


RESEARCH ARTICLE

Assortativity in geometric and scale-free networks

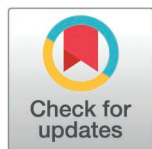
Marc Kaufmann¹, Ulysse Schaller¹, Thomas Bläsius^{2†}, Johannes Lengler^{1‡*}

1 Institut für Theoretische Informatik, ETH Zürich, Zürich, Switzerland, **2** Institut für Theoretische Informatik, Karlsruhe Institute of Technology, Karlsruhe, Germany

 These authors contributed equally to this work.

 Shared senior authorship.

* johannes.lengler@inf.ethz.ch



Abstract

The assortative behavior of a network is the tendency of similar (or dissimilar) nodes to connect to each other. This tendency can have an influence on various properties of the network, such as its robustness or the dynamics of spreading processes. In this paper, we study degree assortativity both in real-world networks and in several generative models for networks with heavy-tailed degree distribution based on latent spaces. In particular, we study Chung-Lu Graphs and Geometric Inhomogeneous Random Graphs (GIRGs). Previous research on assortativity has primarily focused on measuring the degree assortativity in real-world networks using the Pearson assortativity coefficient, despite reservations against this coefficient. We rigorously confirm these reservations by mathematically proving that the Pearson assortativity coefficient does not measure assortativity in any network with sufficiently heavy-tailed degree distributions, which is typical for real-world networks. Moreover, we find that other single-valued assortativity coefficients also do not sufficiently capture the wiring preferences of nodes, which often vary greatly by node degree. We therefore take a more fine-grained approach, analyzing a wide range of conditional and joint weight and degree distributions of connected nodes, both numerically in real-world networks and mathematically in the generative graph models. We provide several methods of visualizing the results. We show that the generative models are assortativity-neutral, while many real-world networks are not. Therefore, we also propose an extension of the GIRG model which retains the manifold desirable properties induced by the degree distribution and the latent space, but also exhibits tunable assortativity. We analyze the resulting model mathematically, and give a fine-grained quantification of its assortativity.

OPEN ACCESS

Citation: Kaufmann M, Schaller U, Bläsius T, Lengler J (2026) Assortativity in geometric and scale-free networks. *PLOS Complex Syst* 3(4): e0000097. <https://doi.org/10.1371/journal.pcsy.0000097>

Editor: Manlio De Domenico, University of Padua: Universita degli Studi di Padova, ITALY

Received: July 4, 2025

Accepted: February 27, 2026

Published: April 20, 2026

Copyright: © 2026 Kaufmann et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All our code and data are available on Github and Zenodo, which allows for all experiments to be reproduced and provides all relevant parameter settings. Code: <https://github.com/thobl/assortativity> <https://zenodo.org/records/16746826> Data: <https://zenodo.org/records/16745980>.

Funding: M.K., J.L. and U.S. gratefully acknowledge support by the Swiss National Science Foundation [grant number 200021 192079]. T.B. gratefully acknowledges support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) -- grant number 524989715. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

The degree of a node in a network, that is, the number of other nodes it is connected to, is a simple measure of “importance” within the network. Whether nodes connect predominantly to similarly important “peers”, or whether there exist hierarchies where less important nodes connect to much more important nodes, has far-reaching consequences for processes that occur within networks as well as for the underlying network structure. This property is often called “assortativity” and typically measured by computing a single numerical value for a network. A lot of information is lost in this process and, to make matters worse, the most widespread way of computing this value has severe statistical flaws. We provide new evidence of these flaws and instead propose a local approach to measuring assortativity, which studies how the degree distribution of a node’s neighbors changes with this node’s degree. We further propose a “tunable” model, which allows to adjust the wiring preferences of nodes based on the degrees of potential neighbors, while at the same time capturing many established structural properties of real networks. We evaluate our new assortativity measure both on real-world networks and theoretical network models including our new tunable models.

1 Introduction

The study of network topologies has been a focal point of network science since its inception. A central feature of social and other networks is the *assortativity* or *homophily* of its nodes. (In keeping with the common naming conventions in graph theory, in the more technical sections we often use the terms *node* and *vertex* interchangeably.) This is the tendency of nodes to form connections with other nodes that have similar properties. For example, a member of a social network will have many connections with people who share profession, hobbies, or kinship with her.

While this feature is interesting in its own right and has been the subject of intensive studies [1–8], it can also be utilized in order to construct artificial networks whose structure mirrors the structure of a large class of real-world networks. The idea is to use a latent space in which the nodes are located. The purpose of this latent space is to model hidden features of the nodes, like interests or profession for social networks, but also geographical distance. Distances in the latent space correspond to similarity of the nodes. Each node also randomly draws an (approximate) degree, which we will refer to as *weight*. This is drawn from a heavy-tailed distribution as they are observed in social networks [9], typically from a power law. Then, nodes form connections, typically referred to as *edges*, based on their distances and their weights. Sometimes, the models have a *temperature* parameter which also allows them to form *weak ties* [10], i.e., nodes can randomly form edges even if not mandated by their distance and weights. This paradigm for generating artificial complex networks has been introduced and re-discovered by many different research communities,

under names like *Geographical Threshold Models* [11], *Hyperbolic Random Graphs* [12], *Scale-Free Percolation* [13], and *Geometric Inhomogeneous Random Graphs* (GIRGs) [14,15] and its variations like MCD-GIRGs [16].

The resulting complex network models have been found to be very expressive. They do not only faithfully reproduce the structural features that are explicitly built in, like sparsity and a power-law degree distribution. They also show a large number of other emerging structural properties that have also been observed in real-world networks. Those include the component structure consisting of one giant connected component and many very small connected components [15], ultra-small graph distances [15] (the *small world property* [15,17]), high clustering coefficients [15], closeness and betweenness centrality [18], small vertex and edge separators [16,19], small entropy per edge [14], high compressibility [14], and a very rich structure of overlapping communities of all sizes [14].

The network models could also be used to explain why some algorithms perform much better on real-world networks than in worst-case instances, including algorithms for computing graph distances, diameters, vertex covers, maximal cliques, and Louvain's algorithm for community detection [20]. We highlight in particular the surprising fact that bidirectional breadth-first search can find the graph distance (hop distance) between two nodes in time that is not only sublinear in the graph size, but that variations of the algorithm need to explore fewer nodes than the largest degree on a shortest connecting path [21]. Another remarkable feature of these network models is that greedy routing schemes that rely only on local information work extremely well and manage to send packages along (almost) shortest paths between the source and the target node [22–24]. This mirrors the outcome of Stanley Milgram's small-world experiment, which became famous under the name *six degrees of separation* [25]. The same complex network models have also been argued to provide a better fit for epidemiological spread of diseases than other models [26,27], and have been used to evaluate potential intervention strategies [28,29]. A particularly welcome feature is that the number of weak ties can be explicitly controlled in the models and influences many of the aforementioned results [18,20,26–28], which reflects the central role of weak ties in many empirical studies [10].

All those similarities with real-world networks emerge without specifying which properties of the nodes are modeled by the various dimensions of the latent space. In fact, such a specification is often difficult, or plainly impossible. While it is sometimes possible to identify a few dimensions in social networks which affect edge formation, it is hardly possible to identify *all* such dimensions, as would be necessary for completely fitting a model to a real network (in some cases reconstruction of the underlying topology has been possible, for example for the autonomous systems of the internet network [22]), but in general this is a very challenging task, and it has been estimated that the dimensionality of the unknown latent space is between 5 and 10 in many cases [30]. Even if some important dimensions are abstractly known, often the according data is not available for the individual nodes. Such information is often even absent for “hard” data like geographic location, and it is very rare for “soft” information like interests and hobbies. Fortunately, many structural insights about complex networks and processes on them can be obtained without such a tight fit of individual nodes [20,21,26,27,29].

While the use of a latent geometric space to model assortative tendencies of the nodes has generally been very successful, the aforementioned network models explicitly exclude a specific type of assortativity, namely *degree assortativity*. Node degree, defined as the number of incident edges, is a prominent feature. Degree assortativity measures whether nodes of low degree tend to have a disproportionately large fraction of their connections to other low-degree nodes, and conversely for high-degree nodes. In fact, most empirical studies on assortativity in network science focus on degree assortativity. This has the simple reason that degrees can be directly inferred from the network structure and are thus always available, while other properties of nodes are usually not available. For the rest of the paper, we will focus on degree assortativity.

Contrarily to most other node properties, positive assortativity (homophily) of *degrees* is not universally found in complex networks. Some such networks also show disassortativity (heterophily) of degrees. A common rule of thumb is that social networks are often positively assortative, whereas biological and technological networks tend to show negative

assortativity [1]. (Note that we can *not* confirm that social networks are positively assortative with respect to the degrees. Rather, we find two opposing trends, a positive and a negative one, and either of the two can dominate. We discuss those findings in Sect 6.2.2) Describing the assortative tendencies of degrees in a network is a fundamental goal of network science in its own right, but they have also been shown to have far-reaching consequences on community structure, network robustness and a multitude of network dynamics:

- Disassortative networks exhibit a higher epidemiological threshold, whereas assortative networks allow for longer periods of intervention before an epidemic spreads in the BTW sandpile model of distress propagation [3].
- Assortativity decreases network robustness [4].
- Assortativity in multilayer networks - connecting hubs of one network with the hubs in another network - can enhance cooperation under adverse conditions [5].
- Assortative degree correlations can strongly improve the sensitivity for weak stimuli in the brain [31].
- Conversely, dynamical processes can also influence the assortativity of an underlying network [6].
- Degree-degree correlations may cause the failure of heuristic algorithms [32].

1.1 Our contributions: Informal discussion

This paper makes several key contributions of varying nature. They range from rigorous mathematical theorems to experimental results. Thus, their level of technicality varies substantially - and as a consequence they may appeal to different audiences. To make our paper accessible, we first discuss all our main results informally, avoiding technical details as far as possible and only providing intuition. Afterwards, we go through the results again. In this second pass, we are still brief and only present the key results, but we formulate them more rigorously and with additional technical details, so that readers can opt to skip the technical parts if desired.

Negative result on Pearson assortativity coefficient. As motivated above, we need good and robust ways to describe and understand assortativity of degrees. However, assortativity is a complex concept and is not easy to measure. The most common approach is to compute a single real-valued coefficient, whose sign then indicates whether the network is assortative (positive sign) or disassortative (negative sign). The most commonly used coefficient is the Pearson assortativity coefficient. However, it has been challenged whether this coefficient is a good measure for assortativity [33]. In this paper we give a negative answer in the strongest possible sense, which is our first main result. We show that for every network with a sufficiently heavy-tailed degree distribution, the Pearson assortativity coefficient is negative, regardless of the assortative behavior of the network. This confirms mathematically that this coefficient does not measure assortativity of degrees if the degree distribution is heavily skewed.

Other coefficients and a new lens on assortativity. There are other coefficients than the Pearson assortativity coefficient which are better suited, in particular Spearman's and Kendall's rank correlation coefficients [33]. Another contribution of this paper is that we investigate those coefficients for a selection of networks from the KONECT database [34]. However, we argue that, in many of these cases, reducing assortativity to a single number is too simplistic. These concerns have been raised in the literature before [35,36]. Instead, we propose an alternative way to study assortativity. In a nutshell, we suggest to consider a random edge $e=\{u,v\}$, and to study how the degree distribution of v changes when we condition on the degree of u . We explore several ways in which this information can be graphically illustrated. We will discuss all alternatives and the full technical details in Sect 6, but already show an example here for illustration in Fig 1. Further examples both for artificially generated networks and for real networks from the KONECT database can also be found in Sect 6, as well as alternative ways for graphical illustration. We note here that other *local* approaches to measuring

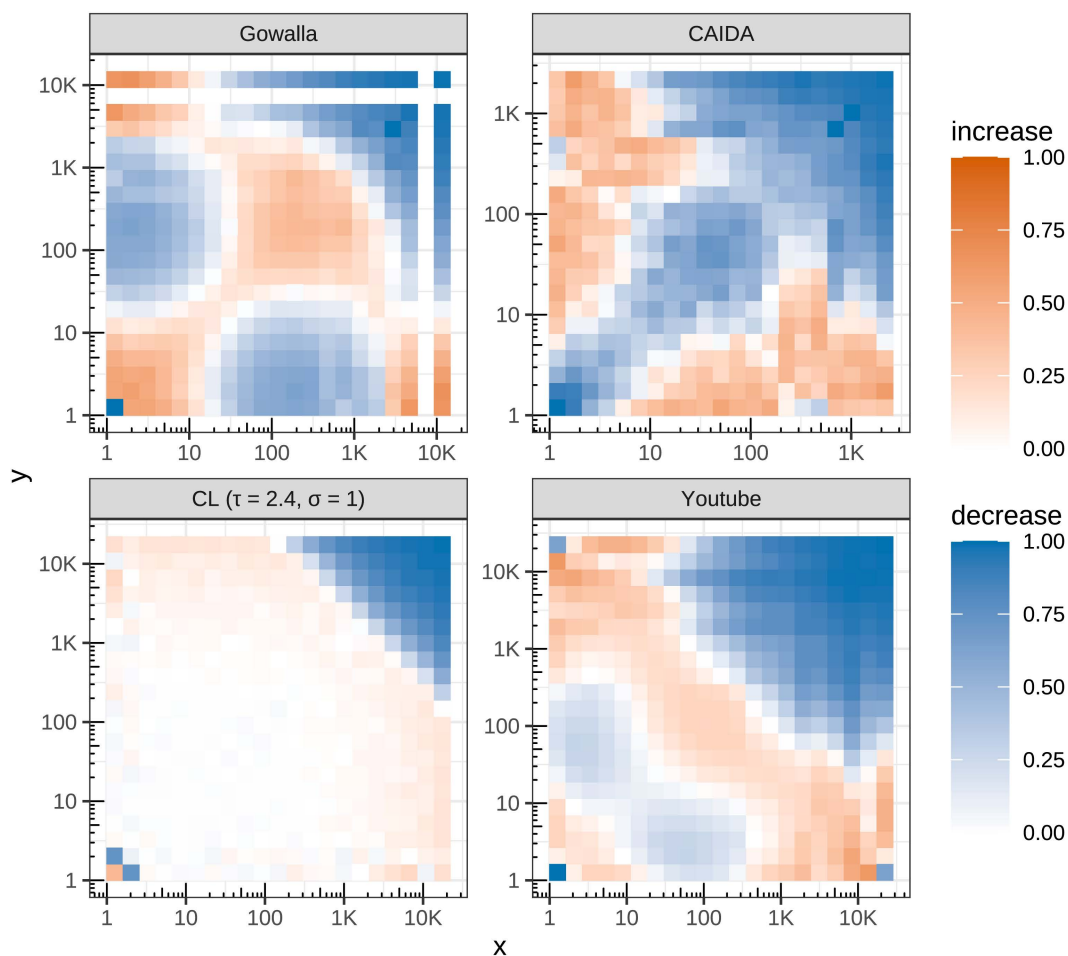


Fig 1. Four heatmaps displaying the effect on a random edge $e=\{u,v\}$ when conditioning on the degree of one endpoint. A red color at coordinate (x,y) indicates that the information $\text{deg}(u) = x$ makes it more likely that the other endpoint of e has degree $\text{deg}(v) = y$. A blue color shows that this becomes less likely. All networks follow a power-law degree distribution. Details on the networks are provided in Sect 6. The Gowalla social network (top-left) shows mostly positive assortativity: the diagonal is red, so nodes of the same degree tend to connect to each other, while most regions elsewhere are blue. The CAIDA Autonomous Systems Network (top-right) shows strong negative assortativity. Two networks (bottom row) with similar degree distributions. Both would be classified as assortativity-neutral by Spearman's and Kendall's assortativity coefficients because both coefficients are close to zero. However, while the artificially generated network (bottom-left) is indeed neutral (the blue triangle is inevitable in power-law networks, see Sect 1.2.3), the Youtube Network (bottom-right) has both positive and negative assortative tendencies which cancel each other out. The heatmap is thus more informative than a single coefficient.

<https://doi.org/10.1371/journal.pcsy.0000097.g001>

assortativity have been tried before [37] (for a survey, see also [38]). Typically, these approaches consider a node's contribution to a global coefficient and thus the drawbacks of coefficients remain unsolved.

Analyzing assortativity in generative models of complex networks. As mentioned before, many complex network models with an underlying latent space do not include degrees in this latent space, but handle them separately. (Some models, like Hyperbolic Random Graphs, do include the weights into the latent hyperbolic space, but in a separable way. Our conclusion still applies to these models: they are mostly assortativity-neutral, and the spurious positive assortativity comes from the same source as for the GIRG model.) This allows to prescribe the degree distribution, for example as a power law. Thus, while the resulting networks are assortative with respect to the features modeled by the latent space, it has been unclear whether their degrees are also assortative. As our third main contribution, we answer this question

for the GIRG model, one of the most prominent complex network models with an underlying geometry and an inhomogeneous degree distribution. We show that it is mostly, but not totally, *assortativity-neutral*. Moreover, we can pinpoint exactly where the small (positive) assortativity comes from. It originates from random fluctuations in the number of nodes with similar properties. Consider that – by chance – there are more nodes than expected with some specific property, i.e., there are more nodes than expected in the same region X of the latent space. Then the nodes in X have higher degrees, but also their neighbors tend to be in X and thus also have higher degree. This gives a small positive contribution to assortativity of degrees. The effect is relevant for nodes of very small degree, but is negligible for nodes of medium or large degree. If we remove this effect by removing the latent space, we show that the resulting *Chung-Lu model* [39] is perfectly assortativity-neutral. Here we measure neutrality not just by a single real number, but in the strongest possible sense: we show that for a random edge $e=\{u,v\}$ the whole probability distribution of $\deg(v)$ is invariant under information about $\deg(u)$. (Except for a ceiling effect for extremely large degrees that we quantify precisely.) This can also be visually seen in [Fig 1\(c\)](#), which displays such a Chung-Lu random graph.

We provide a formal statement that pinpoints in which sense this neutrality is retained in GIRGs. More precisely, we show that the GIRG model is still perfectly assortativity-neutral with respect to *weights*, but not to degrees. Weights are internal parameters of the model that are tightly, but not perfectly, coupled to the degrees. We can show that the only non-neutral contribution, a positive one, to assortativity of degrees comes from this slight mismatch between weights and degrees. If we isolate this effect by removing the degree inhomogeneity (since the effect is negligible for nodes of large degrees) and retaining the latent space, then we can exactly quantify the assortativity of the resulting *Random Geometric Graph* model. Thus we can split the construction of latent space models like GIRGs into two steps, one of which is perfectly assortativity-neutral, while the other one gives a mild positive assortativity that we can quantify.

