Geometric origins of self-similarity in the evolution of real networks

Muhua Zheng,1,2 Guillermo García-Pérez,3,4 Marián Boguñá,1,2 and M. Ángeles Serrano1,2,5

1Departament de Física de la Matèria Condensada, Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain
2Universitat de Barcelona Institute of Complex Systems (UBICS), Universitat de Barcelona, Barcelona, Spain
3QTF Centre of Excellence, Turku Centre for Quantum Physics, Department of Physics and Astronomy, University of Turku, FI-20014 Turun Yliopisto, Finland
4Complex Systems Research Group, Department of Mathematics and Statistics, University of Turku, FI-20014 Turun Yliopisto, Finland
5ICREA, Passeig Lluís Companys 23, E-08010 Barcelona, Spain

(Dated: December 3, 2019)

Abstract

One of the aspirations of network science is to explain the growth of real networks, often through the sequential addition of new nodes that connect to older ones in the graph. However, many real systems evolve through the branching of fundamental units, whether those be scientific fields, countries, or species. Here, we provide empirical evidence for self-similar branching growth in the evolution of real networks and present the Geometric Branching Growth model, which is designed to predict evolution and explain the symmetries observed. The model produces multiscale unfolding of a network in a sequence of scaled-up replicas. Practical applications in real instances include the tuning of network size for best response to external influence and finite-size scaling to assess critical behavior under random link failures.
One-sentence summary: Over large time spans, real networks are found to evolve in a self-similar way that can be explained by geometric branching growth: a process that produces a multiscale unfolding of a real network in a sequence of scaled-up replicas.

INTRODUCTION

From the universal common ancestor, the evolution of life has resulted in the unfolding of a phylogenetic tree of living beings in which populations develop and diversify gradually to become distinct species. The evolutionary drift towards diversification is not exclusive to biological evolution: other processes—such as the growth of human knowledge [1] and recent 20th century historical geopolitical development [2]—also tend towards branching growth in terms of the fundamental units of the system, whether those are species, scientific fields, or countries. In many cases, these evolutionary mechanisms lead to tree-like relationship structures, where contemporary elements are related to one another only through common ancestors. Such trees underlie successful models of complex systems, for instance when coupled to the spatial environment to explain the morphogenesis of branched organs [3] and the metapopulation stability of branching river networks [4]. In other processes, the branching drift does not compel the system to grow in a simple dendritic fashion, but new links can form between the different coexisting components, which interact through a complex network of relationships that typically becomes more dense over time [5].

In the context of network science, growth is most often modeled through the sequential addition of new nodes that connect to older ones in a graph by preferential attachment [6, 7]. Other growth processes have also been considered, such as duplication to explain large-scale proteome evolution [8, 9]. Here, we take an alternative approach and explore the relation between branching growth [10] and geometric renormalization [11] to explain self-similar network evolution. Renormalization in networks, based on the ideas of the renormalization group in statistical physics [12,14], acts as a sort of inverse branching process by coarsegraining nodes and rescaling interactions. Thus, branching growth can be seen as an inverse renormalization transformation: an idea that was introduced in [15] using a purely topological approach to reproduce the structure of fractal networks, where fractality was interpreted as an evolutionary drive towards robustness. However, topological distances in networks are
seriously constrained by the small-world property; while the characterization of fractality in real networks disregards fundamental features of their structure, including clustering and community organization.

Geometric renormalization [11] (GR) is an alternative technique that can be performed by virtue of the discovery that the structure of real networks is underlain by a latent hyperbolic geometry [16, 17]. Thus, the likelihood of interactions between nodes depends on their distances in the underlying space, via a universal connectivity law that operates at all scales and simultaneously encodes short- and long-range connections. This approach has been able to explain many features of the structure of real networks, including the small-world property, scale-free degree distributions, and clustering, as well as fundamental mechanisms such as preferential attachment in growing networks [7] and the emergence of communities [18, 19]. Given a network map, GR produces a multiscale unfolding of the network in scaled-down replicas over progressively longer length scales. This transformation has revealed self-similarity to be a ubiquitous symmetry in real networks, whose structural properties remain scale-invariant as the observational resolution is decreased [11]. This poses the question of whether this self-similarity could be related to the mechanisms driving the growth of real networks and, therefore, whether their evolution could be conceptualized within the framework of the geometric renormalization group.

