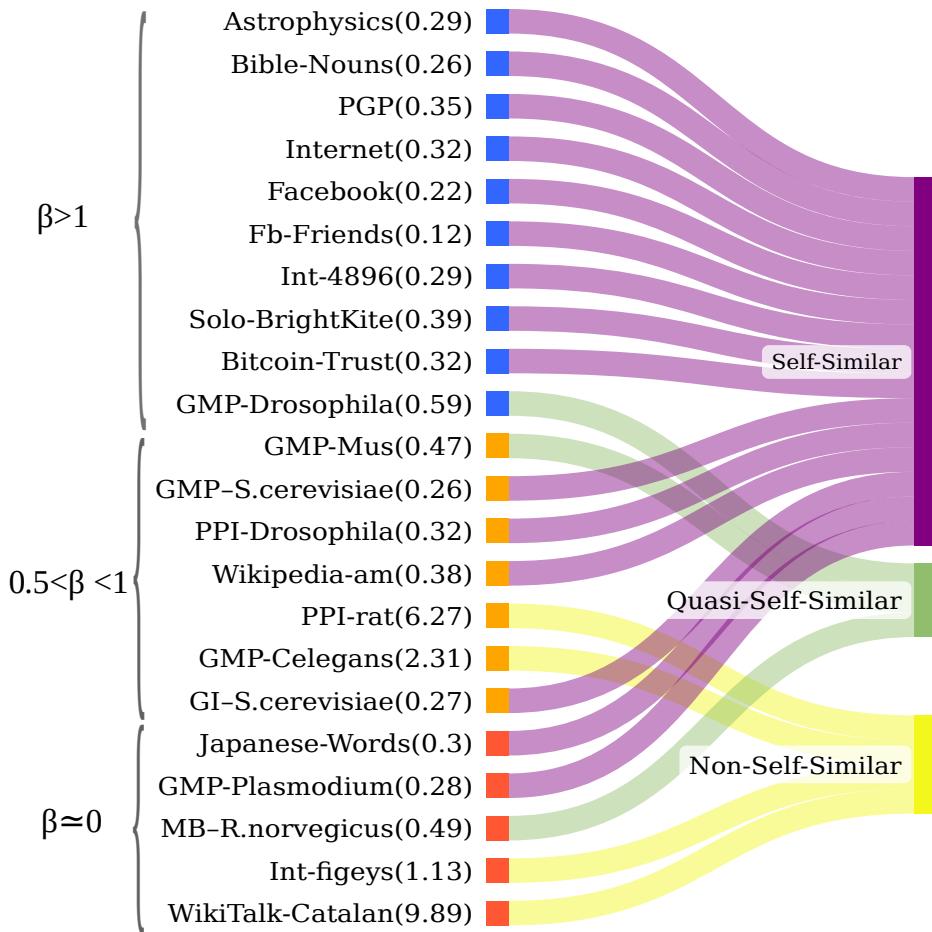


FIG. 16: Sankey diagram of real networks, sorted from the top in descending order by β values. On the left, networks are classified as geometric ($\beta > 1$), quasi-geometric ($0.5 < \beta < 1$), and non-geometric ($\beta \approx 0$). On the right, they are categorized in terms of self-similarity, using their ϵ_{Max}^2 values displayed in parentheses.