Tuning the assortativity of complex network models. As our analysis shows, established models of complex networks such as GIRGs are mostly assortativity-neutral for degrees. This raises the question of how we can build models with a heavy-tailed degree distribution and a latent space (in order to maintain assortativity for *other* properties than degrees), albeit with tunable assortativity of degrees. Our final contribution is to present exactly such a model. We show that we can slightly alter the connection probability of the GIRG model in such a way that assortativity of degrees can be tuned to be either positive or negative, while keeping the basic properties of the model intact. These properties include the tight coupling between weights and degrees, which maintains the prescribed degree distribution, but also a high clustering coefficient, small-world properties, and so on. As for the GIRG model, we can split the degree assortativity into two parts: a minor part coming from mismatches of weights and degrees of low-degree nodes, as before; and a major part that can be described via assortativity of the weights. We precisely quantify this latter part and show that the (positive or negative) assortativity of the model can indeed be tuned by a parameter.

1.2 Our contributions: Technical summary

In this section we elaborate on our contributions from a technical perspective. In the following, assortativity always refers to the degrees of the network's nodes. We denote a neighborhood relation between two nodes u and v by $u \sim v$.

1.2.1 Networks and network models. We consider two types of networks, real-world networks and artificially generated networks. The real-world networks are taken from the KONECT database [34], where we restrict to undirected simple networks whose degrees follow a power law according to [40].

We focus on two generative models for complex networks. The first one is called *Geometric Inhomogeneous Random Graphs (GIRGs)*. In this model, n nodes are placed randomly in the latent space $[0, 1]^d$ for some fixed dimension d . (Often, the torus topology is used, i.e., each dimension wraps around, since this removes boundary effects and makes the model more symmetric.) Weights are drawn from a power-law distribution with law $\mathbb{P}(W = w) = \Theta(w^{-\tau})$ for $w \geq 1$, where the power-law exponent τ is usually in the range $\tau \in (2, 3)$. Afterwards, each pair of nodes u, v connects independently with probability $\Theta(\min\{1, \frac{w_u w_v}{n \|x_u - x_v\|^\sigma}\}^\alpha)$, where w_u, w_v are the weights and x_u, x_v the positions of u and v respectively, and the

parameter $\alpha > 1$, sometimes called the inverse temperature of the model, modulates the proportion of weak ties in the network. The $\Theta(\cdot)$ allows for constant factor deviations. This formula turns out to give exactly the right scaling, as we explain in detail in Sect 3.

The other complex network model we look at is the *Chung-Lu random graph model*. It is similar to the GIRG model, except there is no underlying latent space. So, every node draws a weight as for GIRGs, but then any pair of nodes u, v connects with probability $\Theta(\min\{1, \frac{w_u w_v}{n}\})$, with no geometric information. The two models are closely related, as the marginal connection probability is the same when randomizing over the geometric positions in GIRGs. This means that for two nodes u, v with fixed weights but *random* positions, the marginal connection probability in a GIRG is $\mathbb{P}[u \sim v \mid w_u, w_v] = \Theta(\min\{1, \frac{w_u w_v}{n}\})$, just as for Chung-Lu graphs. As a consequence, many properties transfer from Chung-Lu graphs to GIRGs. In particular, in both models the expected degree of a node coincides with its weight up to a constant proportionality factor. This leads to a degree distribution that follows the same power-law as the weights, ensuring that the neighborhood distribution (obtained by selecting a uniformly random edge e , choosing a random endpoint v of e , and then returning the degree of v) and the small-world property are conserved and remain equivalent for both models. However, due to the latent space, GIRGs also have strong clustering and communities, which Chung-Lu graphs do not have.

We also introduce a variant of GIRGs for which the assortativity coefficient can be tuned to be either positive or negative. We call the resulting model Tunable GIRG or TGIRG. The construction steps are the same as for GIRGs, except that we use a different connection probability: we connect each pair of nodes u, v with probability $\Theta(\min\{1, \frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n \|x_u - x_v\|^\sigma}\}^\alpha)$, where $\sigma \in [0, \tau - 1]$ is a parameter that determines how assortative or disassortative the model is. Here, $\sigma = 1$ yields the GIRG model and thus (almost) neutral assortativity, while smaller values of σ give disassortative networks and larger values give assortative networks.

Finally, we also give a brief analysis of assortativity in *Random Geometric Graphs* (RGG). Those graphs are similar to GIRGs, though the weights do not follow a power law, but are all set to one. Moreover, as we only use those graphs to demonstrate a point for GIRGs, we restrict ourselves to the case of $\alpha = \infty$ (zero temperature), and to the case that the hidden constant factor in the $\Theta(\cdot)$ -term is one. See the discussion of theoretical results on GIRGs below for why we discuss those graphs.

1.2.2 Experimental results. We systematically explore degree assortativity of the real networks and of artificially generated Chung-Lu and GIRG networks. We find, in accordance with our theoretical result, that the Pearson assortativity coefficient is not a good measure for assortativity and provides almost no information. Spearman's and Kendall's correlation coefficients are generally in good agreement with each other. However, we find cases in which those coefficients are close to zero even though there are strong positive and negative assortative effects in the network which happen to cancel out.

We argue that single real coefficients lose too much information, and that a more detailed approach is useful. We offer three visualizations of assortativity. We depict an example of each in Fig 2, but defer their detailed discussion, the systematic exploration of artificial and real networks, and the interpretation of the findings, to Sect 6. For additional examples of networks where intricate wiring patterns are concealed by coefficient values close to zero, we refer the reader to the figures for the Wordnet and Youtube networks in S1 File. The three visualizations are:

- A comparison of line plots showing the conditional probability $\mathbb{P}[\text{deg}(v) \geq x \mid v \sim u]$ on a log-log scale, where u is a random node of prescribed degree d . By inspecting how the curves change for different values of d , one can infer what the typical neighbors of nodes of different degrees look like. Of particular interest is the ordering of the curves and whether such curves cross each other for different d . For artificial networks, our theoretical results show that each such curve can be decomposed into two sub-parts, both of which are linear. The slopes of the two linear parts and the transition point between them are central and should be used as baselines for the curves for real networks.

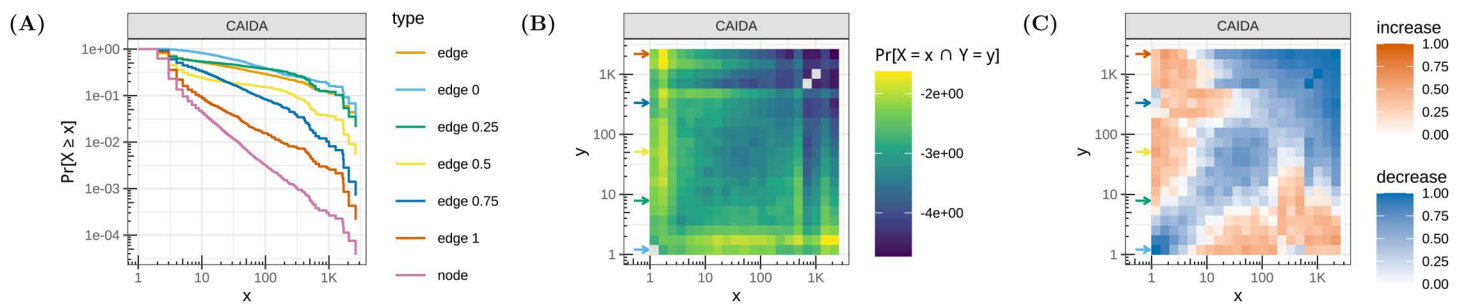


Fig 2. Conditional and Joint Degree Distributions of the CAIDA Autonomous Systems Network. (A) Degree distribution curves. (B) Joint degree heatmap. (C) Conditional degree heatmap.

<https://doi.org/10.1371/journal.pcsy.0000097.g002>

To obtain a proper baseline for such curves it is important to keep in mind the difference between sampling a random node and sampling a random edge. In the former case, one obtains the degree distribution of the network. In the latter case, one obtains the *neighborhood distribution* (equivalently, the neighborhood distribution may be obtained by choosing a uniformly random node u of degree ≥ 1 , go to a uniformly random neighbor v of u , and return the degree of v), which is shifted towards high-degree nodes. For heavy-tailed degree distributions, those two distributions are very different from each other. It is crucial to use the latter distribution as baseline, not the former one. This is closely related to the *friendship paradox* that a random neighbor of a random node has larger average degree than a random node.

- Heatmaps of joint degree frequencies, i.e., heatmaps where the coordinate (x, y) displays the proportion of edges $e=\{u, v\}$ with $\deg(u) = x$ and $\deg(v) = y$. Comparing different rows and columns of those heatmaps reveals preferences of nodes of different degrees.
- The heatmaps get easier to interpret if one uses normalization. Thus, as the third option we display the normalized values $\frac{\mathbb{P}[\deg(v)=y|\deg(u)=x]}{\mathbb{P}[\deg(v)=y]}$ for a random edge $e=\{u, v\}$. The exact definition is slightly more complicated and involves a case distinction for whether the fraction is larger or smaller than one. We defer the details to Sect 6.

A detailed discussion of our findings for social and technical networks can be found in Sect 6. The most important points are:

- All social networks that we consider show an increased likelihood for edges between pairs of low-degree nodes. This is very similar to the effect of adding a latent geometric space to the artificial models, see below. This indicates that such a geometric space may be beneficial for modeling social networks.
- Beyond those pairs of low degrees, there are very different assortativity patterns in social networks. We find at least two types of different patterns. Those can contribute positively or negatively to an assortativity coefficient. If they contribute negatively, then we have opposing contributions to the coefficient, and either contribution may dominate, or contributions may cancel out, so that the complex opposing wiring patterns are invisible through the lens of assortativity coefficients. In general, the Spearman and Kendall assortativity coefficients of social networks range from strongly negative to strongly positive.
- The assortativity patterns for technical networks of autonomous systems strongly differ from those of social networks. They are usually negatively assortative. Other than social networks, those networks do not generally show an increased likelihood in connections between pairs of low-degree nodes. One possible source of strong disassortative patterns in technical networks may stem from underlying hierarchies where connections are formed between different hierarchical levels.

- We can match the different assortativity patterns in real networks to different parameter regimes in the tunable models of artificial networks that we propose (see below).

1.2.3 Theoretical results. Negative result for the Pearson assortativity coefficient. Our first theoretical result is showing that the Pearson assortativity coefficient is not a good measure for assortativity. This has already been argued before [33,40,41], but we give a strong theoretical result:

Theorem 1.1. *Let $2 < \tau < 7/3$ and let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph on n vertices with $\Delta := \max_{v \in \mathcal{V}} \deg(v) = \Theta(n^{1/(\tau-1)})$ such that for a vertex v chosen uniformly at random in \mathcal{V} and for all $0 < k \leq \Delta$ we have $\mathbb{P}(\deg(v) \geq k) = \Theta(k^{1-\tau})$. Then its Pearson assortativity coefficient $r(\mathcal{G})$ is negative, more precisely $r(\mathcal{G}) = \Theta(-n^{-(\tau-2)/(\tau-1)})$.*

This result shows that the Pearson assortativity coefficient does not measure assortativity for power-law networks with a sufficiently heavy tail. It always gives the same result, regardless of how the nodes in the network are wired.

The reason for this result is visible in Fig 1. Note that in all four subplots there is a strongly blue triangle in the upper right corner. This is not an artifact of the networks, but is inevitable in all networks which follow a power-law degree distribution. Such networks simply have too few vertices of large weights, so that the required number of neighbors of large weight for neutrality in this region is impossible to reach. All assortativity coefficients compute a weighted average of the different regions depicted in Fig 1, but the Pearson coefficient puts particular weight on this upper right region, which forces it to be negative if the power-law degree exponent τ is smaller than $7/3$.

Weight assortativity in Chung-Lu graphs and GIRGs. We investigate the assortativity of weights in Chung-Lu graphs and GIRGs. Note that the weights are closely coupled to the degrees, but they are not identical, and we discuss the difference further below. We show that both Chung-Lu graphs and GIRGs are assortativity-neutral for weights, except for a cut-off for extremely large weights. More precisely, consider a vertex u with weight $w_u \leq n$, and consider a random neighbor v of u . Then the conditional weight distribution of v is given by

$$\mathbb{P}[w_v \geq w \mid w_u, v \sim u] = \begin{cases} \Theta(w^{1-\tau}) & \text{if } w_u \leq \frac{n}{w}, \\ \Theta\left(\frac{w^\tau n}{w_u}\right) & \text{if } \frac{n}{w} < w_u \leq n. \end{cases} \quad (1)$$

The full formulation of the statement can be found in Proposition 4.3 (by setting $\sigma = 1$ in the statement). Note that the second case is very exceptional. Most nodes have small (constant) weights, and the maximal weight in the graph is $\Theta(n^{1/(\tau-1)}) = \alpha(n)$. So for most nodes u , the second case is void. Only if the weight w_u is very large, the second case starts occurring for large values for w . For example, if u has an exceptionally large weight of $w_u = \Theta(\sqrt{n})$, then the tail distribution of w_v changes at $w = \Theta(\sqrt{n})$, which again is a very small part of the tail.

The key insight about (1) is that the formula in the first case is independent of w_u . In other words, the distribution of w_v in the neighborhood of u does not depend on the weight of u , except possibly for a cut-off at very large weights. (The reason for the cut-off is that above the cut-off the number of vertices with so high weight in the graph is simply too small to satisfy the formula in the first case, the same effect as discussed in the previous section.) This means that, up to the exception at the cut-off, the models are perfectly assortativity-neutral with respect to the weights.

Degree assortativity in GIRGs and RGGs. We have established that *weights* are assortativity-neutral in GIRGs. If the weight is super-constant, then it is so tightly coupled to the degree that the degrees are necessarily also assortativity-neutral. However, for constant weights the coupling is not so tight, and in this case there is a positive assortative contribution from random fluctuations of the vertex locations. To describe this effect, we strip the GIRG model of nodes of larger weights, which yields the RGG model. Moreover, in order to keep the exposition simple we only consider the case of zero temperature. In Proposition 5.1, we derive an exact formula for the assortativity in this resulting model. Essentially, for a vertex u with given degree, the degree of a random neighbor v of u is the sum of two independent Poisson random variables. One of them is independent of $\deg(u)$, while the expectation of the other is proportional to $\deg(u)$, which yields a positive constant assortativity.

We believe, without formulating a mathematical theorem, that assortativity in GIRGs is best regarded as the result of two consecutive processes: for a random vertex u the *weight* of a random neighbor v is not coupled to w_u . However, the translation of weight into degree gives positive assortativity, modulated by exactly the same process as for RGG. This modulation is only relevant for constant weights and degrees. This is also strongly confirmed by experiments, which show positive assortativity for constant degrees, and neutral assortativity otherwise.

1.2.4 TGIRG: A complex network model with tunable assortativity. As our final contribution, we propose a complex network model that builds on the GIRG model and inherits its wealth of structural properties, and at the same time can be tuned to have either positive or negative assortativity. The idea is to use a slightly different probability kernel. In the GIRG model, the connection probability increases with the product of the weights $w_u \cdot w_v$. This product can equivalently be written as $(w_u \wedge w_v) \cdot (w_u \vee w_v)$. The idea is now to scale up or down the smallest of those two factors with an exponent $\sigma \in [0, \tau - 1)$. (This range is chosen so that the expected degree of a vertex is of the same order as its weight as in the original GIRG model, see Lemma 3.2.) This follows an idea of [42], and in fact this model has been studied before [43,44], but not with respect to assortativity.

In this paper we show two things: firstly, we analyze basic properties of the resulting TGIRG model, including its marginal connection probability (Proposition 4.2). We also show that it shares properties of the GIRG and the Chung-Lu model, in particular that it has the same neighborhood distribution as both (Proposition 4.1) and that it has a large clustering coefficient of $\Omega(1)$ as GIRGs (Proposition 3.6). And secondly we show that assortativity of the resulting model is indeed tunable. More precisely, let us fix as before a vertex u of fixed weight w_u , and let v be a random neighbor of u . Then we show the following formula, where for simplicity we ignore cases that only hold for exceptionally large weights (the full statement can be found in Proposition 4.3):

$$\mathbb{P}[w_v \geq w \mid w_u, v \sim u] = \Theta(w^{1-\tau}(w \wedge w_u)^{\sigma-1}). \tag{2}$$

In order to understand the formula, let us first consider $\sigma > 1$ and use the formula $\Theta(w^{1-\tau})$ of the assortativity-neutral GIRG model as baseline. Compared to the baseline, the probability is boosted by a factor $(w \wedge w_u)^{\sigma-1}$. For $w \leq w_u$, this factor simplifies to $w^{\sigma-1}$, and it is capped at $w_u^{\sigma-1}$ for $w > w_u$. The first thing to note is that this factor is increasing in w , so larger weights w get boosted more. Moreover, for larger w_u the boost is stronger in two ways: firstly there is a prolonged interval $[1, w_u]$ in which the boost factor is increasing; and secondly the maximum boost factor $w_u^{\sigma-1}$ is higher. Both effects imply that for larger weight w_u , the distribution puts more emphasis on the tail of the distribution. Thus the resulting model is assortative.

On the other hand, if $\sigma < 1$ then the exponent $\sigma - 1$ is negative. This reverses the effect and makes the network disassortative. (Note that this is assortativity of *weights*. The same effect as for GIRGs will add a small positive assortative effect for *degrees* that is only relevant for small degrees.) We confirm these findings by experiments. Fig 3 illustrates the effects of increasing σ on a graph with the same vertex set and approximately the same average degree.

1.3 Related work

In this subsection, we give an overview of related work, including known results on TGIRGs.

Origins of the line of research and early results. The study of assortativity through the lens of network science was initiated by Newman and various collaborators in a series of papers [1,2,45]. In these early works, they introduced the Pearson correlation coefficient as a measure for network assortativity and observed experimentally that social networks tend to be positively assortative, while biological and technological networks exhibit negative assortativity [1]. Newman also computed the (vanishing) limits of the Pearson correlation coefficient for Erdős-Rényi random graphs and preferential attachment graphs as well as for the grown graph model of Callaway. In the same work, he also proposed a positively assortative network model, for which he runs simulations according to which the phase transition for the existence of a

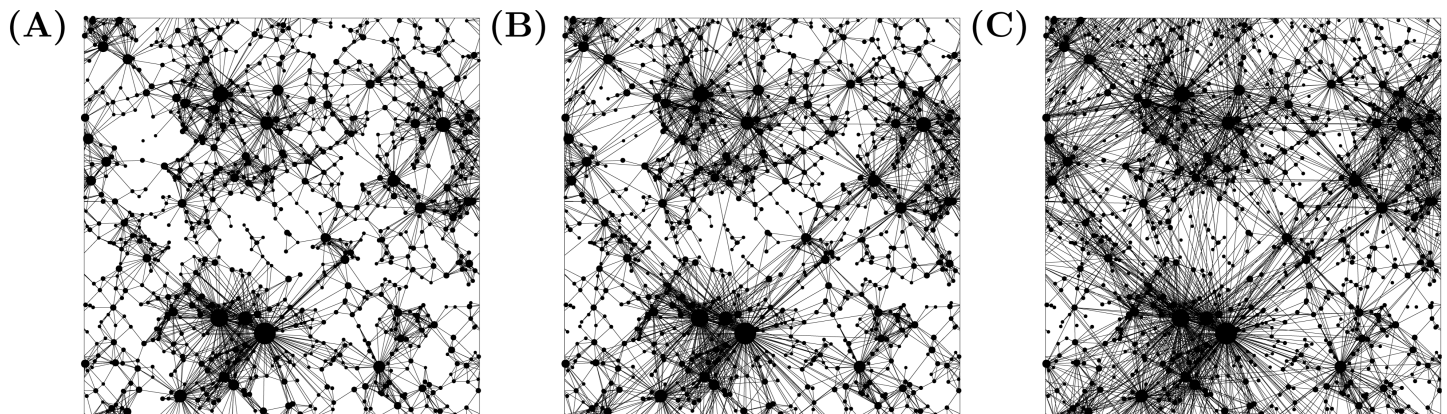


Fig 3. TGIRGs with $n=1000$, $\tau = 2.8$, $\alpha = \infty$, and average degree 10 on the same vertex set for varying σ : (A) $\sigma = 0.2$, (B) $\sigma = 1$, (C) $\sigma = 1.8$.