In this work, we show that real networks—citations between scientific journals [20, 21] and international trade [22]—have evolved in a self-similar way over time spans of more than 100 years, meaning that their local, mesoscale, and global topological properties remain in a steady state as time goes by, with a moderate increase of the average degree. We demonstrate that the observations can be modeled by a self-similar metric expansion produced by a geometric branching growth (GBG) process. Beyond the capacity of the model to explain and predict the self-similar evolution of real networks effectively, the technique is flexible and allows us to produce scaled-up network replicas that, when combined with scaled-down network replicas [11], provide a full up-and-down self-similar multiscale unfolding of complex networks that covers both large and small scales. We illustrate the use of GBG multiscale unfolding in real network instances via the tuning of network size for optimal response to an external influence, referred to here as the optimal mass, and a finite-size scaling analysis of critical behavior under random link failures.
We consider the evolution of the journal citation network \cite{20} (JCN) and of the world trade web \cite{22} (WTW) over time spans of more than 100 years.

The evolution of journal citation networks offers a quantitative proxy for the development of contemporary science and the emergence of a vast number of new scientific fields and subfields, driven by diversification and specialization \cite{20, 21, 23-25}. Here, we analyze data from \cite{20}, where the time period 1900-2013 is divided into time windows of ten years before 1970 and of five years thereafter. One citation network is reconstructed for each time window, where journals are represented as nodes that are linked whenever citations between their publications exist.

An increase of the number of actors is also a hallmark of the evolution of the international trade system. The number of states in the world increased from 42 in 1900 to 195 in 2016 \cite{26}, mainly due to processes such as decolonization, the dismantlement of large or multicultural states such as the USSR and Yugoslavia (1991) into a number of smaller states, the parliamentary split of an existing state into two as happened in Czechoslovakia (1993), and independence processes after civil wars, like that of the Republic of South Sudan and the Republic of the Sudan (2011). Here, we use networks in the World Trade Atlas \cite{22}: a collection of annual world trade network maps in hyperbolic geometry, which provide information on the long-term evolution of the international trade system from 1870 to 2013, where nodes represent countries linked by bilateral trade relationships. The maps revealed that globalization, hierarchization, and localization are main forces shaping the trade space, which far from being flat is hyperbolic, as a reflection of its complex architecture. More details of the two datasets are available in the Supplementary Materials (SM) and the main statistical properties are in Tables S1 and S2.

The size $N$ of the two evolving networks increases over time ranging from 118 journals in 1900-1910 to 21460 in 2008-2013, and from 24 countries in 1870 to 189 in 2013, Fig. S1a-b in SM. After World War II, the average degree $\langle k \rangle$ only shows a moderate increase in the JCN and almost flat behavior in the WTW, Fig. S1c-d. Degree distributions, clustering spectra, degree-degree correlations and the community structure of some snapshots are shown for the JCN in Fig. 1A-D and for the WTW in Fig. 1E-H (results for all snapshots are in SM, Figs. S2-S4). We observe clear-cut self-similar behavior with the curves for different
networks overlapping when the degrees of the nodes are rescaled by the average degree of the corresponding network. Fig. 1D and H shows the modularity, $Q$, of the optimal partitions detected by the Louvain method [27], and the adjusted mutual information AMI [28] between the optimal partitions of two consecutive snapshots, in which we only considered the nodes that exist in both. The level of modularity remains stable throughout the evolution of the systems and the overlap between communities in the consecutive snapshots is consistently very high. This indicates that the community structure is mostly preserved as time goes by. Hence, the empirical evidence presented so far indicates that these real networks grow in a self-similar fashion.
THE GEOMETRIC BRANCING GROWTH MODEL

To model the observed self-similar evolution of real networks, we propose the GBG transformation that produces self-similar multiscale unfolding of a network in a shell of scaled-up replicas of progressively increasing size, as illustrated in Fig. 2A. The GBG transformation acts on network maps: geometric representations that reveal the manifest latent hyperbolic geometry of network structure [16, 29]. To describe network maps, we employ the S^1 model [16] which, in contrast to the isomorphic hyperbolic version H^2 [29], makes the similarity dimension explicit. In the S^1 model, each node, i, is assigned a hidden degree, κ_i, or popularity, and an angular position, θ_i, or similarity, in a one-dimensional sphere representing the similarity space; and every pair of nodes, i and j, is connected with probability:

\[ p_{ij} = \frac{1}{1 + \chi_{ij}^\beta} = \frac{1}{1 + \left(\frac{R \Delta \theta_{ij}}{\mu \kappa_i \kappa_j}\right)^\beta} \tag{1} \]

so that more popular (larger κ) or more similar (lower Δθ) nodes are more likely to form connections. The similarity circle has radius R, adjusted to maintain a constant density of nodes, equal to one, without loss of generality. A network map, that is, the set of hidden variables \{κ_i, θ_i\} together with the parameters β and μ controlling the local clustering coefficient \langle c \rangle and the average degree \langle k \rangle, can be obtained by finding the coordinates that maximize the probability for the observed network to be generated as an instance of the S^1 model. Due to the aforementioned isomorphism between S^1 and H^2, these model parameters and hidden variables also define the corresponding hyperbolic map [30]. To produce the maps we used the embedding tool Mercator [31], which infers the coordinates of the nodes and parameters R, β, and μ from the topology of the network. More details can be found in SM.

The GBG transformation can be controlled to adjust the growth in the number of nodes and also the flow of the average degree, embodying a family of models that includes non-inflationary and inflationary growth. Non-inflationary growth produces a sequence of progressively magnified layers with decreasing average degree that comply with GR. Inflationary growth means that scaled-up shell layers are produced with an average degree that does not decrease very fast, or even increases. The GBG transformation is compliant with GR if, when GR is applied to the layer obtained after the GBG transformation, the result is the original network. In brief, the GR transformation [11] proceeds by defining non-overlapping blocks of consecutive nodes of equal size r around the similarity circle, which are then coarse
grained into a single node in the renormalized lower-resolution map, where pairs of nodes are connected with a link if any of their precursor nodes were connected in the original layer. As a result, multiscale unfolding of self-similar scaled-down network replicas is obtained, except for the average degree of the renormalized layers, which typically grows exponentially in real networks (more details in SM).

The first step to generate a GBG scaled-up map is to split every node in the original layer into \( r \) descendants with probability \( p \), so that the population increases as \( N' = N(1 + p(r - 1)) = bN \) with branching rate \( b \). For mathematical convenience, we will continue the description with \( r = 2 \) (\( b = 1 + p \)). We can use parameter \( b \) in combination with the number of layers in the multiscale unfolding to adjust the growth of the number of nodes over the evolution of the network. Every branching node produces a pair of descendants that require the assignment of similarity coordinates in the \( S^1 \) circle and hidden degrees. The radius of the circle is rescaled as \( R' = bR \), so that the density of nodes remains equal to one.

**Assigning coordinates to descendants.** One of the requirements for self-similar growth is the preservation of the ordering of nodes in the circle and their concentration across specific angular sectors, which defines the geometric community organization of networks [18, 19]. To this end, the simplest means to model growth is to place the descendants at angular coordinates \( \theta_i^+ \) and \( \theta_i^- \) to the left and right of the angular position of their corresponding ancestor, \( i \), with uniform probability within a small angular separation \( \Delta \theta^\pm \). The values \( \Delta \theta^\pm \) are bounded by the total number of nodes in the descendant layer and by the proximity (to the left or right) of consecutive nodes to the ancestor in the similarity circle. We set \( \Delta \theta^\pm = \min\{\frac{2\pi}{N'}, \frac{\Delta \theta_{ij}}{2}\} \), where \( \Delta \theta_{ij} = \pi - |\pi - |\theta_i - \theta_j|| \) is the angular distance between the branching node \( i \) and its consecutive neighboring node \( j \) (to left or right) in the ancestor layer.