-
- [1] M. Boguñá, I. Bonamassa, M. D. Domenico, S. Havlin, D. Krioukov, and M. Á. Serrano, *Nature Reviews Physics* **3**, 114 (2021).
- [2] Z. Wu, G. Menichetti, C. Rahmede, and G. Bianconi, *Scientific Reports* **5**, 10073 (2015).
- [3] L. Bombelli, J. Lee, D. Meyer, and R. D. Sorkin, *Phys. Rev. Lett.* **59**, 521 (1987).
- [4] S. Surya, *Living Reviews in Relativity* **22** (2019), 10.1007/s41114-019-0023-1.
- [5] D. Krioukov, M. Kitsak, R. S. Sinkovits, D. Rideout, D. Meyer, and M. Boguñá, *Scientific Reports* **2**, 793 (2012).
- [6] M. Penrose, *Random Geometric Graphs* (Oxford University Press, 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, 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 Mathematics* **12**, 2 (2016).
- [10] M. E. Newman, *Physical Review 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* **433**, 392 (2005).
- [15] C. Song, S. Havlin, and H. A. Makse, *Nature Physics* **2**, 275 (2006).
- [16] J. S. Kim, K.-I. Goh, B. Kahng, and D. Kim, *New Journal of Physics* **9**, 177 (2007).
- [17] J. Alvarez-Hamelin, L. Dall'Asta, A. Barrat, and A. Vespignani, *Networks and Heterogeneous Media* **3**, 371 (2008).
- [18] G. García-Pérez, M. Boguñá, and M. Á. Serrano, *Nature Physics* **14**, 583 (2018).
- [19] R. Jankowski, A. Allard, M. Boguñá, and M. Á. Serrano, *Nature Communications* **14**, 7585 (2023).
- [20] G. Budel, M. Kitsak, R. Aldecoa, K. Zuev, and D. Krioukov, *Physical Review E* **109**, 054131 (2024).
- [21] E. A. Bender and E. Canfield, *Journal of Combinatorial Theory, Series A* **24**, 296 (1978).
- [22] M. Molloy and B. Reed, *Random Structures & Algorithms* **6**, 161 (1995).
- [23] J. P. Gleeson, *Phys. Rev. E* **80**, 036107 (2009).
- [24] P. Colomer-de Simón, M. Á. Serrano, M. G. Beiró, J. I. Alvarez-Hamelin, and M. Boguñá, *Sci. Rep.* **3**, 2517 (2013).
- [25] M. Boguñá, D. Krioukov, P. Almagro, and M. Á. Serrano, *Phys. Rev. Res.* **2**, 023040 (2020).
- [26] P. Colomer-de Simón and M. Boguñá, *Phys. Rev. X* **4**, 041020 (2014).
- [27] J. van der Kolk, M. A. Serrano, and M. Boguñá, *Communications Physics* **5**, 245 (2022).
- [28] L. Gugelmann, K. Panagiotou, and U. Peter, in *Autom Lang Program (ICALP 2012, Part II)*, LNCS **7392** (2012).
- [29] E. Candellero and N. Fountoulakis, *Internet Math.* **12**, 2 (2016).
- [30] N. Fountoulakis, P. van der Hoorn, T. Müller, and M. Schepers, *Electron. J. Probab.* **26**, 1 (2021).
- [31] M. A. Abdullah, N. Fountoulakis, and M. Bode, *Internet Math.* **1** (2017), 10.24166/im.13.2017.
- [32] T. Friedrich and A. Krohmer, *SIAM J. Discrete Math.* **32**, 1314 (2018).
- [33] T. Müller and M. Staps, *Adv. Appl. Probab.* **51**, 358 (2019).
- [34] M. E. J. Newman, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 404 (2001).
- [35] C. Römhild and C. Harrison, “Bible cross-references,” <http://www.chrisharrison.net/index.php/Visualizations/BibleViz>, accessed: 2024-10-02.
- [36] M. Boguñá, R. Pastor-Satorras, A. Díaz-Guilera, and A. Arenas, *Phys. Rev. E* **70**, 056122 (2004).
- [37] K. Claffy, Y. Hyun, K. Keys, M. Fomenkov, and D. Krioukov, in *2009 Cybersecurity Applications and Technology Conference for Homeland Security* (2009) pp. 205–211.
- [38] M. Boguñá, F. Papadopoulos, and D. Krioukov, *Nature Communications* **1**, 62 (2010).
- [39] B. Rozemberczki, C. Allen, and R. Sarkar, *Journal of Complex Networks* **9**, cnab014 (2021), <https://academic.oup.com/comnet/article-pdf/9/2/cnab014/40435146/cnab014.pdf>.
- [40] P. Sapiezynski, A. Stopczynski, D. D. Lassen, and S. Lehmann, *Scientific Data* **6**, 315 (2019).
- [41] M. Zitnik, R. Sosić, M. W. Feldman, and J. Leskovec, *Proceedings of the National Academy of Sciences* **116**, 4426 (2019), <https://www.pnas.org/doi/pdf/10.1073/pnas.1818013116>.
- [42] 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, NY, USA, 2011) p. 1082–1090.
- [43] S. Kumar, F. Spezzano, V. Subrahmanian, and C. Faloutsos, in *Data Mining (ICDM), 2016 IEEE 16th International Conference on* (IEEE, 2016) pp. 221–230.
- [44] 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, 2018) pp. 333–341.
- [45] M. De Domenico, M. A. Porter, and A. Arenas, *Journal of Complex Networks* **3**, 159 (2014), <https://academic.oup.com/comnet/article-pdf/3/2/159/1070864/cnu038.pdf>.
- [46] H.-W. Tang, K. Spirohn, Y. Hu, T. Hao, I. A. Kovács, Y. Gao, R. Binari, D. Yang-Zhou, K. H. Wan, J. S. Bader, D. Balcha, W. Bian, B. W. Booth, A. G. Coté, S. de Rouck, A. Desbuleux, K. Y. Goh, D.-K. Kim, J. J. Knapp, W. X. Lee, I. Lemmens, C. Li, M. Li, R. Li, H. J. Lim, Y. Liu, K. Luck, D. Markey, C. Pollis, S. Rangarajan, J. Rodiger, S. Schlabach, Y. Shen, D. Sheykhhari, B. TeeKing, F. P. Roth, J. Tavernier, M. A. Calderwood, D. E. Hill, S. E. Celinker, M. Vidal, N. Perrimon, and S. E. Mohr, *Nature Communications* **14**, 2162 (2023).

- [47] J. Kunegis, in *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13 Companion (Association for Computing Machinery, New York, NY, USA, 2013) p. 1343–1350.
- [48] Y. Hu, A. Vinayagam, A. Nand, A. Comjean, V. Chung, T. Hao, S. E. Mohr, and N. Perrimon, *Nucleic Acids Research* **46**, D567 (2017), <https://academic.oup.com/nar/article-pdf/46/D1/D567/23162375/gkx1116.pdf>.
- [49] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenstat, M. Sheffer, and U. Alon, *Science* **303**, 1538 (2004), <https://www.science.org/doi/pdf/10.1126/science.1089167>.
- [50] P. H. M. Huss, *IET Systems Biology* **1**, 280 (2007).
- [51] R. M. Ewing, P. Chu, F. Elisma, H. Li, P. Taylor, S. Climie, L. McBroom-Cerajewski, M. D. Robinson, L. O'Connor, M. Li, R. Taylor, M. Dharsee, Y. Ho, A. Heilbut, L. Moore, S. Zhang, O. Ornatsky, Y. V. Bukhman, M. Ethier, Y. Sheng, J. Vasilescu, M. Abu-Farha, J. Lambert, H. S. Duewel, I. I. Stewart, B. Kuehl, K. Hogue, K. Colwill, K. Gladwish, B. Muskat, R. Kinach, S. Adams, M. F. Moran, G. B. Morin, T. Topaloglou, and D. Figeys, *Molecular Systems Biology* **3**, 89 (2007), <https://www.embopress.org/doi/pdf/10.1038/msb4100134>.