<https://doi.org/10.1371/journal.pcsy.0000097.g003>

giant component occurs earlier with increasing assortativity. He further observed numerically that contrarily to other networks, the attack strategy of removing a small number of high-degree nodes in order to disconnect the giant component is relatively inefficient in assortative networks because high-degree nodes cluster together [1]. Degree-degree correlations, as well as non-degree assortativity were studied in [2], including mixing by ethnicity or age in heterosexual partnerships, indicating strong positive correlations. The paper also proposes Monte-Carlo algorithms for generating networks of a desired assortativity level. Simulation results again point to an earlier emergence of the giant component with increasing assortativity. At the same time, the size of the giant component is smaller for assortative graphs than for neutral and disassortative instances. This is credited to the fact that the well-connected core of high-degree nodes already leads to a giant component at lower densities, but these lower densities also make it less likely that the giant component extends to regions in the graph outside this core. (Non-degree) assortativity was subsequently investigated as a plausible cause of emerging communities in networks [46], accompanied by an algorithm based on the betweenness centrality measure, which was later applied successfully also to non-human social networks [47]. Newman and Park demonstrated that the reverse implication connecting communities and assortativity sometimes also holds: community structure can lead to degree-degree correlations [48]. These early research articles already illustrate the difficulty of finding models which are both tunable and mathematically tractable.

Assortativity tuning. Numerous approaches have been tried in order to generate networks with varying levels of assortativity. The most general procedures are the ones introduced by Newman [2] and Boguñá and Pastor-Satorras [49] which devise two different schemes to construct general correlated networks with prescribed correlations. Simpler approaches focus on imposing merely the intuitive requirement that “nodes with similar degrees connect preferably”, instead of “hard-coding” the desired correlations [50]. Another classical procedure relies on an idea from the so-called configuration model [51], which allows to generate networks with a desired degree sequence by assigning to each node the number of “stubs” that correspond to its desired degree and then randomly connecting stubs. This idea can be extended to a rewiring procedure where pairs of edges uv and rs are selected randomly replaced by ur and vs if and only if the assortativity of the network is increased (or decreased) through this rewiring [2,50,52]. Finally, there exist various modifications of the preferential attachment model, which was originally introduced by Barabási and Albert [9]. In this model of a *growing network*, nodes are added one-by-one, connecting to existing nodes with a probability that is proportional to their respective degrees. I.e., the “rich get richer” and newly arriving (and therefore low-degree) nodes connect preferably to nodes of high degree. By modifying the attachment rule, one can directly change the wiring preferences of nodes [1,53]. The iterative generation of such networks introduces a lot of dependencies which make the rigorous mathematical analysis of networks

difficult. Although we focus here on degree-assortativity, we mention that for attribute-assortativity, besides the latent space approach that underlies GIRGs, there exist several mechanisms which allow to adjust the assortativity levels of a network, such as the stochastic block model and its variants [54–56].

Structural and process implications of (dis)assortativity. A key motivation for studying assortativity of networks is the belief that the level of assortativity can have far-reaching structural consequences for the network in question as well as the evolution of dynamical processes on it. Structurally, degree assortativity is said to lead to tightly-knit cores of influential nodes, often referred to as a “rich-club” [57]. We remark here that such a well-connected core consisting of high-degree nodes can also be produced by other mechanisms, as exhibited by GIRGs [15]. Another assertion echoed in many papers is that network robustness is greatly affected by the (dis)assortativity of a network [58,59]. Dynamical processes on networks have also been studied intensively through the lens of assortativity [60]. This has led to the observation that degree correlations impact the epidemic threshold [61] and the development of models for spreading which take assortativity into account [38]. Synchronization phenomena in complex networks have also been investigated [62,63].

Challenges of research on assortativity. Since the proposed models are often tailored to explain a single structural aspect of real-world networks or study a specific dynamical process, this has resulted in a wealth of otherwise sparsely studied models which are difficult to compare. From a practitioner’s perspective, any insight relating degree-degree correlations to structural properties such as network robustness, which is crucial in critical infrastructures, also requires that the examined model captures more than just one aspect of real-world networks in order to be valid. As outlined in the introduction, GIRGs jointly exhibit many such desirable properties stemming from the latent geometric space. As we discuss in the next subsection, a lot is already known about our proposed extension of GIRGs as they fit well into existing frameworks. For the description of further properties of TGIRGs, we refer the reader to Sect 3; most notably they have a large local clustering coefficient of $\Theta(1)$.

1.3.1 Related work on TGIRGs. A variety of properties of TGIRGs have been shown for different ranges of σ . Some of these properties have been derived using a different parametrization of the model, known as the weight-dependent random connection models with the interpolation kernel (see for example [43]). Our choice of parametrization has the advantage that weights are intuitive proxies for expected degrees and that τ directly controls the decay of their power-law distribution, whereas the alternative parametrization only provide this indirectly. There also exist parametrizations which assign different exponents σ_1 and σ_2 to the smaller and larger weight in the connection probability. These can however always be reparametrized (as long as $\sigma_1 > 0$) to give $\sigma_1 = 1$ by changing the power-law exponent τ accordingly [42].

We first give the conversion formulas between our parametrization (Definition 2.6) and the alternative parametrization, which uses marks in the interval $[0, 1]$ instead of power-law distributed weights. Apart from the dimension d of the underlying space (for which the mark parametrization does not differ from ours), three parameters are used, $\gamma_M, \delta_M, \alpha_M$, where the subscript M serves to distinguish the parameters of the mark parametrization from the parameters of our parametrization (in particular M since both parametrization use the symbol α , but for different purposes). The conversion formulas between the two parametrizations are as follows:

$$\begin{aligned} \tau = 1 + \frac{1}{\gamma_M} &\iff \gamma_M = \frac{1}{\tau-1}, \\ \alpha &= \delta_M, \\ \sigma = \frac{\alpha_M}{\gamma_M} &\iff \alpha_M = \frac{\sigma}{\tau-1}. \end{aligned}$$

The existence and uniqueness of a giant component is well understood [43, Proposition 2.4 and Corollary 2.14]:

- If $\sigma > \tau - 2$ or $\tau < 2 + \frac{1}{\alpha}$, whp (we say an event occurs *with high probability (whp)* if it occurs with probability tending to 1 as n tends to ∞) the graph contains a unique giant component.
- If $d=1, \alpha > 2, \tau > 2 + \frac{1}{\alpha-1}$, and $\sigma < \tau - 2$, whp the graph does not contain a giant component.

- For other parameter regions, the presence of a giant component depends on the constants hidden in the Θ -notation of the connection probability in (8).

Additionally, the cluster size decay of TGIRGs is studied in [42,64]. Regarding the average distance in the giant component, unlike with standard GIRGs for which this is always doubly logarithmic in the number of nodes, various scaling regimes can occur, depending on the model parameters. When $\sigma \in [0, \tau - 2]$, TGIRGs satisfy Assumptions 1.1 and 1.2 in [65], and hence Theorems 1.1 and 1.2 therein apply, which yields that the average distance is whp $\frac{4+\sigma(1)}{\log^{1/(\alpha(\tau-2))}} \log \log n$ if $\tau < 2 + \frac{1}{\alpha}$ and $\omega(\log \log n)$ if $\tau > 2 + \frac{1}{\alpha}$. When $\sigma \in [1, \tau - 1)$, the proof in [15] can be adapted (the proof is for GIRGs, i.e., for $\sigma = 1$, and for larger σ the connection probability between two given vertices is larger than for GIRGs), yielding a doubly logarithmic average distance in this case as well. On the other hand, if $\alpha > 2$, $\tau > 2 + \frac{1}{\alpha-1}$, and $\sigma < \tau - 2$, then the average distance is linear in the number of nodes [66]. In some parameter regimes, TGIRGs also exhibit an average distance that is neither linear nor ultra-small. For example, the connection probability in TGIRGs is lower bounded by the connection probability in long-range percolation, where the average distance is at most polylogarithmic when $\alpha < 2$ [67], yielding at most polylogarithmic distances for any TGIRG with $\alpha < 2$. This holds for any choice of τ , σ and d , including choices where the average distance is $\omega(\log \log n)$.

Organization of the paper. The remainder of the paper is organized as follows. In Sect 2, we formally introduce the graph models we analyze and state some of their key properties in Sect 3. In Sect 4, we derive a comprehensive set of conditional and joint weight distributions for vertices of Chung-Lu graphs and GIRGs. In Sect 5, we prove that degrees of neighboring vertices in Random Geometric Graphs are positively assortative and discuss degree-degree correlations in Chung-Lu graphs and GIRGs. Sect 6 contains experimental evaluations for a variety of real-world networks and for artificially generated networks from the network models. In Sect 7 we prove Theorem 1.1, which shows that the Pearson correlation coefficient is negative for every networks whose degree distribution has a sufficiently heavy tail. We also discuss the use of alternative coefficients and report those coefficients for real-world networks.

2 Graph models

We denote by $[n]$ the set $[n] := \{1, \dots, n\}$. We will often use the notations

$$x \wedge y := \min\{x, y\} \quad \text{and} \quad x \vee y := \max\{x, y\}$$

to denote minima and maxima. We consider simple undirected graphs with vertex set $\mathcal{V} = [n]$ and edge set denoted by \mathcal{E} . The set of neighbors of a vertex $v \in \mathcal{V}$ is denoted by $\Gamma(v)$. We use the short-hand notation uar for (events occurring) uniformly at random and iid when random variables are independent and identically distributed. In most graphs models we study, the degrees are distributed according to a power law, which is defined as follows.

Definition 2.1. Let $\tau > 1$. A discrete random variable X with values in \mathbb{N} is said to follow a power-law with exponent τ if $\mathbb{P}(X = x) = \Theta(x^{-\tau})$ for $x \in \mathbb{N}$. A continuous random variable X with values in $[1, \infty)$ is said to follow a power-law with exponent τ if it has a density function f_x satisfying $f_x(x) = \Theta(x^{-\tau})$.

Our focus will be on scale-free graph models, i.e., graphs that have a power-law degree distribution with exponent $\tau \in (2, 3)$. However, for the sake of generality, we give the definitions for any choice of $\tau > 2$. We start by defining Chung-Lu graphs, which are standard graphs yielding such a power-law degree distribution.

Definition 2.2 (Chung-Lu graph [39]). Let $\tau > 2$ and let \mathcal{D} be a power-law distribution on $[1, \infty)$ with exponent τ . A Chung-Lu graph is obtained by the following two-step procedure:

- (1) Every vertex $v \in \mathcal{V}$ draws iid a weight $w_v \sim \mathcal{D}$.
- (2) For every two distinct vertices $u, v \in \mathcal{V}$, add an edge between u and v in \mathcal{E} independently with probability

$$\Theta\left(\min\left\{\frac{w_u w_v}{n}, 1\right\}\right). \tag{3}$$

Classically, the $\Theta(\cdot)$ simply hides a factor 1, but this more general definition of the model also captures similar random graphs, like the Norros-Reittu model [68], while important properties stay asymptotically invariant [69].

Very frequently, real-world networks have an (implicit) underlying geometry. The Random Geometric Graph (RGG) model is the simplest model that includes geometry. In RGGs, each vertex is assigned coordinates in an underlying ground space. Pairs of vertices are then connected independently of other pairs if their distance is below a global threshold [70]. We will take the d -dimensional unit hypercube $[0, 1]^d$ equipped with the torus topology as the ground space. In particular, we define the distance between two points $x = (x_1, \dots, x_d), x' = (x'_1, \dots, x'_d)$ in $[0, 1]^d$ as

$$\|x - x'\| := \max_{1 \leq i \leq d} \min\{|x_i - x'_i|, 1 - |x_i - x'_i|\}. \tag{4}$$

We remark here that, while in the subsequent we fix the norm for our models to be the max-norm, all our results, except for the expected volume of the intersection of balls of influence in Proposition 5.1, also hold for any other choice of norm.

Definition 2.3 (RGG). Fix a node set \mathcal{V} of order $|\mathcal{V}| = n$ and a function $r = r(n) = \Theta(n^{-\frac{1}{d}})$. A threshold Random Geometric Graph is obtained by the following two-step procedure:

1. Every node $v \in \mathcal{V}$ draws independently and uar a position x_v in the hypercube $[0, 1]^d$.
2. Connect each pair of distinct vertices $u, v \in \mathcal{V}$ by an edge iff $\|x_u - x_v\| \leq r$

Note that the condition on r yields a sparse graph, that is, an expected number of edges which is linear in the number of vertices. We refer to the geometric region of points with distance at most r from a node as its *ball of influence* or *box of influence*.

Geometric Inhomogeneous Random Graphs (GIRGs) combine the degree inhomogeneity of Chung-Lu graphs with the geometric component of RGGs. The vertices are assigned both a weight and a position in a given ground space. We will again take the d -dimensional unit hypercube $[0, 1]^d$ equipped with the torus topology as ground space.

Definition 2.4 (GIRG [14]). Let $\tau > 2, \alpha > 1$ and $d \in \mathbb{N}$ and let \mathcal{D} be a power-law distribution on $[1, \infty)$ with exponent τ . A Geometric Inhomogeneous Random Graph (GIRG) is obtained by the following three-step procedure:

- (1) Every vertex $v \in \mathcal{V}$ draws iid a weight $w_v \sim \mathcal{D}$.
- (2) Every vertex $v \in \mathcal{V}$ draws independently and uar a position x_v in the hypercube $[0, 1]^d$.
- (3) For every two distinct vertices $u, v \in \mathcal{V}$, add an edge between u and v in \mathcal{E} independently with probability

$$p_{uv} = \Theta\left(\min\left\{\frac{w_u w_v}{n \|x_u - x_v\|^d}, 1\right\}^\alpha\right). \tag{5}$$

We also allow $\alpha = \infty$ and in this case require that

$$p_{uv} = \begin{cases} \Theta(1), & \text{if } \|x_u - x_v\| \leq O\left(\left(\frac{w_u w_v}{n}\right)^{1/d}\right), \\ 0, & \text{if } \|x_u - x_v\| \geq \Omega\left(\left(\frac{w_u w_v}{n}\right)^{1/d}\right), \end{cases} \tag{6}$$

where the constants hidden by O and Ω do not have to match, i.e., there can be an interval $[c_1(\frac{w_u w_v}{n})^{1/d}, c_2(\frac{w_u w_v}{n})^{1/d}]$ for $\|x_u - x_v\|$ where the behaviour of p_{uv} is arbitrary.

We now introduce extensions of the Chung-Lu and the GIRG models that allow to modulate the assortativity of the graph. This is achieved by varying the influence of the smaller weight in the connection probability via a parameter $0 \leq \sigma \leq \tau - 1$. We note here that the choice $\sigma = 1$ recovers the original models. We begin with the extension of the Chung-Lu model, which we call Tunable Chung-Lu graphs.

Definition 2.5 (Tunable Chung-Lu graph). *Let $\tau > 2$, $\sigma \in [0, \tau - 1]$ and let \mathcal{D} be a power-law distribution on $[1, \infty)$ with exponent τ . A Tunable Chung-Lu graph is obtained by the following two-step procedure:*

- (1) Every vertex $v \in \mathcal{V}$ draws iid a weight $w_v \sim \mathcal{D}$.
- (2) For every two distinct vertices $u, v \in \mathcal{V}$, add an edge between u and v in \mathcal{E} independently with probability

$$\Theta\left(\min\left\{\frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n}, 1\right\}\right). \tag{7}$$

Next we define the analogous extension of GIRGs, which we call TGIRGs. This model has been studied in [42,64,71,72] and by Lüchtrath in his PhD Thesis [43], but never in the context of assortativity.

Definition 2.6 (TGIRG). *Let $\tau > 2$, $\alpha > 1$, $\sigma \in [0, \tau - 1]$ and $d \in \mathbb{N}$ and let \mathcal{D} be a power-law distribution on $[1, \infty)$ with exponent τ . A Tunable Geometric Inhomogeneous Random Graph (TGIRG) is obtained by the following three-step procedure:*

- (1) Every vertex $v \in \mathcal{V}$ draws iid a weight $w_v \sim \mathcal{D}$.
- (2) Every vertex $v \in \mathcal{V}$ draws independently and uar a position x_v in the hypercube $[0, 1]^d$.
- (3) For every two distinct vertices $u, v \in \mathcal{V}$, add an edge between u and v in \mathcal{E} independently with probability

$$\Theta\left(\min\left\{\frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n \|x_u - x_v\|^d}, 1\right\}^\alpha\right). \tag{8}$$

Analogously to GIRGs, also here we allow $\alpha = \infty$, requiring in this case that

$$p_{uv} = \begin{cases} \Theta(1), & \text{if } \|x_u - x_v\| \leq O\left(\left(\frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n}\right)^{1/d}\right), \\ 0, & \text{if } \|x_u - x_v\| \geq \Omega\left(\left(\frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n}\right)^{1/d}\right). \end{cases} \tag{9}$$

We remark that for $\sigma = 0$ we recover the (soft) Boolean model [65,73,74], which is basically a Random Geometric Graph where each vertex has an associated random radius drawn from a power-law distribution, while for $\sigma = \tau - 2$ we recover the age-dependent random connection model introduced in [44] as an approximation to the spatial preferential attachment model [75]. We also note that (T)GIRGs sometimes use a slightly different parametrization, replacing the long-range parameter $\alpha \in (1, \infty]$ by its inverse $T := \frac{1}{\alpha} \in [0, 1)$, called the *temperature* of the model.

3 General properties of the graph models

In this section, we describe some general and important properties of Chung-Lu graphs and GIRGs. They hold in fact for a general class of graph models described in [15]. They also generalize to Tunable Chung-Lu graphs and TGIRGs (remember that we get the original Chung-Lu graph model and GIRG model by setting $\sigma = 1$ in the definition of the corresponding tunable model), hence for the sake of conciseness we state each result directly for the more general tunable models.

The first result gives the marginal probability that an edge between two vertices with given weights is present in TGIRGs.