To assign the hidden degrees \( \kappa^+ \) and \( \kappa^- \), we impose two conditions. First, the hidden degrees of ancestors and descendants need to comply with GR. This implies that the relation between hidden degrees of ancestors and descendants should be compliant with the GR transformation \( z = z^+ + z^- \), where \( z = \kappa^\beta \). Second, the hidden degrees of descendants must be independent and identically distributed random variables with a distribution of hidden degrees that preserves that of the ancestor layer, \( \rho(\kappa) \) (equivalently \( \rho(z) \)). Taking
FIG. 2. Sketch of the GBG model of complex networks. (A) In each layer of the self-similar upwards multiscale unfolding, the size of each node is proportional to the logarithm of its hidden degree, \( \kappa \). Different colors represent different geometric communities. Solid gray lines on the blue disks represent links in each layer. Dashed lines connect ancestors to their descendants along the flow (blue arrows). A pair of nodes, \( i, j \), with hidden degree \( \kappa_i \) and \( \kappa_j \) has been highlighted, for which the angular separation, \( \Delta \theta_{ij} \), represents their similarity distance. In the lower left corner is a magnified sketch of the branching process from an ancestor, \( i \), to its two corresponding descendants. Also shown is the complementary cumulative distribution of \( z \) snapshots, 1965-1975 in the JCN (B) and 1960 in WTW (D), together with their corresponding stable distribution fittings. (C) and (E) show the \( P_c(z^\pm) \) distribution of different descendant layers (\( l = 0 \) indicates the original network) in the JCN and WTW, respectively.

The two conditions together, the transformed hidden degrees \( z^\pm \) of descendants should satisfy:

\[
\int \int dz^+ dz^- \rho(z^+)\rho(z^-)\delta \left( z - (z^+ + z^-) \right) = \rho(z). \tag{2}
\]

This equation implies that \( \rho(z) \) is a stable distribution [32], meaning that the linear combination of two independent variables with probability distribution \( \rho(z) \) has the same distribution, up to scaling and location factors. Stable distributions admit multiple parametrizations but are always defined by four parameters \( f(z; \alpha, \eta, c, d) \): the tail exponent \( \alpha \in (0, 2] \) and skewness \( \eta \in [-1, 1] \) which control the shape; and \( c \) and \( d \) for scale and location. Stable
distributions conform a rich family of models including Gaussian ($\alpha = 2$), Cauchy ($\alpha = 1$ and $\eta = 0$), Lévy ($\alpha = 1/2$ and $\eta = 1$), and Landau ($\alpha = 1$ and $\eta = 1$) distributions. Stable distributions are infinitely divisible and are the only possible limit distributions for properly normalized and centered sums of iid random variables (generalized Central Limit Theorem) \[33\]. In addition, they can accommodate fat tails and asymmetry, and therefore often offer a very good fit for empirical data \[34\]–\[37\].

We proceed by fitting a stable distribution to the distribution of hidden degrees in the original layer \[38\]–\[39\] (see technicalities in SM). Fig. 2B and D show a very good fit for the JCN and WTW (see Fig. S5 for more empirical networks, and the corresponding fitting parameters in Table S3). So, the distribution for descendants, $f(z^\pm; \alpha^\pm, \eta^\pm, c^\pm, d^\pm) = f(z^\pm; \alpha, \eta, c/2^{1/\alpha}, d/2)$, follows immediately from Eq. (2) and basic properties of the stable distribution, with the shape parameters remaining invariant, and scale and location being adjusted so that the stable distribution of the ancestor layer is recovered when we sum the hidden variables $z^\pm$ of the descendants. These functions and Bayes rule can be used to generate numerically the values of $z^+$ from the probability of hidden degrees of descendants, conditional on the degree of the ancestor $\rho(z^+|z)^{nor}$, normalized to ensure that the hidden degrees of descendants are non-negative. Finally, $z^-$ is calculated deterministically using $z^- = z - z^+$, and the variables $z^\pm$ are transformed back into $\kappa^\pm$ using $\kappa = z^{1/\beta}$. Fig. 2C and E demonstrates that the distribution of hidden variables $z^\pm$ of the descendants has the same shape as the ancestor layer for the JCN and WTW. See Fig. S5 for more empirical networks.