Lemma 3.1. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a TGIRG and $v \in \mathcal{V}$ be a vertex with fixed position $x_v \in [0, 1]^d$. Then all edges uv for $u \neq v$ are independently present with probability

$$\mathbb{P}(uv \in \mathcal{E} \mid x_u, w_u, w_v) = \Theta\left(\min\left\{\frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n}, 1\right\}\right). \tag{10}$$

In particular, we can also remove the conditioning on x_u (by integrating over $[0, 1]^d$) and get

$$\mathbb{P}(uv \in \mathcal{E} \mid w_u, w_v) = \Theta\left(\min\left\{\frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n}, 1\right\}\right). \tag{11}$$

Proof. For classical GIRGs, this corresponds to Lemma 4.2 and Theorem 7.3 in [15]. We give the proof for TGIRGs in the supporting information. □

The next lemma says that the expected degree of a vertex is of the same order as its weight, thus allowing us to treat a given weight sequence $(w_v)_{v \in \mathcal{V}}$ as a sequence of expected degrees.

Lemma 3.2. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a Tunable Chung-Lu graph or a TGIRG and $v \in \mathcal{V}$ be a vertex with fixed weight $w_v \in [1, \infty)$. Then we have $\mathbb{E}[\deg(v) \mid w_v] = \Theta(\min\{w_v, n\})$, or equivalently $\mathbb{P}(uv \in \mathcal{E} \mid w_v) = \Theta(\min\{\frac{w_v}{n}, 1\})$ for a vertex u chosen uniformly at random in \mathcal{V} .

Proof. For classical GIRGs and Chung-Lu graphs, this corresponds to Lemma 4.3 in [15]. The general result was proved using a different parametrization in [43]. For convenience of the reader and in the spirit of self-containment, we provide a proof with our parametrization in the supporting information. □

By observing that $\mathbb{E}[w_v] = \Theta(1)$ for all vertices $v \in \mathcal{V}$, this then naturally yields the connection probability between two random vertices.

Lemma 3.3. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a Tunable Chung-Lu graph or a TGIRG and $u, v \in \mathcal{V}$ be vertices chosen uniformly at random. Then $\mathbb{P}(uv \in \mathcal{E}) = \Theta(1/n)$.

For vertices of sufficiently large weight, their degree is concentrated around its expectation, as stated precisely in the next lemma.

Lemma 3.4. Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a Tunable Chung-Lu graph or a TGIRG. Then the following properties hold with probability $1 - n^{-\omega(1)}$:

- (i) $\deg(v) = O(w_v + \log^2 n)$ for all $v \in \mathcal{V}$.
- (ii) $\deg(v) = (1 + o(1))\mathbb{E}[\deg(v)] = \Theta(w_v)$ for all $v \in \mathcal{V}$ with $w_v = \omega(\log^2 n)$.

Proof. For classical GIRGs and Chung-Lu graphs, this corresponds to Lemma 4.4 in [15]. The same proof can be used for the tunable models. □

The above results imply in particular that Chung-Lu graphs and GIRGs have a power-law degree distribution with exponent τ . Next we will see that, just like classical GIRGs, TGIRGs exhibit clustering. Before stating the result formally, we give the required definition of the (average) clustering coefficient of a graph.

Definition 3.5. In a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ the local clustering coefficient of a vertex $v \in \mathcal{V}$ is defined as

$$cc(v) := cc_{\mathcal{G}}(v) := \begin{cases} \frac{|\{\{u, u'\} \subseteq \Gamma(v), uu' \in \mathcal{E}\}|}{\binom{\deg(v)}{2}} & \text{if } \deg(v) \geq 2, \\ 0 & \text{otherwise,} \end{cases}$$

and the average clustering coefficient, or simply clustering coefficient, of \mathcal{G} is given by

$$cc(\mathcal{G}) := \frac{1}{|\mathcal{V}|} \sum_{v \in \mathcal{V}} cc(v).$$

Proposition 3.6 below then states that the average clustering coefficient of TGIRGs is constant. Note that the statement places no further restrictions on σ . In particular, for $\sigma = 1$, we recover the classical GIRG model for which the analogous result was proven originally in [14]. We remark that in Tunable Chung-Lu graphs the situation is completely different and due to the lack of geometry their clustering coefficient tends to 0 with increasing n .

Proposition 3.6. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a TGIRG. Then whp its clustering coefficient satisfies $cc(\mathcal{G}) = \Theta(1)$.*

4 Conditional distributions of weights

In this section, we give conditional and joint weight distributions for Tunable Chung-Lu graphs and TGIRGs. Note that all results directly apply to classical Chung-Lu graphs and GIRGs (often with a simplified statement) by taking $\sigma = 1$. We begin with the weight distribution of a (vertex) endpoint of an edge drawn uniformly at random. The key ingredient for deriving the results of this section is Bayes' Formula. Full proofs are given in [S1 File](#).

Proposition 4.1. *Let $\tau > 2\sigma \in [0, \tau - 1)$ and let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a Tunable Chung-Lu graph or a TGIRG whose weight distribution \mathcal{D} is a power-law with exponent τ . Denote by $f_{W_v}(w \mid uv \in \mathcal{E})$ the density of the weight of the endpoint of an edge uv chosen uniformly at random in \mathcal{E} . Then*

$$f_{W_v}(w \mid uv \in \mathcal{E}) = \begin{cases} \Theta(w^{1-\tau}) & \text{if } w \leq n, \\ \Theta(w^{-\tau}n) & \text{if } w > n. \end{cases} \tag{12}$$

Note that the second case in [Eq \(12\)](#) basically never occurs, since whp the largest weight is of order $n^{1/(\tau-1)} = o(n)$. Proposition 4.1 therefore tells us that the weight of the endpoint of a random edge follows a power-law with exponent $\tau - 1$ (as opposed to the weight of random vertex, which also follows a power-law, but with exponent τ). The next proposition gives the joint weight density of two endpoints u, v of a randomly chosen edge.

Proposition 4.2. *Let $\tau > 2\sigma \in [0, \tau - 1)$ and let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a Tunable Chung-Lu graph or a TGIRG whose weight distribution \mathcal{D} is a power-law with exponent τ . Denote by $f_{(W_u, W_v)}(w_u, w_v \mid uv \in \mathcal{E})$ the joint density of the weights of the endpoints of an edge uv chosen uniformly at random in \mathcal{E} . Then*

$$f_{(W_u, W_v)}(w_u, w_v \mid uv \in \mathcal{E}) = \begin{cases} (w_u \wedge w_v)^{\sigma-\tau} (w_u \vee w_v)^{1-\tau} & \text{if } (w_u \wedge w_v)^\sigma (w_u \vee w_v) \leq n, \\ n(w_u w_v)^{-\tau} & \text{if } (w_u \wedge w_v)^\sigma (w_u \vee w_v) > n, \end{cases}$$

i.e., $f_{(W_u, W_v)}(w_u, w_v \mid uv \in \mathcal{E}) = \Theta(((w_u \wedge w_v)^\sigma (w_u \vee w_v) \wedge n) \cdot (w_u w_v)^{-\tau})$.

We remark here that the case distinction stems from the cutoff at 1 in the connection probability. There is a multigraph version of the model where we do not cap the expression at 1 but simply require that the expected number of edges connecting u and v is $\frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n}$. For the vast majority of the vertices, this expectation will naturally be at most 1, and for the remaining vertex pairs, one can for example randomly assign $\lfloor \frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n} \rfloor$ or $\lceil \frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n} \rceil$ edges, with the probabilities chosen to match the desired expectation. In this multigraph version, the first line in the display equation in Proposition 4.2 holds for all weights, not just for weights with $(w_u \wedge w_v)^\sigma (w_u \vee w_v) \leq n$. In the simple graph version that is the focus of this paper, the first line of the display still applies for a large majority of vertex pairs. Note that for the classical Chung-Lu and GIRG models (obtained by choosing $\sigma = 1$), this line simplifies to $f_{W_u, W_v}(w_u, w_v \mid uv \in \mathcal{E}) = w_u^{1-\tau} w_v^{1-\tau}$, and by Proposition 4.1 this is the same as $f_{W_u}(w_u \mid uv \in \mathcal{E}) \cdot f_{W_v}(w_v \mid uv \in \mathcal{E})$. In other words, (under the cutoff) the two endpoint weights are independent, which already is a strong indication of the neutral assortativity of these classical models. The next proposition gives the weight density of a vertex v as a function of the weight w_u of its neighbor u , which gives more insight in the assortative behavior of Tunable Chung-Lu graphs and TGIRGs in general.

Proposition 4.3. Let $\tau \in (2, 3)$, $\sigma \in [0, \tau - 1)$ and let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a Tunable Chung-Lu graph or a TGIRG whose weight distribution \mathcal{D} is a power-law with exponent τ . Consider an edge $uv \in \mathcal{E}$ connecting a vertex u with weight w_u to some vertex v . Then the conditional distribution of the weight W_v of v satisfies

$$f_{W_v}(w \mid w_u, uv \in \mathcal{E}) = \begin{cases} \Theta(w^{1-\tau}(w \wedge w_u)^{\sigma-1}) & \text{if } (w \wedge w_u)^\sigma (w \vee w_u) \leq n, \\ \Theta\left(\frac{w^{-\tau}n}{w_u}\right) & \text{if } w_u \leq n \leq (w \wedge w_u)^\sigma (w \vee w_u), \\ \Theta(w^{-\tau}) & \text{if } w_u > n. \end{cases}$$

Again, the last case of the display equation in Proposition 4.3 is basically void since whp all weight are much smaller than n . Moreover, the second case only applies to exceptionally large weights. For the classical choice $\sigma = 1$, the first case gives $f_{W_v}(w \mid w_u, uv \in \mathcal{E}) = \Theta(w^{1-\tau})$ and in particular this conditional distribution is actually independent of w_u , which shows the neutral assortativity of Chung-Lu graphs and GIRGs. Compared to this baseline, we see that by choosing $\sigma > 1$ the distribution gets a boost of order $(w \wedge w_u)^{\sigma-1}$. In particular, this boosting factor is increased for larger w and w_u , which translates to positive assortativity. On the other hand, choosing $\sigma < 1$ makes the exponent $\sigma - 1$ negative, and thus yields an opposite effect, which shows the disassortative behavior of the corresponding graph models.

5 Degree assortativity in RGGs, Chung-Lu graphs and GIRGs

So far, we have analyzed assortativity with respect to the vertex weights of our random graph models. However, our ulterior motive is to understand assortativity with respect to vertex *degrees*. In the graph models of interest here, the weight of a vertex corresponds to its expected degree, and hence Propositions 4.1-4.3 are also informative about the degree assortativity in typical instances of Chung-Lu graphs and GIRGs. Moreover, the degrees of vertices of sufficiently large (polylogarithmic) weights are whp of the same order as their weight (see, e.g., Lemma 3.4), making these vertices completely assortativity-neutral in terms of degree as well.

In GIRGs however, there is an additional effect caused by random vertex-density fluctuations in certain regions, i.e., by chance, some regions in the ground space of the GIRG may contain more nodes. These effects are noticeable for low-degree nodes. In order to explain this effect, we consider the simpler model of Random Geometric Graphs. We note here that the same analysis applies to GIRGs but comes at the price of much higher technicality. We omit the details but note qualitatively that the larger the weights of the involved nodes, the smaller the effect.

5.1 Degree assortativity in random geometric graphs

In the following, we derive the conditional degree distribution for vertices given the degree of one of their neighbors in Random Geometric Graphs, demonstrating that degrees of neighboring vertices are positively correlated. In this section, we will denote by $B_r(x)$ the ball of radius r centered at $x \in [0, 1]^d$. By $\text{Vol}(A)$ we denote the Euclidean volume of a set $A \subseteq [0, 1]^d$ and by $Po(\mu)$, following the usual convention, a Poisson random variable with mean μ . Recall from our definition of RGGs that the latent space we are using is equipped with the torus topology and distances are measured using the max-norm $\|\cdot\|_\infty$. **Proposition 5.1.** Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a Random Geometric Graph on n vertices in the d -dimensional hypercube, let $uv \in \mathcal{E}$ be an edge selected uniformly at random from \mathcal{E} , and let the degree of u be $\text{deg}(u) = k$. Then $\text{deg}(v)$ is distributed as

$$\text{deg}(v) \sim 1 + Po\left((k-1) \frac{\text{Vol}(B_r(x_u) \cap B_r(x_v))}{\text{Vol}(B_r(x_u))}\right) + Po\left((n-1-k) \frac{\text{Vol}(B_r(x_u) \setminus B_r(x_v))}{\text{Vol}([0, 1]^d)}\right),$$

where

$$\mathbb{E}\left[\frac{\text{Vol}(B_r(x_u) \cap B_r(x_v))}{\text{Vol}(B_r(x_u))}\right] = \left(\frac{3}{4}\right)^d.$$

In particular, $\text{deg}(v)$ can be stochastically lower-bounded by a random variable X distributed as

$$X \sim 1 + \text{Po}(\Theta(k)).$$

Proposition 5.1 tells us that RGGs have positive degree assortativity in the following sense. The degree of a random neighbor v of some vertex u can be decomposed into two independent Poisson random variables. The expectation of the first one does not change with $\text{deg}(u)$, while the expectation of the second one scales linearly with $\text{deg}(u)$ (Fig 4).

5.2 Degree assortativity in GIRGs

We are now in a position to summarize our findings. We can distinguish the following cases. Pick a vertex u of weight w_u and one of its neighbors v of weight w_v .

- Whenever $w_u \cdot w_v \leq n$, the conditional weight distribution of w_v is independent of w_u (Proposition 4.3). Hence the only assortative effect on the degree of v comes from the geometric (positive) assortativity: The degree of a vertex conditional on the degree of its neighbor depends positively on the intersection of their balls of influence - whereby here the ball of influence is the geometric region (which is a function of w_v) in which v connects to every vertex with at least constant probability. (Note that whenever $w_v \gg w_u$, then the ball of influence of u is typically contained in the ball of influence of v .) For constant-weight vertices, this effect of the geometry indeed plays a prominent role. This influence decays as w_v grows. Since whp no weight is of order $\Omega(n)$, this scenario occurs in particular whenever $w_v = O(1)$.
- If $w_u \cdot w_v > n$, the conditional density resp. tail of the distribution of w_v decays (proportionally) as w_u increases. Since, as mentioned, degrees are concentrated for large weights, our theoretical results suggest that the negatively assortative effect of the weights exceed the positive (but vanishing) assortative effect of the geometry.

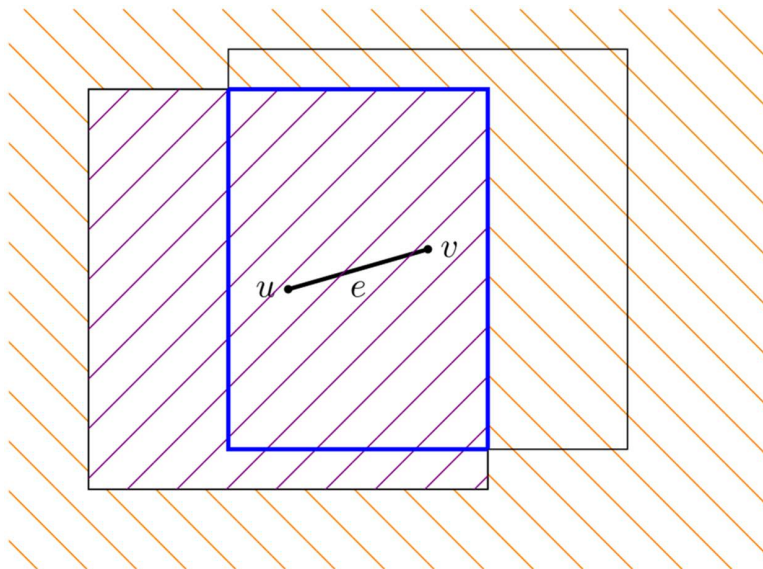


Fig 4. Illustration of the intersection of the boxes of influence $B_r(x_u) \cap B_r(x_v)$ of two vertices u, v . If we know that $\text{deg}(u) = k$, then the density of vertices inside $B_r(x_u)$ (area shaded in purple) is proportional to k , while the density of vertices outside $B_r(x_u)$ (shaded in orange) decreases (slightly) with k . Note that the intersection $B_r(x_u) \cap B_r(x_v)$ (delimited in blue) makes up a constant fraction of the area of $B_r(x_v)$, which leads to positive assortativity in RGGs.

<https://doi.org/10.1371/journal.pcsy.0000097.g004>

- If either w_v or w_u is larger than n , the weights have no assortative effect anymore (cf. Proposition 4.3). However, whp this case does not occur across the whole graph.

6 Experiments

Here we complement our theoretical considerations with an empirical evaluation. Besides other interesting observations, our experiments in particular support our claims that (i) our tunable models indeed control the assortativity in the desired way and that (ii) using just a single number to describe how vertices of different degrees tend to connect does not do the rich underlying structure justice.

We start in Sect 6.1 by studying assortativity coefficients based on the correlation measures between the two endpoints of an edge. As correlation measures, we consider Pearson’s, Spearman’s, and variants of Kendall’s coefficients. Our analysis confirms previous findings [33] that Spearman’s rank correlation coefficient (and similarly Kendall’s) provides a more appropriate measure of assortativity than Pearson’s correlation coefficient, although we will see in Sect 6.2 that reducing the measure assortativity to a single number nonetheless has serious limitations. Additionally, we show how the σ parameter in our tunable models can be leveraged to regulate these assortativity coefficients.

In Sect 6.2, we take a more fine-grained look at the joint distribution of vertex degrees. This provides deeper insights into how vertices of different degrees tend to connect that go beyond what can be expressed with just a single number. **Overview of real-world networks.** Before presenting the experimental results, we provide brief background information on the real-world network dataset examined here; also see Table 1. More detailed descriptions of the individual networks can be found in S1 File. In order to allow for meaningful results, we restrict our analysis to networks which are verifiably power-law. The rigorous detection of power laws in empirical degree distributions in real-world networks is a challenging task – hence we rely on the evaluations of [40], who analyzed 35 undirected real-world networks without multi-edges collected from the KONECT database [34]. Two of the networks, the Catster/Dogster network and the Route views contain

Table 1. The real-world networks we use in our experiments. We report the number of vertices, the number of edges, the average degree, the Spearman assortativity coefficient and the power-law exponent τ estimated by the Hill estimator [76] as reported in [40].

graph	vertices	edges	avg degree	τ	assortativity
LiveJournal	5 204 176	49 174 464	18.90	3.86	0.4
Orkut	3 072 441	117 184 899	76.28	3.58	0.36
Gowalla	196 591	950 327	9.67	2.80	0.25
Brightkite	58 228	214 078	7.35	3.51	0.22
Bible names	1 773	9 131	10.30	3.09	0.08
WordNet	146 005	656 999	9.00	2.86	0.03
Human PPI	3 133	6 726	4.29	3.04	-0.03
Youtube	1 134 890	2 987 624	5.27	2.48	-0.08
Skitter	1 696 415	11 095 298	13.08	2.38	-0.14
Proteins	1 870	2 277	2.44	3.09	-0.16
Amazon	334 863	925 872	5.53	3.99	-0.19
Catster/Dogster	623 766	15 699 276	50.34	2.10	-0.23
Dogster	426 820	8 546 581	40.05	2.15	-0.3
Route views	6 474	13 895	4.29	2.13	-0.5
CAIDA	26 475	53 381	4.03	2.10	-0.53
Hyves	1 402 673	2 777 419	3.96	2.98	-0.55
Catster	149 700	5 449 275	72.80	2.09	-0.56

<https://doi.org/10.1371/journal.pcsy.0000097.t001>

self-loops. The methodology of [40] ensures in particular statistically consistent and robust estimations of the tail exponent τ of the power law from the measured degree sequence. We further required that the networks were non-bipartite, and had estimated power-law parameters within or close to the range $2 < \tau < 3$, for all three employed estimators: Hill, Moments and Kernel estimator. For the experiments comparing assortativity coefficient values, we additionally include networks which are of strong power-law type but whose exponent τ is larger than 3.

Code and data. The code used for these experiments is freely available on GitHub (<https://github.com/thobl/assortativity>) and Zenodo (<https://zenodo.org/records/16746826>). All used and produced data is available on Zenodo (<https://zenodo.org/records/16745980>).

6.1 Assortativity coefficients

An assortativity coefficient measures the correlation between the degrees of the vertices of an edge. To make this more precise, consider an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. For each edge $\{u, v\} \in \mathcal{E}$ we consider the pairs $(\text{deg}(u), \text{deg}(v))$ and $(\text{deg}(v), \text{deg}(u))$. For this set of pairs, we consider three types of correlation measures: Pearson's coefficient, Spearman's coefficient, and Kendall's coefficient. All coefficients take values in $[-1, 1]$, with -1 denoting maximally negative assortativity and 1 denoting maximally positive assortativity. We defer their precise definitions to [S1 File](#).

6.1.1 Assortativity coefficients of real-world networks. [Fig 5](#) shows and compares the different assortativity coefficients for the surveyed networks. Consistent with our theoretical result (Theorem 1.1), the sign of the Pearson correlation coefficient is always negative for networks with $\tau < \frac{7}{3}$ and essentially non-positive for all surveyed networks including those with $\tau > 3$. For Spearman and Kendall, we obtain a much wider range of different assortativity values. For all considered networks, the sign of the assortativity is the same for Spearman and Kendall – and their magnitudes track each other closely. This strongly supports the previous suggestion of Litvak and van der Hofstad [33] to use Spearman and not Pearson to measure assortativity. In the following, unless explicitly mentioned otherwise, the term assortativity coefficient always refers to the Spearman correlation.

Overall we can observe that most networks have negative assortativity. In particular for networks with power-law exponent $\tau < 3$, the three networks with largest assortativity coefficients are Gowalla (0.25), WordNet (0.03) and Youtube (-0.08).

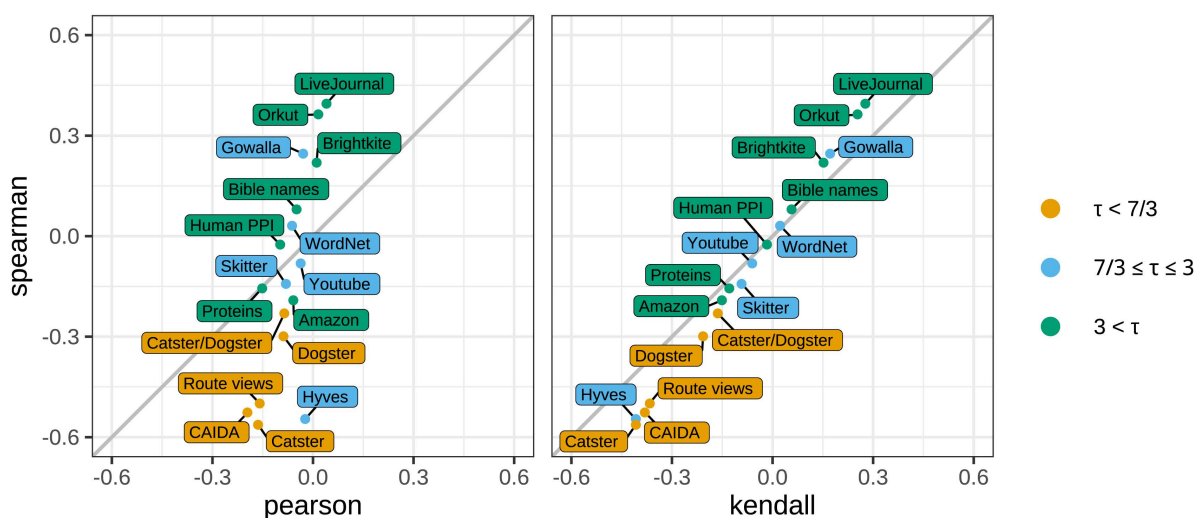


Fig 5. Comparison of assortativity coefficients (Pearson, Kendall, Spearman) for real networks.

<https://doi.org/10.1371/journal.pcsy.0000097.g005>

6.1.2 Our tunable models. The goal with our tunable models was to have a parameter σ that controls the assortativity coefficient. Before analyzing the plots in detail, let us briefly recall the role of σ in the connection probability, which we restate here for the case of Tunable Chung-Lu graphs – the role of σ in the connection probability of TGIRGs is analogous:

$$\Theta\left(\min\left\{\frac{(w_u \wedge w_v)^\sigma (w_u \vee w_v)}{n}, 1\right\}\right). \tag{13}$$

Fig 6 shows the assortativity depending on σ . Focusing on the Spearman coefficient (bottom row) for now, one can clearly see that changing σ has the desired effect that the assortativity changes monotonically depending on σ . Note that from our theoretical considerations, we know that we only get a power-law degree distribution for $\tau \geq \sigma + 1$, which corresponds to the change of behavior with sudden upticks for $\sigma \in \{1.2, 1.4, 1.6\}$ for the curves representing $\tau \in \{2.2, 2.4, 2.6\}$, respectively. We note that for $\tau = 2.2$, in the range $\sigma \leq \tau - 1$, also the value of Spearman, despite increasing as σ increases, never becomes positive. Our theoretical results (Proposition 4.2), which hold for all $2 < \tau < 3$ indicate that this is another artifact of the use of coefficients and not a sign that there is no assortative wiring behavior occurring in the graph.

Comparing Chung-Lu graphs with GIRGs, we can see that the geometry facilitates a larger assortativity coefficient. While the effect is not very strong for most parameter combinations, it is particularly pronounced for small values of σ and large values of τ . The latter is predicted by our theoretical results: we know that the positive effect on assortativity from the latent geometry is restricted to low-degree vertices (Sect 5), on those are more dominant for larger τ .

We note that the assortativity coefficient is rather stable with respect to scaling the graph size, see Fig 7. We also refer to Sect 6.2 for degree distributions showing that we indeed still get a power-law distribution if $\tau \geq \sigma + 1$.

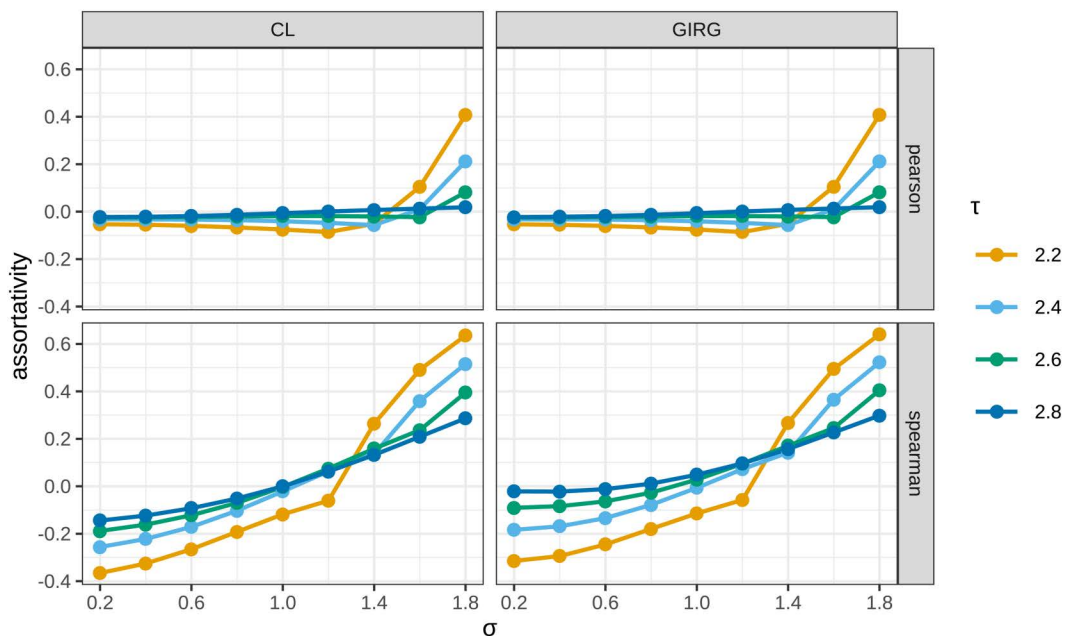


Fig 6. Assortativity coefficients of Tunable Chung-Lu graphs and TGIRGs as a function of σ for different power-law exponents τ . Each point is the mean of five generated networks. Each generated network has 200 000 vertices and expected average degree 15. For the TGIRGs we use dimension 2 and temperature 0.

<https://doi.org/10.1371/journal.pcsy.0000097.g006>

The top row of [Fig 6](#) is only there to again show the shortcomings of Pearson correlation in the context of assortativity. The coefficient only slightly changes and actually decreases for increased σ , except for the regimes where σ is too large to actually yield a power-law distribution.

6.2 Joint distribution

Here we take a more detailed look at the joint distribution of edge degrees, i.e., when drawing a random edge $\{u, v\}$, $X = \text{deg}(u)$ and $Y = \text{deg}(v)$ are random variables and we are interested in their joint distribution. We propose three different ways of looking at this joint distribution, resulting in three different types of plots, which we introduce in the following, using a Chung-Lu graph with different values of σ as an example; see [Fig 8A–8C](#).

Joint distribution heatmap. [Fig 8A](#) directly shows the joint distribution of X and Y , i.e., the color at coordinate (x, y) indicates the probability that $X=x$ and $Y=y$. Note that the plot uses logarithmic axes and that the degrees are grouped into buckets. To be more precise, we use 21 buckets and for $i \in [0, 20]$, the i th bucket B_i is the range $B_i = [b^i, b^{i+1})$. The base b is chosen such that the maximum degree d_{max} is in the last bucket, i.e., $b = \sqrt[21]{d_{max} + 1}$.

Comparing the three plots in [Fig 8A](#), we can clearly see the effect of σ on the joint distribution, with $\sigma < 1$ (disassortative) facilitating edges between vertices of different degrees while $\sigma > 1$ leads to more edges with endpoints of similar degree. However, we note that this trend becomes only apparent due to the comparison of the plots for different values of σ . When just looking at the plot for $\sigma = 1$, it is not obvious that the corresponding graph has neutral assortativity. Indeed, the slightly brighter L-shape might be perceived as negative assortativity, which would be a misinterpretation; also see [Fig 6](#). Thus, although this is certainly the simplest representation of the joint distribution, we recommend to use one of the other representations introduced below.

Complementary cumulative distributions. [Fig 8B](#) shows the complementary cumulative distribution of the random variable $X = \text{deg}(u)$ for two different random trials. The line labeled node shows $\mathbb{P}[X \geq x]$ if u is a random node. The line labeled edge shows $\mathbb{P}[X \geq x]$ if u is the random endpoint of a random edge $\{u, v\}$. For the latter, we are also interested in conditioning on the degree $Y = \text{deg}(v)$ of the other endpoint v , i.e., $\mathbb{P}[X \geq x | Y]$. For this, we consider five different values of Y . More precisely, we condition on Y being in one of the five evenly distributed buckets $B_0, B_5, B_{10}, B_{15},$ or B_{20} , as defined above. In [Fig 8B](#), this is labeled as edge c where $c \in \{0, 0.25, 0.5, 0.75, 1\}$ refers to the bucket B_{20c} . Also note the colored arrows on the left side of the plots in [Fig 8A](#). They indicate on which value of Y the corresponding line in [Fig 8B](#) conditions.

Before we discuss the line plots, we briefly comment on the scaling of the plots and the interpretation of the curves. As is common for networks whose degree sequence tail follows a power law, we use a logarithmic scale on both axes. In this scaling, conforming to a power law then means that the curve is a straight line and its gradient corresponds to $-\tau + 1$, where τ is the exponent of the power law.

Considering the node line, one can see that changing σ only slightly changes the curve, i.e., we get the desired power-law distribution consistent with our theoretical results that the (expected) degree of a node is of the same order as its weight and hence does not depend on σ (Lemma 3.2). Concerning assortativity considerations, note that complete independence of the two endpoint-degrees of an edge would mean that conditioning on the degree Y should not change the distribution of X . In terms of the curves, this would mean that all curves edge c should coincide with edge. Note that in the central plot of [Fig 8B](#) with $\sigma = 1$ (neutral assortativity), this is indeed the case for $c \in \{0.25, 0.5, 0.75\}$. Moreover, for $c \in \{0, 1\}$, it is also the case for small values of x . For larger values of x , however, conditioning on Y being very large ($c = 1$) makes it less likely that $X \geq x$. This makes sense as in a Chung-Lu graph, there are not enough high-degree vertices such that many high-degree vertices can connect to other high-degree vertices, see Proposition 4.3 and also the discussion after the re-statement of Theorem 1.1 in Sect 7.1. Conversely, the probability $X \geq x$ is slightly increased for large x when conditioning on Y being very small ($c = 0$).

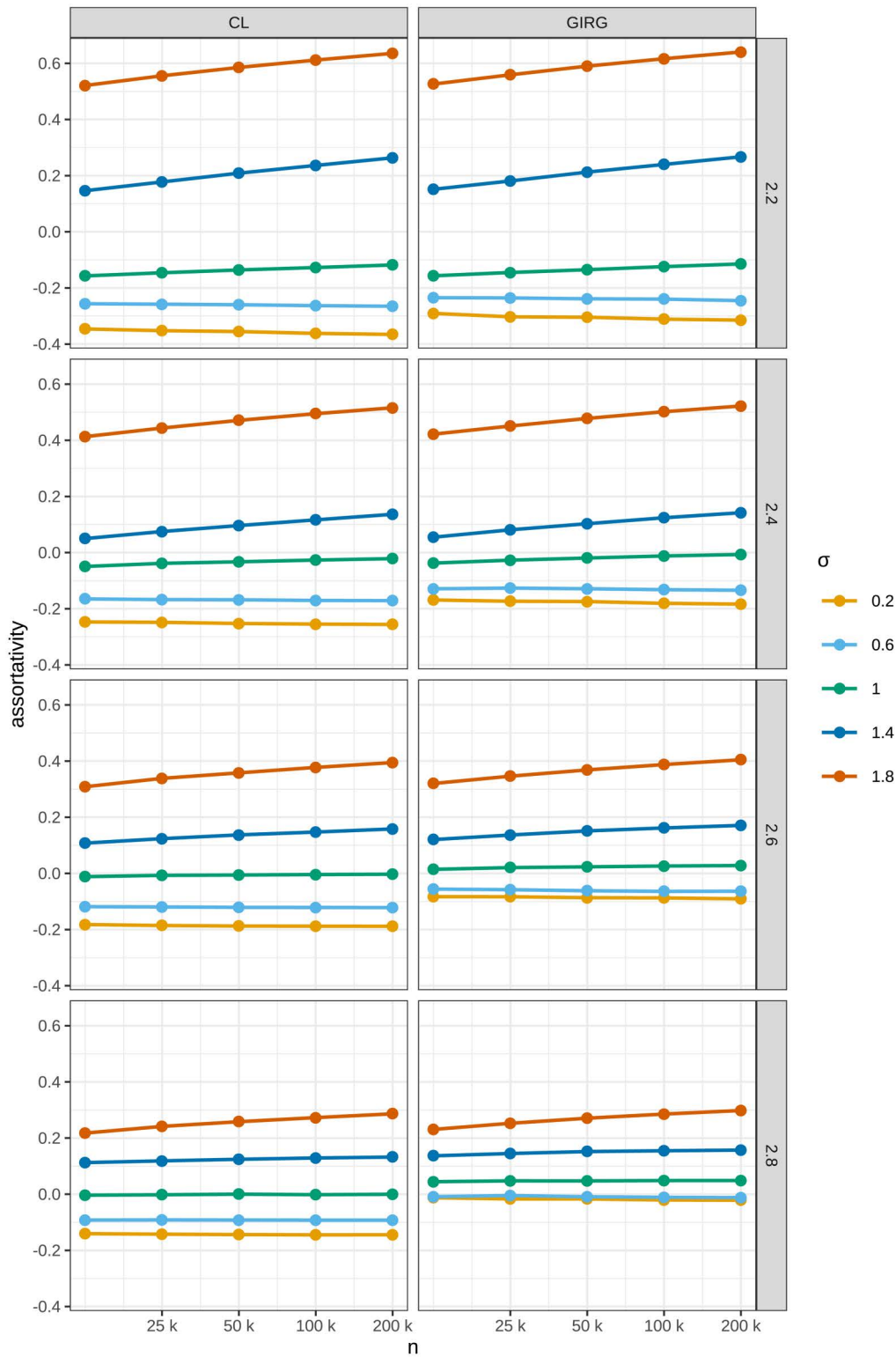
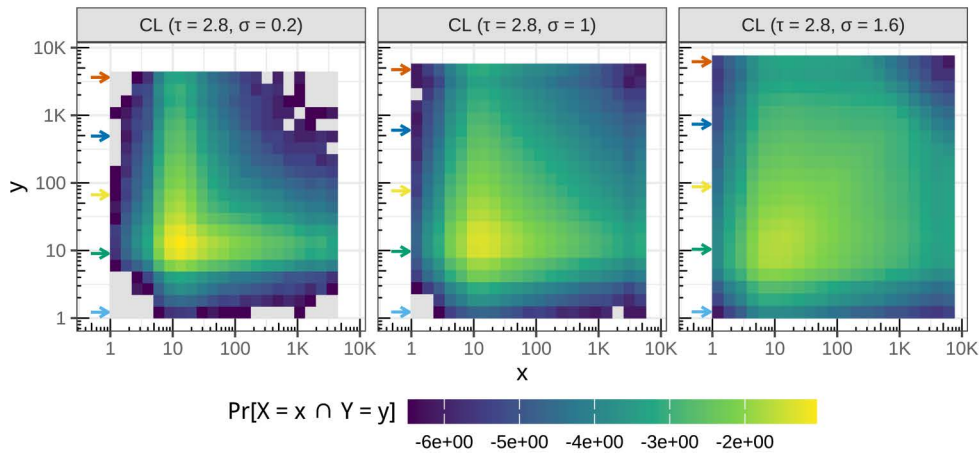


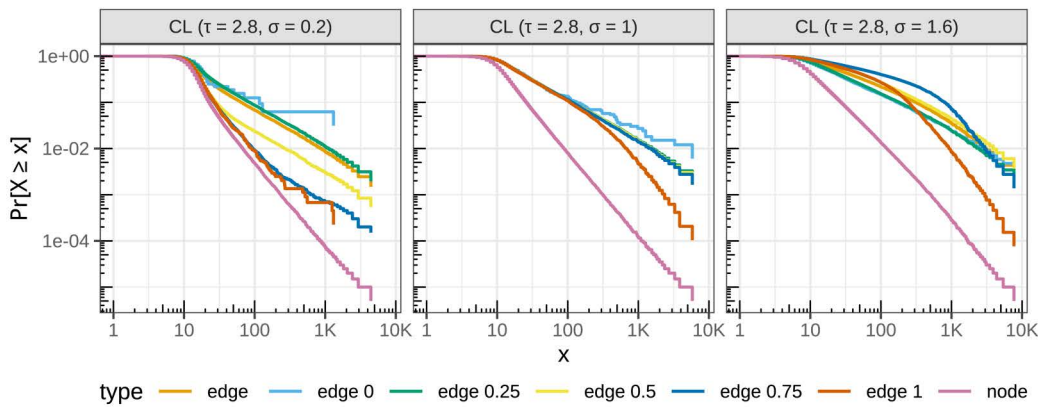
Fig 7. Spearman Assortativity coefficient values in Tunable Chung-Lu Graphs and TGIRGs for varying network sizes and levels of σ .