**Connecting nodes in the descendant layer.** Once coordinates have been assigned to nodes, connections between descendants in the new layer are implemented such that the resulting network belongs to the $S^1$ ensemble. In what we call the non-inflationary limit, we also require that the new network is compliant with GR, that is, GR applied to the descendant layer should result in the ancestor layer. We use the probability of connection $p_{ij}$ Eq. (6) as in the $S^1$ model, rescaling $\mu$ in the new layer to control the flow of the average degree, and with $\beta$ remaining invariant as in the GR transformation. We use $\mu' = b\mu$ and connect descendants branching from the same ancestor with probability $p_{ij}(\mu')$. Then, for every pair of connected ancestors, we establish potential links among their descendants with the same probability $p_{ij}(\mu')$, but making sure that at least one link is formed between them (see details in SM).
FIG. 3. **GBG is compliant with GR in the non-inflationary limit.** (A) Sketch illustrating the GR-GBG transformation in the Internet network. More specifically, we renormalize the original network from layer 4 to layer 0 and recover the statistically equivalent network applying GBG to layer 0 with $b = 1.2$. The colors of the nodes indicate the community structure detected by the Louvain algorithm. (B) Flow of the average clustering. (C) Flow of the average degree. (D) Flow of the connection probability $p^{(l)}(\chi^{(l)}_{ij})$ as a function of the effective distance $\chi^{(l)}_{ij}$, where $l$ indicates the layer. (E) Sketch illustrating the GBG-GR transformation in the human metabolic network. We transform the original network applying GBG with $b = 1.2$ to produce 4 layers and recover the original network with GR applied to the last layer. (F to H) Analogous to (B to D).

As a result, non-inflationary GBG is a statistical inverse of GR. We support this claim with the results shown in Fig. 3 and Figs. S6-S13 in SM for different real networks and branching rates $b$, which show that the results are robust. In Fig. 3A-D, we show that, after applying GR to the Internet, the original network can be recovered with high fidelity (in a statistical sense) by applying the non-inflationary GBG to the renormalized layer. Conversely, if we first apply the non-inflationary GBG technique, to obtain the scaled-up network, and then recover the networks by geometric renormalization, the result is analogous; see Fig. 3E-H for the metabolic network (details of the datasets in SM). This means that non-inflationary GBG and GR flows produce the same values of average clustering, average degree and empirical connection probability, among other properties. Notice that the GBG
transformation also preserves the original community structure, as detected by the Louvain algorithm \cite{27}.

In the inflationary version of GBG, we first proceed as in the non-inflationary case. Once we have a non-inflationary GBG map, we set $\mu_a = a\mu' = ab\mu, (a \geq 1)$ to adjust the average degree to a larger value by adding extra links between any pair of nodes that remained unconnected using probability:

$$\pi_{ij} = \frac{p_{ij}(\mu'_a) - p_{ij}(\mu')}{1 - p_{ij}(\mu')}.$$  \hspace{1cm} (3)

These steps ensure that: i) all pairs of descendants in the GBG layer are connected with probability $p_{ij}(\mu'_a)$, with the original form Eq. (6) in the ancestor layer, and hence the resulting network belongs to the $S^1$ ensemble; ii) links exist between descendants of connected ancestors; and iii) the non-inflationary limit is recovered for $a = 1$, that is, in this case, $\pi_{ij} = 0$ and no extra links are formed so that GBG complies with GR and there are only connections in the descendant layer between descendants of the same ancestor or of connected ancestors.

**Behavior of the average degree.** In the non-inflationary GBG model ($a = 1$), we can use the inverse of the GR relation between the average degrees in a descendant layer and in the ancestor layer \cite{11}, using $\mu' = b\mu$ to obtain $\langle k \rangle^{(l)} = (b^{-\nu})^l \langle k \rangle^{(0)}$, where the scaling factor $\nu$ depends on the connectivity structure of the original network, and $\langle k' \rangle^{(l)}$ (the mean degree of layer $l$ and $l = 0$) refers to the original network. Typically, as the scaling factor, $\nu$, is positive in real networks \cite{11}, the average degree of the descendant layers decreases exponentially.

In the inflationary regime, $\mu' = ab\mu$ and following the same derivations as in \cite{11}, we find:

$$\langle k \rangle^{(l)}_a = a^l \langle k \rangle^{(l)} = (ab^{-\nu})^l \langle k \rangle^{(0)} = \left[ \frac{N^{(l)}}{N^{(0)}} \right]^{-\nu + \frac{\ln a}{\ln b}} \langle k \rangle^{(0)},$$  \hspace{1cm} (4)