<https://doi.org/10.1371/journal.pcsy.0000097.g007>

(A)



(B)



(C)

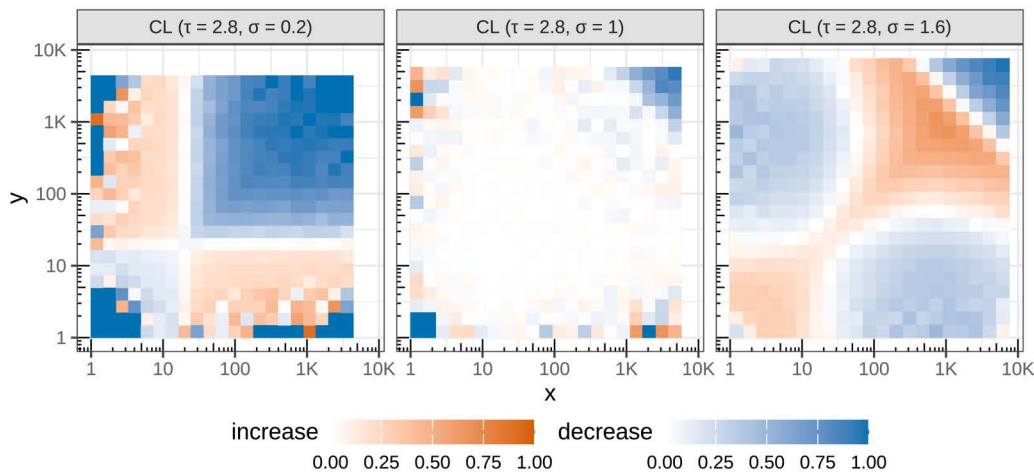


Fig 8. Three different types of plots illustrating the joint distribution. Each column shows a Chung-Lu graph with 200 000 vertices, average degree 15, power-law exponent $\tau = 2.8$, and varying $\sigma \in \{0.2, 1, 1.6\}$, controlling the assortativity. **(A)** Joint distribution of vertex degrees of a random edge. The colors use a logarithmic scale. **(B)** Complementary cumulative distribution of the degree of a random node (node) and of a random endpoint of a random

edge (edge). The numbers indicate conditioning on the degree of the other endpoint. **(C)** For a random edge with degrees (X, Y) , the color indicates how $\mathbb{P}[X = x]$ changes when conditioning on $Y = y$.

<https://doi.org/10.1371/journal.pcsy.0000097.g008>

Varying σ clearly changes the picture. For $\sigma = 0.2$ (negative assortativity), one can see in the left plot of Fig 8B that the lines corresponding to a large value of Y ($c \in \{0.5, 0.75, 1\}$) lie significantly below the edge line, i.e., conditioning on Y being large decreases the probability that $X \geq x$. This matches the intuition of what negative assortativity is supposed to mean. For $\sigma = 1.8$ (positive assortativity) in the right plot, one can see that for small values of x , the same lines ($c \in \{0.5, 0.75, 1\}$) lie above the edge line, which again makes sense for positive assortativity. For larger values of x , however, the lines for $c=1$ and later also for $c=0.75$ fall below the edge line, which again comes from the fact that there are just not enough high-degree vertices to support enough edges between vertices of very high degree. This agrees with the derivation in Proposition 4.3, where we prove that the tail of the curves is composed of two different linear pieces of different slopes. As predicted in Proposition 4.3, the second slopes starts very late for small values of c and is thus not visible, but it becomes observable for $c \in \{0.75, 1\}$.

While these plots in Fig 8B are somewhat difficult to read, we would argue that they are more informative than the heatmaps shown in Fig 8A.

Conditional heatmaps. Finally, the plots in Fig 8C show how $\mathbb{P}[X = x]$ changes when conditioning on $Y = y$ (again using the same buckets as before). We normalize this change as follows. If $\mathbb{P}[X = x | Y = y] \geq \mathbb{P}[X = x]$, we say that we have an *increase* of $1 - \frac{\mathbb{P}[X=x]}{\mathbb{P}[X=x|Y=y]}$. Otherwise, we have a *decrease* of $1 - \frac{\mathbb{P}[X=x|Y=y]}{\mathbb{P}[X=x]}$. Note that this normalizes the values to lie between 0 and 1 with values close to 1 indicating a large change due to conditioning on $Y = y$. Moreover, 0 increase is equivalent to 0 decrease, indicating independence. We note that the plots are symmetric as $\frac{\mathbb{P}[X=x]}{\mathbb{P}[X=x|Y=y]} = \frac{\mathbb{P}[Y=y]}{\mathbb{P}[Y=y|X=x]}$. For brevity, we call these heatmaps *conditional heatmaps*.

We believe these conditional heatmaps to be easy to read and insightful. In Fig 8C, one can nicely see that for neutral assortativity ($\sigma = 1$), the plot is mostly white. Moreover, the effect of varying σ is very apparent. We discuss these effects in more detail in the following section.

6.2.1 Discussion – Generated networks. The three rows of Fig 9 show the conditional heatmaps for tunable Chung-Lu as well as TGIRGs with temperature $T=0$ and $T=0.7$, corresponding to $\alpha = \infty$ and $\alpha \approx 1.4$ respectively. The columns show different values from $\sigma \in \{0.2, 1, 1.6\}$. Thus, considering the different columns aids our understanding of σ 's influence on the assortativity of the underlying graphs. Moreover, the different rows let us study the impact of the geometry.

For the central column ($\sigma = 1$), we can see that the conditional heatmap is mostly white, indicating neutral assortativity. Notice that for Chung-Lu (first row, central column) there is some variation for the low-degree vertices (in particular for degrees 1 and 2). This is most likely just noise: As there are in general only very few vertices of such low degree, a slight absolute change results in a big relative change. The completely blue square in the bottom-left corner for example tells us that there are no edges where both endpoints have degree 1 or 2, which is not surprising if there are in general few vertices of such low degree. For GIRGs the picture is similar, except that we see an increase in edges where both endpoints have low degree (red bottom-left corner). This can be explained by random fluctuation in the density of nodes in the underlying geometry, as derived in Sect 5. One can also nicely see how TGIRGs with higher temperature ($T=0.7$) lie between Chung-Lu and TGIRG with temperature zero.

For varying σ , we observe two extremal patterns. As one extreme, we see an L-shape, with a vertical red stripe on the left hand of the heatmap and horizontal red stripe on the bottom of the heatmap ($\sigma = 0.2$). This reflects a strongly disassortative wiring pattern where high-degree vertices prefer to connect to low-degree vertices and vice versa. It indicates that low values of σ successfully produce disassortative networks (while preserving key properties of classical Chung-Lu graphs and GIRGs respectively). Again, for TGIRGs, the geometry makes connections between low-degree vertices more assortative, resulting in the red lower left corner of the conditional heatmap.

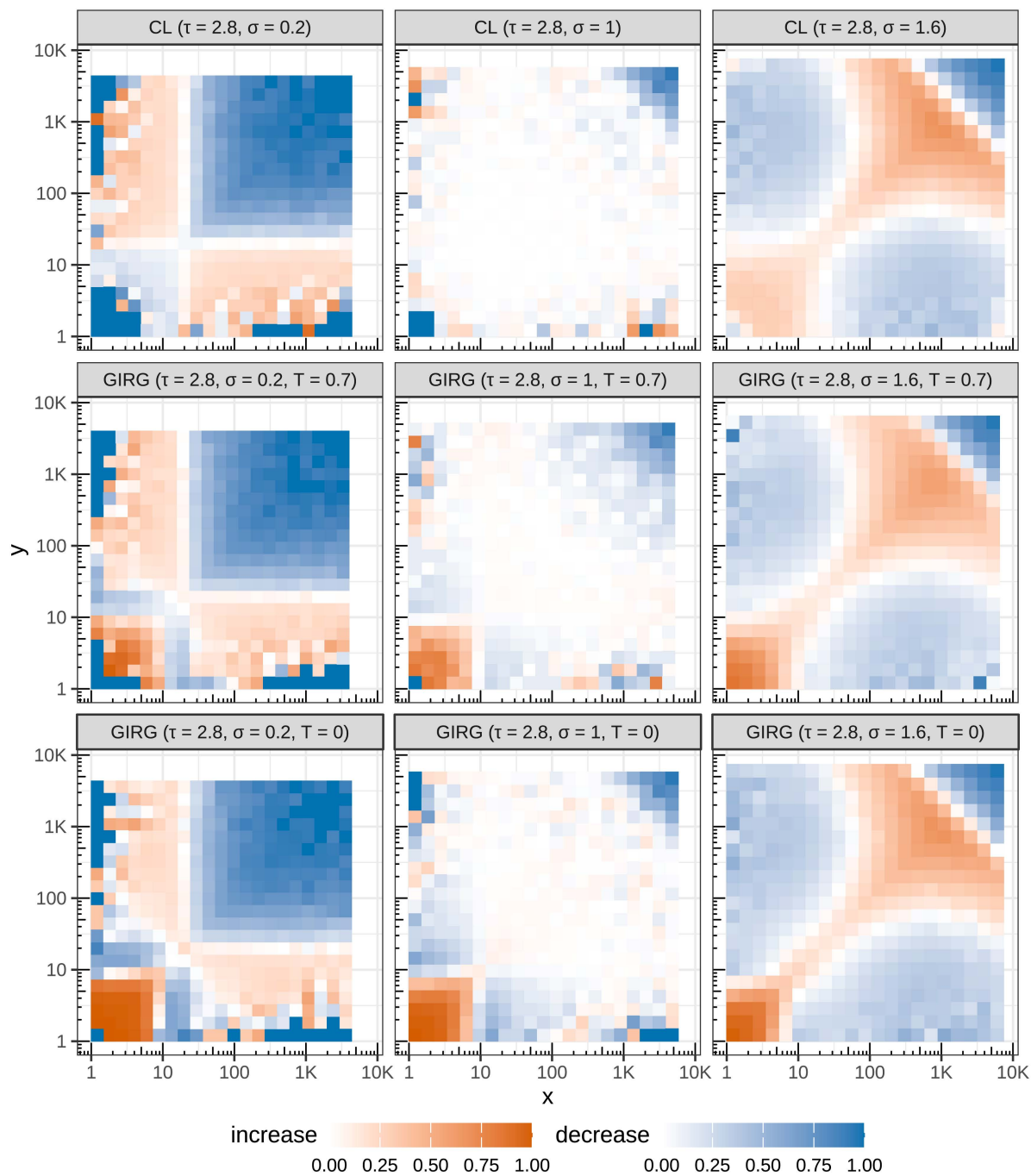


Fig 9. Conditional heatmaps for Tunable Chung-Lu (top), TGIRGs with temperature $T=0.7$ (middle row), and $T=0$ (bottom) for $\sigma = 0.2$ (left), $\sigma = 1$ (middle column), and $\sigma = 1.6$ (right). Each graph has 200 000 vertices, average degree 15, and power-law exponent $\tau = 2.8$. The dimension for the TGIRGs is 2.

<https://doi.org/10.1371/journal.pcsy.0000097.g009>

On the other extreme, in highly assortative networks one would expect heatmaps to display a distinguished red diagonal (connecting the lower left corner of the heatmap to its upper right corner). Indeed, this would indicate that vertices in the corresponding network of degree, say k , preferably connect to vertices of degree k , whereas connecting to vertices of any other degree is less likely. And indeed, such assortative behavior can be observed qualitatively for a large range of degrees for both Tunable Chung-Lu graphs and TGIRGs when $\sigma = 1.6$; see the right column in [Fig 9](#). Notably, this effect now also reaches low-degree nodes in Tunable Chung-Lu graphs. Note that for very high-degree nodes, this idealized assortative wiring indicated by the red diagonal can not be maintained. The blue triangle, which is always present in power-law networks with exponent $\tau > 2$ now leads to a lateral spreading of the probability mass which would be on the diagonal to the nodes with highest degrees that are available. We remark again that the Pearson correlation coefficient puts special emphasis on these very-high-degree vertices which are forced to connect to lower-degree vertices and thus contribute negatively to the overall picture, which is highly problematic for scale-free networks in general and renders it useless for small values of τ . This is made rigorous by our results in [Sect 7](#).

6.2.2 Discussion – Real-world networks. Changing focus now to the real-world networks, the wiring patterns are much more nuanced. Nonetheless, some large-scale phenomena can be made out in the conditional heatmaps shown in [Fig 10](#). Similar to the generated graphs, all plots show a blue triangle in the top-right corner. Due to normalization, this also means that the remaining regions in the same rows and columns appear more red. To compensate for this effect we need to compare with an assortative-neutral baseline, for which we can use Chung-Lu networks with matching power-law exponent τ . [Fig 11](#) shows such Chung-Lu networks for various values of τ . The size of the blue region is larger for smaller τ . For large τ the resulting effect is negligible since then the blue region is too small to have much impact, but it is notable for small τ .

Social networks. Considering the social networks in the first two rows of [Fig 10](#), we want to point out that their Spearman assortativity coefficients cover a large range. In fact, from all networks in our data set with power-law exponent $\tau \in (2, 3]$, the `Catster` and `Gowalla` networks have the smallest (-0.56) and largest (0.25) assortativity coefficients, respectively; see [Fig 5](#) and [Table 1](#). This contradicts the common belief that social networks tend to have positive assortativity [1]. Nonetheless, the conditional heatmaps in [Fig 10](#) reveal a pattern that is common among the social networks in our data set: Low degree vertices in the range [1,10] preferably connect to vertices of similar degree and to vertices with much higher degree but not so much to vertices in the intermediate range. Note that the colors especially in the three petster networks (`Catster`, `Dogster`, `Catster/Dogster`) and in the `Gowalla` network are substantially more pronounced than in the baseline in [Fig 11](#). Thus the pattern of preferences is not just a normalization artifact.

This pattern can at least partially explain the wide range in assortativity values. Connections between pairs of low-degree vertices contribute to positive assortativity while edges with one low-degree and one high-degree endpoint yield negative assortativity. These are conflicting effects and depending on which is stronger, one might obtain positive, negative, or neutral assortativity.

Beyond these similarities between the networks in the wiring patterns of low-degree vertices, there are also some differences, in particular for the medium degree vertices. In the `Catster` network, we can observe that vertices in the range [10, 100] preferably connect to vertices of much higher degree. In contrast to that, in the `Gowalla` network, we also get an increased number of connections between medium-degree vertices, resulting in a slightly red diagonal. We note that this fits to the fact that `Catster` has negative and `Gowalla` has positive assortativity. The `Youtube` network, shows a similar behavior as discussed before, with low-degree vertices (range [1,7]) connecting to other low-degree vertices and to high-degree vertices (degree > 400). In this network, vertices of intermediate degree in the range [20, 400] predominantly connect to vertices in the same range. However, we do not see a red diagonal as there is a gap where vertices in the range [7,20] are slightly less likely to connect with vertices in the same range. We note that, while the assortativity coefficient of the `Youtube` network is close to 0, we observe a highly interesting wiring pattern that is far from being neutral, i.e., there is a dependence between the vertex degrees of an edge that it is too intricate to be well described by a single number.

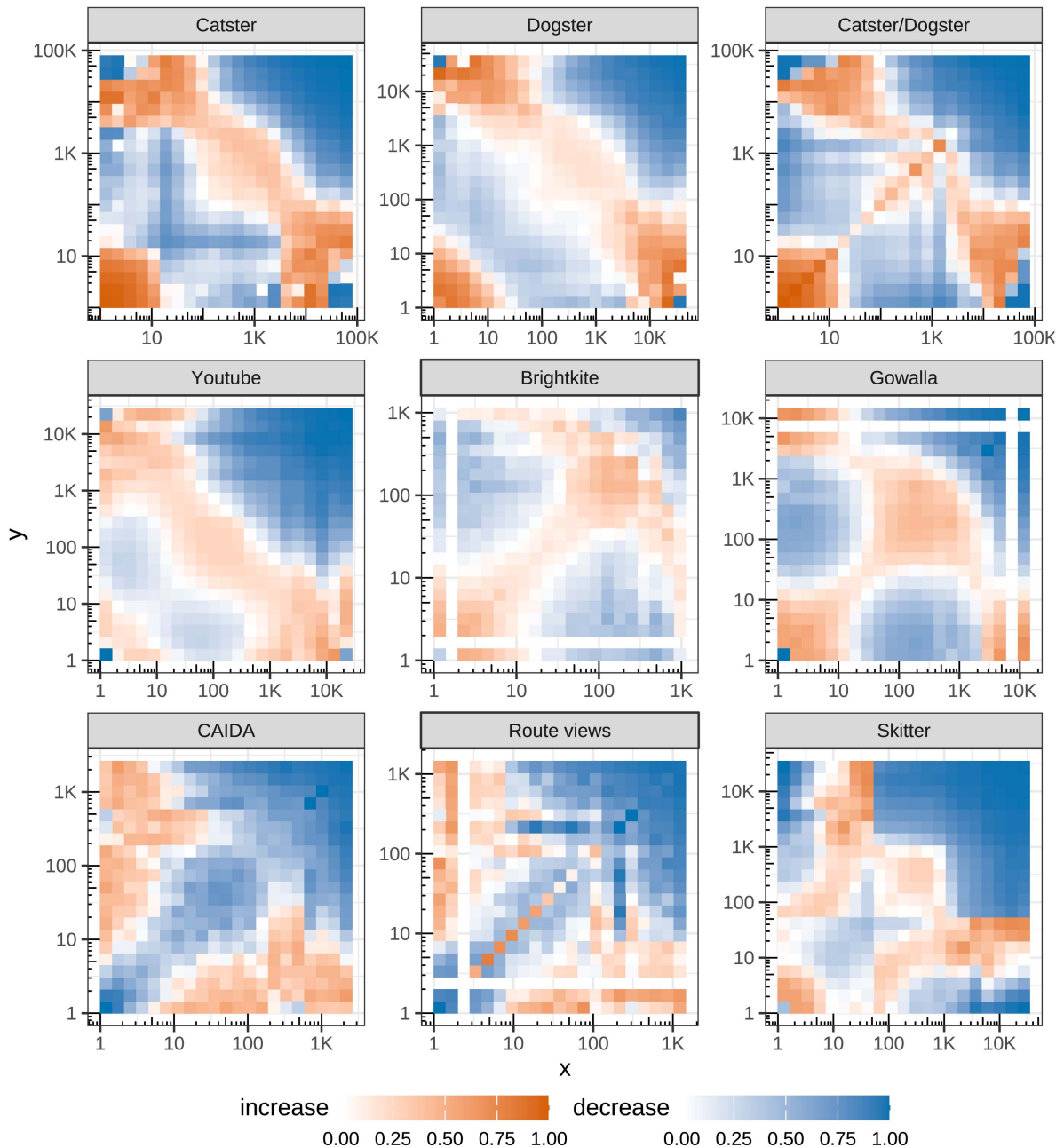


Fig 10. Conditional heatmaps of nine real-world networks. The first six are social networks, the bottom three are technical networks representing connections between autonomous systems. The networks are ordered by their Spearman assortativity coefficient: -0.56 (Catster), -0.30 (Dogster), -0.23 (Catster/Dogster), -0.08 (Youtube), 0.22 (Brightkite), 0.25 (Gowalla) for the social networks and -0.53 (CAIDA), -0.50 (Route views), -0.14 (Skitter) for the autonomous systems.

<https://doi.org/10.1371/journal.pcsy.0000097.g010>

Autonomous systems. The three networks in the bottom row of Fig 10 show computer networks where each vertex represents an autonomous systems and edges indicate direct connections. Interestingly, we see a very different wiring pattern compared to the social networks. In the first two networks, CAIDA and Route views, we get a very similar picture.

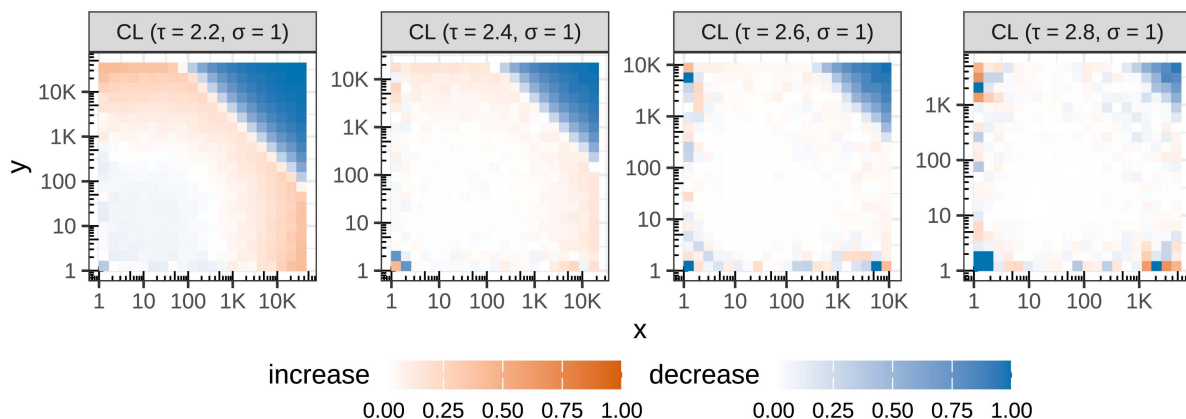


Fig 11. Conditional heatmaps for two Chung-Lu networks with different power law exponents τ . We use those as assortative-neutral baselines.

<https://doi.org/10.1371/journal.pcsy.0000097.g011>

The main difference is the sharp increase on the diagonal for `Route` views, which is an artifact from self-loops in the network data. Moreover, `Route` views is more noisy than `CAIDA`, which is likely due to the fact that the network is substantially smaller (6.5k and 26k vertices, respectively). We thus focus on `CAIDA` in the following, although similar observations are true for `Route` views.

The overall pattern for `CAIDA` is that the diagonal is strikingly sparse and generally most connections are between vertices of low degree (<10) and vertices of higher degree (>10). Examining the wiring preferences more in detail, the only exception to this pattern are vertices with degree in the range [200, 600] that also have an increased connection probability to vertices in the range [10,30], i.e., slightly above 10. This overall negative assortativity of the `CAIDA` network, which is also reflected by the Spearman assortativity coefficient of -0.53 , is consistent with the general belief that technological networks tend to be disassortative.

For the `Skitter` network, the picture is more nuanced. There is an increased number of edges where both endpoints have a similar degree for the ranges [1,7] and [100, 600]. Beyond this assortative wiring, we mostly get disassortative patterns: Vertices with degree in the range [1,7] have an increased number of connections to vertices in the range [50, 200]. Vertices of slightly higher degree in range [7,50] tend to connect to vertices of even higher degree (> 100). Although less prominent than for `CAIDA`, this overall yields mostly disassortative wiring, which is also reflected in the assortativity coefficient of -0.14 .

We believe that the differences between `CAIDA` and `Skitter` are quite interesting and we can offer some potential explanations (which should be viewed as educated guesses rather than definite truth). These networks of autonomous systems form strong hierarchical structures with connections between different levels of the hierarchy. Moreover, vertices higher up in the hierarchy tend to be central hubs with high degree while vertices lower down in the hierarchy have fewer connections. It thus makes sense that we get disassortative patterns if most connections are between different levels of the hierarchy. The core difference between `CAIDA` and `Skitter` is their size, with 26k vertices for `CAIDA` and 1.7M vertices for `Skitter`. In a smaller network like `CAIDA`, it is conceivable to have a flat hierarchy with vertices of very low degree directly connecting to vertices of very high degree. However, this likely becomes infeasible in the larger `Skitter` network. This is consistent with the rather extreme power-law exponents of $\tau = 2.1$ for `CAIDA` and $\tau = 2.13$ for `Route` views, while `Skitter` has $\tau = 2.38$. With this, the patterns we observe for the `Skitter` network between vertices of different degrees (top-left and bottom-right region of the plot) could come from a hierarchy with multiple levels, where connections are predominantly between adjacent levels containing vertices with different (but not extremely different) degrees. Moreover, the two red regions on the diagonal could come from connections on

the same level of the hierarchy or from adjacent levels of the hierarchy containing vertices of similar degree (like, e.g., in a regular tree).

Although *Skitter* is less disassortative than the other two, we would say that our results are in agreement with the general belief that technological networks tend to be disassortative. However, the detailed picture is again more nuanced than can easily be described with a single number.

Comparison with the models. We have already seen in the previous paragraphs that real-world networks in Fig 10 can exhibit a variety of assortativity patterns, and that it makes sense to distinguish between social networks (first two rows in Fig 10) and the technical networks stemming from autonomous systems (last row in Fig 10).

For social networks, we universally observe that low-degree vertices show an increased preference to connect to other low-degree vertices, which contributes positively to assortativity. On the other hand, assortativity between other pairs of vertices may vary widely. In Sect 5.2, we showed for our generative models that the effect of the latent geometric space on assortativity is to increase assortativity between pairs of vertices of low degree, while not affecting assortativity between other pairs of vertices. The fact that all studied social networks show such an increased assortativity between low-degree vertices suggests that the changes induced by a latent space are compatible with the observed assortativity in social networks.

In general the real-world networks show a more complex structure than the network models. When we compare individual social networks with generated networks of the same parameters, the fit is not tight. Fig 12 shows a comparison for

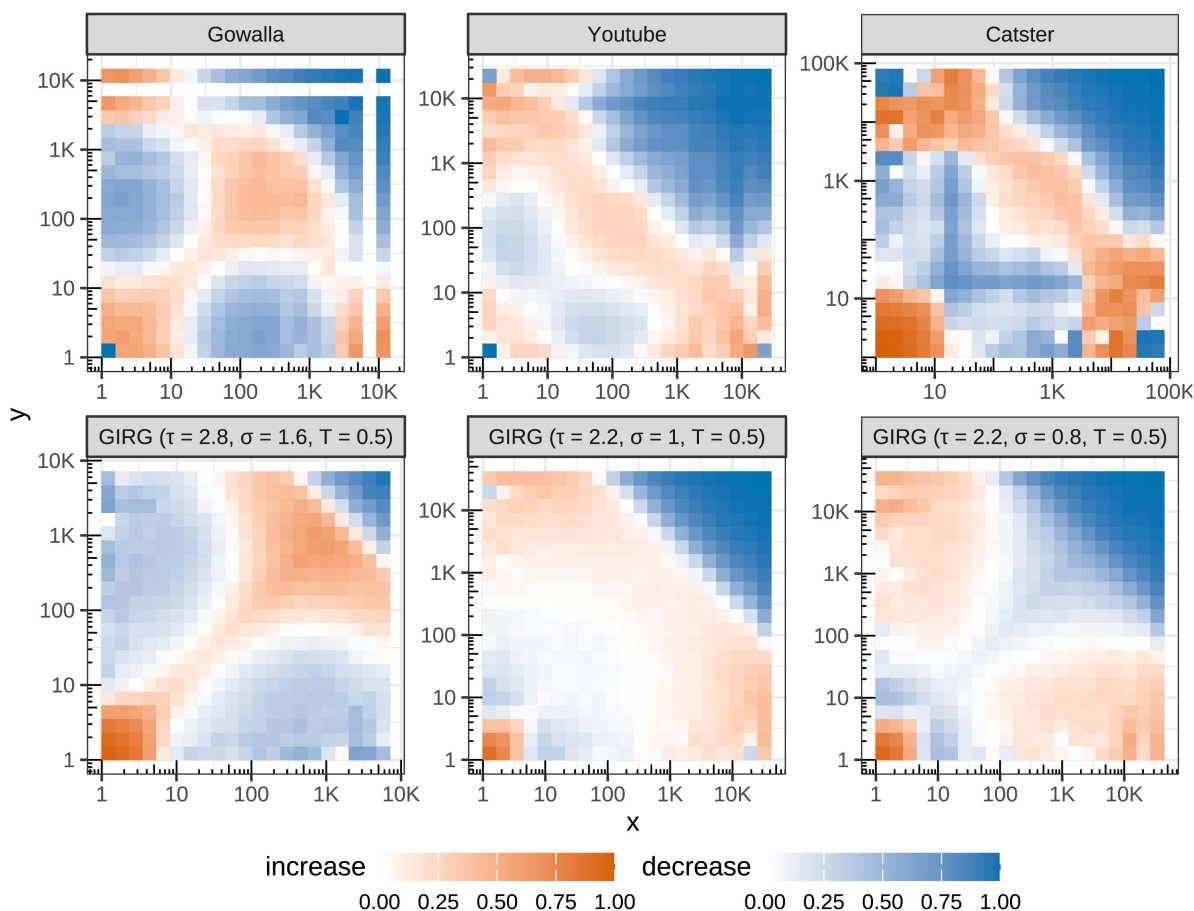


Fig 12. Direct comparison of social networks with generated networks of corresponding parameters.

<https://doi.org/10.1371/journal.pcsy.0000097.g012>

several social networks. The *Gowalla* network shows a red λ -shape in its conditional heatmap in Fig 10 which has some resemblance to the λ -shape of the TGIRG with the same power-law exponent $\tau = 2.8$ and with $\sigma = 1.6$ (left bottom panel). However, apart from the more nuanced structure in the real network we want to highlight one important difference: for the generated network the blue triangle is substantially smaller and the upper diagonal of the red λ -shape does not extend to the corners, while it does so for the *Gowalla* network. This means that vertices of low degree (which are most vertices) have an increased probability to form an edge with a vertex of very high degree in the *Gowalla* network, but not in the corresponding TGIRG. This effect might partially stem from an imperfect power-law in the *Gowalla* network, see the corresponding figure in S1 File. But even for generated models with a smaller value of τ , the probability is not as much increased as in the *Gowalla* network, cf. Fig 11. This indicates that the wiring pattern of vertices of very high degree is not adequately captured by the models. There is only a small number of such vertices, but due to their high degrees they contribute a substantial fraction of all edges.

The situation for the *Youtube* network is similar to *Gowalla*. Fitting the parameters gives a loose fit (middle row, bottom panel), but it does not capture the finer structures and is not a good quantitative fit in all regions. Note that in the plot we chose $\tau = 2.2$ because that gives a better fit better the size of the blue triangle in the upper right, but it does not correspond to the exponent $\tau = 2.48$ of the *Youtube* network.

Among social networks, we get the worst fit for the petster networks. Since those have negative assortativity coefficient, this suggest using a tuning parameter $\sigma < 1$, but this does not result in a similar conditional heatmap, see the right column in Fig 12. Arguably, the gestalt shape for neutral $\sigma = 1$ (Fig 12, middle row bottom panel) or even $\sigma = 1.2$ (Fig 13 left) fits better, but even then the quantitative match is not good. In particular, it underestimates the number of connections between vertices of low and of very high degree. This suggests that the assortative structure of the petster networks is richer and can not be captured by the uniform application of a single tuning parameter as in the TGIRG model.

For the autonomous systems networks (bottom row in Fig 10), we get somewhat similar wiring patterns when using $\sigma = 0.8$ (Fig 13 middle and right). Due to the fact that an underlying geometry increases the probability for edges between low-degree vertices, the Tunable Chung-Lu seems to be a better fit for CAIDA and Route views, while TGIRG seems to be a better fit for Skitter. A notable difference to the artificial networks is that *Skitter* has an increased probability for edges where both vertices are in the range [100,600], which is not present in the models.

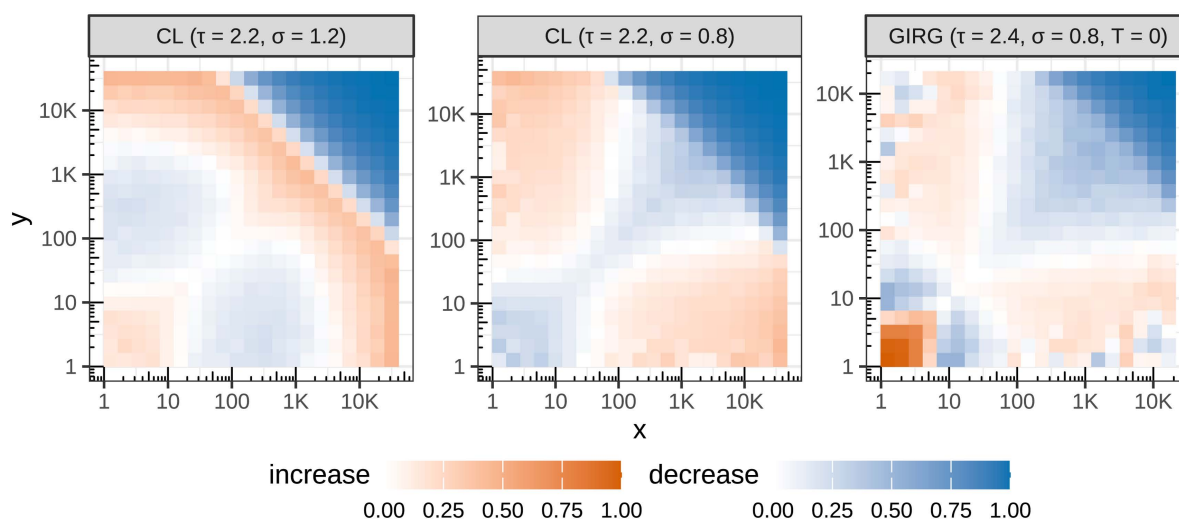


Fig 13. Conditional heatmaps for three models that, to some degree, show similar patterns as observed in the real-world networks.

<https://doi.org/10.1371/journal.pcsy.0000097.g013>

7 Scale-free networks and assortativity coefficients

This section is dedicated to the analysis of scale-free networks using assortativity coefficients. In the first subsection, we show that for a substantial range of τ , all networks whose degree distribution follow a power-law with parameter τ have negative Pearson correlation coefficient. It has been argued previously that the Pearson coefficient is ill-suited to model degree-degree correlations in scale-free networks. Here, we provide a much stronger result which supports this claim. Intuitively, our main theorem of this section means that the Pearson correlation coefficient can never give meaningful results in any network with with power-law degree distribution and exponent close to 2, independent of any homophily or heterophily in the network. In the second subsection, we provide a discussion of experimental results as well as alternative coefficients - and why the analysis of conditional and joint distributions is better suited to understand assortativity than coefficients.

7.1 Severe shortcoming of Pearson correlation coefficient for the analysis of scale-free networks

We next provide the formal definition of the Pearson correlation coefficient, as given by Newman in [1].

Consider a network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and define a pair (X_0, X_1) of random variables as follows. We sample an edge $e = uv \in \mathcal{E}$ uniformly at random, and then set $(X_0, X_1) = (\deg(u) - 1, \deg(v) - 1)$ with probability $\frac{1}{2}$ and $(X_0, X_1) = (\deg(v) - 1, \deg(u) - 1)$ with probability $\frac{1}{2}$. These $\deg(\cdot) - 1$ terms correspond to the *remaining degree* of the vertex, i.e., the number of additional edges that are incident to the vertex, discounting the edge e that led there in the first place. Note that X_0 and X_1 are identically distributed but not necessarily independent. One way to quantify the dependence between the endpoint degrees of a random edge is to compute the Pearson correlation coefficient of X_0 and X_1 . The *Pearson assortativity coefficient* of the graph \mathcal{G} is defined as such:

$$r(\mathcal{G}) := \frac{\text{Cov}(X_0, X_1)}{\sqrt{\text{Var}(X_0)}\sqrt{\text{Var}(X_1)}} = \frac{\mathbb{E}[X_0X_1] - \mathbb{E}[X_0]\mathbb{E}[X_1]}{\text{Var}(X_0)}, \quad (14)$$

where the equality above holds because X_0 and X_1 have the same distribution. Note in particular that the denominator is always positive by definition. For convenience, we restate the main theorem of this section.

Intuitively, the negativity of $r(\mathcal{G})$ comes from the contribution of high-degree vertices. More precisely, because of the power-law degree distribution, any vertex of degree $\omega(n^{1/\tau})$ has to connect to vertices of smaller degree, because there are $\omega(n^{1/\tau})$ vertices of degree $\omega(n^{1/\tau})$. This effect is even stronger for vertices of maximum degree $\Theta(n^{1/(\tau-1)})$, because almost all of their neighbors have much smaller degree. If the tail of the degree distribution is heavy enough (which translates to the $\tau < 7/3$ condition), the contribution of these vertices dominates the numerator in (14).

8 Materials and methods

In this brief section, for the convenience of the reader, we give references to the relevant passages of the paper that provide sufficient detail to reproduce our findings.

Networks and network models. The formal definitions of the graph models we use can be found in Sect 2, with crucial model properties stated in Sect 3. We describe the sampling process for TGIRGs in S1 File. An overview of the dataset of real networks used in our experiments is given at the beginning of Sect 6, with detailed descriptions of the real networks given in S1 File.

Assortativity coefficients. An overview of the assortativity coefficients we use is given in Sect 6.1, with the definition of Pearson's coefficient given in Sect 7.1 and additional details given in S1 File.

Experimental setup. The setup for our experiments is described in Sect 6, with details regarding the heatmaps given in Sect 6.2. The code used for our experiments is freely available on GitHub (<https://github.com/thobl/assortativity>) and

Zenodo (<https://zenodo.org/records/16746826>). All used and produced data is available on Zenodo (<https://zenodo.org/records/16745980>).

Supporting information

S1 File. Supporting information for assortativity in geometric and scale-free networks. Contains the proofs omitted from Sects [3](#), [4](#), [5](#) and [7](#), the formal definitions of the correlation measures used in our analysis, additional information on the real networks in our dataset, additional figures and guidelines for generating TGIRGs.
(PDF)

Author contributions

Conceptualization: Johannes Lengler.

Data curation: Thomas Bläsius.

Formal analysis: Marc Kaufmann, Ulysse Schaller, Thomas Bläsius, Johannes Lengler.

Funding acquisition: Thomas Bläsius, Johannes Lengler.

Investigation: Marc Kaufmann, Ulysse Schaller, Thomas Bläsius, Johannes Lengler.

Methodology: Marc Kaufmann, Ulysse Schaller, Thomas Bläsius, Johannes Lengler.

Project administration: Marc Kaufmann, Ulysse Schaller.

Software: Thomas Bläsius.

Supervision: Thomas Bläsius, Johannes Lengler.

Validation: Thomas Bläsius.

Visualization: Marc Kaufmann, Ulysse Schaller, Thomas Bläsius, Johannes Lengler.

Writing – original draft: Marc Kaufmann, Ulysse Schaller.

Writing – review & editing: Marc Kaufmann, Thomas Bläsius, Johannes Lengler.

References

1. Newman MEJ. Assortative mixing in networks. *Phys Rev Lett.* 2002;89(20):208701. <https://doi.org/10.1103/PhysRevLett.89.208701> PMID: [12443515](#)
2. Newman MEJ. Mixing patterns in networks. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2003;67(2 Pt 2):026126. <https://doi.org/10.1103/PhysRevE.67.026126> PMID: [12636767](#)
3. D'Agostino G, Scala A, Zlatić V, Caldarelli G. Robustness and assortativity for diffusion-like processes in scale-free networks. *EPL.* 2012;97(6):68006. <https://doi.org/10.1209/0295-5075/97/68006>
4. Zhou D, Stanley HE, D'Agostino G, Scala A. Assortativity decreases the robustness of interdependent networks. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2012;86(6 Pt 2):066103. <https://doi.org/10.1103/PhysRevE.86.066103> PMID: [23368000](#)
5. Duh M, Gosak M, Slavinec M, Perc M. Assortativity provides a narrow margin for enhanced cooperation on multilayer networks. *New J Phys.* 2019;21(12):123016. <https://doi.org/10.1088/1367-2630/ab5cb2>
6. Bialonski S, Lehnertz K. Assortative mixing in functional brain networks during epileptic seizures. *Chaos.* 2013;23(3):033139. <https://doi.org/10.1063/1.4821915> PMID: [24089975](#)
7. Newman MEJ. The structure and function of complex networks. *SIAM Review.* 2003;45(2):167–256. <https://doi.org/10.1137/S003614450342480>
8. Chang H, Su B-B, Zhou Y-P, He D-R. Assortativity and act degree distribution of some collaboration networks. *Physica A: Statistical Mechanics and its Applications.* 2007;383(2):687–702. <https://doi.org/10.1016/j.physa.2007.04.045>
9. Barabasi A, Albert R. Emergence of scaling in random networks. *Science.* 1999;286(5439):509–12. <https://doi.org/10.1126/science.286.5439.509> PMID: [10521342](#)
10. Granovetter MS. The strength of weak ties. *American Journal of Sociology.* 1973;78(6):1360–80. <https://doi.org/10.1086/225469>

11. Masuda N, Miwa H, Konno N. Geographical threshold graphs with small-world and scale-free properties. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2005;71(3 Pt 2A):036108. <https://doi.org/10.1103/PhysRevE.71.036108> PMID: [15903494](https://pubmed.ncbi.nlm.nih.gov/15903494/)
12. Krioukov D, Papadopoulos F, Kitsak M, Vahdat A, Boguñá M. Hyperbolic geometry of complex networks. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2010;82(3 Pt 2):036106. <https://doi.org/10.1103/PhysRevE.82.036106> PMID: [21230138](https://pubmed.ncbi.nlm.nih.gov/21230138/)
13. Deijfen M, van der Hofstad R, Hooghiemstra G. Scale-free percolation. *Ann Inst H Poincaré Probab Statist.* 2013;49(3). <https://doi.org/10.1214/12-aihp480>
14. Bringmann K, Keusch R, Lengler J. Geometric inhomogeneous random graphs. *Theoretical Computer Science.* 2019;760:35–54. <https://doi.org/10.1016/j.tcs.2018.08.014>
15. Bringmann K, Keusch R, Lengler J. Average distance in a general class of scale-free networks. *Adv Appl Probab.* 2024;57(2):371–406. <https://doi.org/10.1017/apr.2024.43>
16. Lengler J, Todorovic L. Existence of small separators depends on geometry for geometric inhomogeneous random graphs. 2017. <https://arxiv.org/abs/1711.03814>
17. Watts DJ, Strogatz SH. Collective dynamics of “small-world” networks. *Nature.* 1998;393(6684):440–2. <https://doi.org/10.1038/30918> PMID: [9623998](https://pubmed.ncbi.nlm.nih.gov/9623998/)
18. Dayan B, Kaufmann M, Schaller U. Expressivity of geometric inhomogeneous random graphs—metric and non-metric. *Springer proceedings in complexity.* Springer Nature Switzerland. 2024. p. 85–100. https://doi.org/10.1007/978-3-031-57515-0_7
19. Kaufmann M, Ravi RR, Schaller U. Sublinear cuts are the exception in BDF-GIRGs. 2024. <https://arxiv.org/abs/2405.19369>
20. Bläsius T, Fischbeck P. On the external validity of average-case analyses of graph algorithms. *ACM Trans Algorithms.* 2024;20(1):1–42. <https://doi.org/10.1145/3633778>
21. Cerf S, Dayan B, De Ambroggio U, Kaufmann M, Lengler J, Schaller U. Balanced bidirectional breadth-first search on scale-free networks. 2024. <https://arxiv.org/abs/241022186>
22. Boguñá M, Papadopoulos F, Krioukov D. Sustaining the Internet with hyperbolic mapping. *Nat Commun.* 2010;1:62. <https://doi.org/10.1038/ncomms1063> PMID: [20842196](https://pubmed.ncbi.nlm.nih.gov/20842196/)
23. Bringmann K, Keusch R, Lengler J, Maus Y, Molla AR. Greedy routing and the algorithmic small-world phenomenon. In: *Proceedings of the ACM Symposium on Principles of Distributed Computing*, 2017. p. 371–80. <https://doi.org/10.1145/3087801.3087829>
24. Bläsius T, Friedrich T, Katzmann M, Krohmer A. *Algorithm Engineering and Experiments (ALENEX)*, 2018. p. 199–208.
25. Milgram S. The small world problem. *Psychology Today.* 1967;1(1):60–7.
26. Komjáthy J, Lapinskas J, Lengler J, Schaller U. Polynomial growth in degree-dependent first passage percolation on spatial random graphs. *Electron J Probab.* 2024;29(none). <https://doi.org/10.1214/24-ejp1216>
27. Komjáthy J, Lapinskas J, Lengler J, Schaller U. Four universal growth regimes in degree-dependent first passage percolation on spatial random graphs I. 2023. <https://arxiv.org/abs/230911840>
28. Jorritsma J, Hulshof T, Komjáthy J. Not all interventions are equal for the height of the second peak. *Chaos Solitons Fractals.* 2020;139:109965. <https://doi.org/10.1016/j.chaos.2020.109965> PMID: [32863609](https://pubmed.ncbi.nlm.nih.gov/32863609/)
29. Koch C, Lengler J. Bootstrap percolation on geometric inhomogeneous random graphs. *Internet Mathematics.* 2021;18995.
30. Friedrich T, Göbel A, Katzmann M, Schiller L. Real-world networks are low-dimensional: theoretical and practical assessment. In: *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, 2024. p. 2036–44. <https://doi.org/10.24963/ijcai.2024/225>
31. Schmeltzer C, Kihara A, Sokolov I, Rüdiger S. Degree correlations optimize neuronal network sensitivity to sub-threshold stimuli. *PLoS ONE.* 2015. <https://doi.org/10.1371/journal.pone.0121794>
32. Vázquez A, Weigt M. Computational complexity arising from degree correlations in networks. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2003;67(2 Pt 2):027101. <https://doi.org/10.1103/PhysRevE.67.027101> PMID: [12636856](https://pubmed.ncbi.nlm.nih.gov/12636856/)
33. Litvak N, van der Hofstad R. Uncovering disassortativity in large scale-free networks. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2013;87(2):022801. <https://doi.org/10.1103/PhysRevE.87.022801> PMID: [23496562](https://pubmed.ncbi.nlm.nih.gov/23496562/)
34. Kunegis J. *Proceedings of the 22nd International Conference on World Wide Web*, 2013. p. 1343–50. <https://doi.org/10.1145/2487788.2488173>
35. Cinelli M, Peel L, Iovanella A, Delvenne J-C. Network constraints on the mixing patterns of binary node metadata. *Phys Rev E.* 2020;102(6–1):062310. <https://doi.org/10.1103/PhysRevE.102.062310> PMID: [33466011](https://pubmed.ncbi.nlm.nih.gov/33466011/)
36. Karimi F, Oliveira M. On the inadequacy of nominal assortativity for assessing homophily in networks. *Sci Rep.* 2023;13(1):21053. <https://doi.org/10.1038/s41598-023-48113-5> PMID: [38030623](https://pubmed.ncbi.nlm.nih.gov/38030623/)
37. Piraveenan M, Prokopenko M, Zomaya AY. Local assortativeness in scale-free networks. *Europhys Lett.* 2010;89(4):49901. <https://doi.org/10.1209/0295-5075/89/49901>
38. Noldus R, Van Mieghem P. Assortativity in complex networks. *jcomplexnetw.* 2015;3(4):507–42. <https://doi.org/10.1093/comnet/cnv005>
39. Chung F, Lu L. The average distances in random graphs with given expected degrees. *Proc Natl Acad Sci U S A.* 2002;99(25):15879–82. <https://doi.org/10.1073/pnas.252631999> PMID: [12466502](https://pubmed.ncbi.nlm.nih.gov/12466502/)
40. Voitalov I, van der Hoorn P, van der Hofstad R, Krioukov D. Scale-free networks well done. *Phys Rev Research.* 2019;1(3). <https://doi.org/10.1103/physrevresearch.1.033034>

41. van der Hofstad R, Litvak N. Degree-degree dependencies in random graphs with heavy-tailed degrees. *Internet Mathematics*. 2014;10(3–4):287–334. <https://doi.org/10.1080/15427951.2013.850455>
42. Jorritsma J, Komjáthy J, Mitsche D. Cluster-size decay in supercritical kernel-based spatial random graphs. 2023. <https://arxiv.org/abs/230300724>
43. Lühtrath L. Percolation in weight-dependent random connection models. University of Cologne; 2022.
44. Gracar P, Heydenreich M, Mönch C, Mörters P. Recurrence versus transience for weight-dependent random connection models. *Electron J Probab*. 2022;27(none). <https://doi.org/10.1214/22-ejp748>
45. Newman ME. The structure of scientific collaboration networks. *Proc Natl Acad Sci U S A*. 2001;98(2):404–9. <https://doi.org/10.1073/pnas.98.2.404> PMID: [11149952](https://pubmed.ncbi.nlm.nih.gov/11149952/)
46. Newman MEJ, Girvan M. Mixing patterns and community structure in networks. *Statistical mechanics of complex networks*. Berlin, Heidelberg: Springer; 2003. p. 66–87.
47. Lusseau D, Newman MEJ. Identifying the role that animals play in their social networks. *Proc Biol Sci*. 2004;271 Suppl 6(Suppl 6):S477–81. <https://doi.org/10.1098/rsbl.2004.0225> PMID: [15801609](https://pubmed.ncbi.nlm.nih.gov/15801609/)
48. Newman MEJ, Park J. Why social networks are different from other types of networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2003;68(3 Pt 2):036122. <https://doi.org/10.1103/PhysRevE.68.036122> PMID: [14524847](https://pubmed.ncbi.nlm.nih.gov/14524847/)
49. Boguñá M, Pastor-Satorras R. Class of correlated random networks with hidden variables. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2003;68(3 Pt 2):036112. <https://doi.org/10.1103/PhysRevE.68.036112> PMID: [14524837](https://pubmed.ncbi.nlm.nih.gov/14524837/)
50. Xulvi-Brunet R, Sokolov IM. Reshuffling scale-free networks: from random to assortative. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*. 2004;70(6 Pt 2):066102.
51. Molloy M, Reed B. A critical point for random graphs with a given degree sequence. *Random Struct Algorithms*. 1995;6(2–3):161–80. <https://doi.org/10.1002/rsa.3240060204>
52. Maslov S, Sneppen K. Specificity and stability in topology of protein networks. *Science*. 2002;296:910–3. <https://doi.org/10.1126/science.1065103>
53. Vázquez A. Growing network with local rules: preferential attachment, clustering hierarchy, and degree correlations. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2003;67(5 Pt 2):056104. <https://doi.org/10.1103/PhysRevE.67.056104> PMID: [12786217](https://pubmed.ncbi.nlm.nih.gov/12786217/)
54. Holland PW, Laskey KB, Leinhardt S. Stochastic blockmodels: first steps. *Social Networks*. 1983;5(2):109–37. [https://doi.org/10.1016/0378-8733\(83\)90021-7](https://doi.org/10.1016/0378-8733(83)90021-7)
55. Karrer B, Newman MEJ. Stochastic blockmodels and community structure in networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2011;83(1 Pt 2):016107. <https://doi.org/10.1103/PhysRevE.83.016107> PMID: [21405744](https://pubmed.ncbi.nlm.nih.gov/21405744/)
56. Airoldi EM, Blei DM, Fienberg SE, Xing EP. Mixed membership stochastic blockmodels. *J Mach Learn Res*. 2008;9:1981–2014. PMID: [21701698](https://pubmed.ncbi.nlm.nih.gov/21701698/)
57. Colizza V, Flammini A, Serrano M, Vespignani A. Detecting rich-club ordering in complex networks. *Nat Phys*. 2006;2. <https://doi.org/10.1038/nphys209>
58. Holme P, Kim B, Yoon C, Han SK. Attack vulnerability of complex networks. *Physical Review E, Statistical, Nonlinear, and Soft Matter Physics*. 2002;65:056109. <https://doi.org/10.1103/PhysRevE.65.056109>
59. Schneider CM, Moreira AA, Andrade Jr JS, Havlin S, Herrmann HJ. Mitigation of malicious attacks on networks. *Proc Natl Acad Sci U S A*. 2011;108(10):3838–41. <https://doi.org/10.1073/pnas.1009440108> PMID: [21368159](https://pubmed.ncbi.nlm.nih.gov/21368159/)
60. Newman MEJ. Spread of epidemic disease on networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2002;66(1 Pt 2):016128. <https://doi.org/10.1103/PhysRevE.66.016128> PMID: [12241447](https://pubmed.ncbi.nlm.nih.gov/12241447/)
61. Boguñá M, Pastor-Satorras R, Vespignani A. Absence of epidemic threshold in scale-free networks with degree correlations. *Phys Rev Lett*. 2003;90(2):028701. <https://doi.org/10.1103/PhysRevLett.90.028701> PMID: [12570587](https://pubmed.ncbi.nlm.nih.gov/12570587/)
62. Gómez-Gardeñes J, Moreno Y, Arenas A. Paths to synchronization on complex networks. *Phys Rev Lett*. 2007;98(3):034101. <https://doi.org/10.1103/PhysRevLett.98.034101> PMID: [17358685](https://pubmed.ncbi.nlm.nih.gov/17358685/)
63. Kelly D, Gottwald GA. On the topology of synchrony optimized networks of a Kuramoto-model with non-identical oscillators. *Chaos*. 2011;21(2):025110. <https://doi.org/10.1063/1.3590855> PMID: [21721788](https://pubmed.ncbi.nlm.nih.gov/21721788/)
64. Jorritsma J, Komjáthy J, Mitsche D. Large deviations of the giant in supercritical kernel-based spatial random graphs. *arXiv preprint*. 2024. <https://arxiv.org/abs/240402984>
65. Gracar P, Grauer A, Mörters P. Chemical distance in geometric random graphs with long edges and scale-free degree distribution. *Commun Math Phys*. 2022;395(2):859–906. <https://doi.org/10.1007/s00220-022-04445-3>
66. Lühtrath L. All spatial random graphs with weak long-range effects have chemical distance comparable to Euclidean distance. 2024. <https://arxiv.org/abs/241212796>
67. Biskup M. On the scaling of the chemical distance in long-range percolation models. *Ann Probab*. 2004;32(4). <https://doi.org/10.1214/009117904000000577>
68. Norros I, Reittu H. On a conditionally Poissonian graph process. *Advances in Applied Probability*. 2006;38(1):59–75. <https://doi.org/10.1239/aap/1143936140>

69. Janson S. Asymptotic equivalence and contiguity of some random graphs. *Random Structures and Algorithms*. 2010;36. <https://doi.org/10.1002/rsa.20297>
70. Penrose M. *Random geometric graphs*. OUP Oxford; 2003.
71. Gracar P, Lühtrath L, Mönch C. Finiteness of the percolation threshold for inhomogeneous long-range models in one dimension. 2022. <https://arxiv.org/abs/2203.11966>
72. van der Hofstad R, van der Hoorn P, Maitra N. Scaling of the clustering function in spatial inhomogeneous random graphs. *J Stat Phys*. 2023;190(6). <https://doi.org/10.1007/s10955-023-03122-6>
73. Hall P. On continuum percolation. *Ann Probab*. 1985;13(4). <https://doi.org/10.1214/aop/1176992809>
74. Yukich JE. Ultra-small scale-free geometric networks. *Journal of Applied Probability*. 2006;43(3):665–77. <https://doi.org/10.1239/jap/1158784937>
75. Jacob E, Mörters P. Spatial preferential attachment networks: power laws and clustering coefficients. *Ann Appl Probab*. 2015;25(2). <https://doi.org/10.1214/14-aap1006>
76. Hill BM. A simple general approach to inference about the tail of a distribution. *Ann Statist*. 1975;3(5). <https://doi.org/10.1214/aos/1176343247>