where we have used $l = \frac{\ln N^{(l)}}{\ln b}$, with $N^{(l)}$ and $N^{(0)}$ being the network sizes on layer $l$ and 0, respectively; and $N^{(l)} = bN^{(l-1)}$. Notice that the inflationary process was applied here to every layer in the flow. If, instead, it is applied in a single step to the last layer produced in a non-inflationary GBG transformation, then $\langle k \rangle^{(l)}_a = a^l \langle k \rangle^{(l)} = a(b^{-\nu})^l \langle k \rangle^{(0)}$. From Eq. (4), the average degree $\langle k \rangle^{(l)}_a$ increases as a power of $N^{(l)}$. Fig. S14 shows the high degree of congruency between this theoretical prediction, the empirical data, and simulations (as explained below) of the inflationary version of GBG, in the JCN and WTW.
FIG. 4. The GBG model predicts the self-similar evolution of real networks. (A) and (B) Evolution of the average degree \( \langle k \rangle \) vs network size \( N \). The estimation, validation, and projection sections are separated by vertical dashed lines. Blue circles, green squares and red stars represent empirical data, validation points, and projection from the model, respectively. The data in the estimation section are used to find values of \( a \) (see Methods). When the branching rates, \( b \), are fixed to 1.5 in JCN and 1.2 in WTW, the corresponding values of \( a \) are: 1.452 and 1.130, see Fig. S15 in SM. For validation purposes, we grow the network from 1965-1975 in JCN and 1960 in WTW, using GBG and compared the resulting networks with empirical snapshots of the same size. (C to J) Comparison of the topological properties of simulated and empirical networks. (C, E) Complementary cumulative distribution of rescaled degrees. (D, F) Degree-dependent clustering coefficient over rescaled-degree classes, insets: Degree–degree correlations; (G, I), the modularity \( Q \); (H, J), local rich-club effect in the JCN and the WTW, respectively.

Predicting the evolution of real networks. The inflationary GBG model reproduces the self-similar evolution of JCN and WTW. To support this claim, we divide the empirical
data into two consecutive time windows: the first for estimation purposes and the second for validation purposes. Note that the JCN and WTW data from before World War II are not used due to the high fluctuations of the network properties, see Figs. S2 and S3. We fix a value of \( b \) in the range \( 1 < b < 2 \) to adjust the rate of growth in our GBG simulation in such a way that we can produce enough snapshots to compare with the real data. With this value of \( b \), we estimate parameter \( a \) from the empirical evolution of the average degree vs network size (see SM). We find that \( a \) remains stable over time (see Fig. S15 in SM), consistent with the empirical observation that the average degree grows as a power of the system size, see Fig. S19A and B. Next, we use the network snapshot at the end of the estimation period as the initial layer in GBG multiscale unfolding to simulate a number of scaled-up layers that we then compare to empirical snapshots of approximately the same size in the validation set. The comparisons of degree distributions, clustering, degree–degree correlations, and modularity are shown in Fig. S19C, D, G and E, F, I. We also measured the local rich-club and nested self-similarity effects, reported in Fig. S19H and J and Figs. S16 and S17. We name as "local rich-club effect" and "nested self-similarity effect" the observation in real networks that the nested hierarchy of subgraphs produced by progressively thresholding the degrees of the nodes presents, respectively, an increasing internal average degree and self-similar structure [16, 40]. This is a highly non-trivial property with crucial implications, such as the absence of a critical threshold in any phase transition whose critical point depends monotonously on the average degree, including percolation, epidemic spreading processes and the Ising model [41]. The results show that all the networks analyzed in this paper, including JCN and WTW, present the two effects, see Figs. S16 and S17 in SM. Notice that standard growing network models, including the Barabási-Albert model [6] and the Popularity–Similarity Optimization model in hyperbolic space [7], have a constant average degree as the network grows, and they also present a constant average degree of the subgraphs in the nested hierarchy, see Fig. S18 in SM. Therefore, they lack the local rich-club effect. In fact, if those models were adjusted to increase the average degree over time, as happens in the real networks that we analyze in this work, the flow of the average degree in the nested hierarchy would be decreasing, see Fig. S18b in SM. In addition, the results are robust for different values of \( b \), and for different starting times, see Fig. S19 in SM. Therefore, the GBG model reproduces the self-similar evolution of the structure of the two networks with high fidelity. More comparisons between the model and empirical observations are also shown in
SCALED-UP REAL NETWORK REPLICA

One of the practical applications of the GBG model is the production of magnified replicas of real networks: versions where the number of nodes is increased while preserving the statistical properties of the original network, in particular its average degree $\langle k^{(0)} \rangle$. Using GBG, the procedure is straightforward and involves adjusting the parameter $b$, the number of layers $l$, and the inflationary parameter $a$. The idea is to single out a specific scale after a certain number of non-inflationary GBG steps and to tune $a$ to increase the average degree to the target value by adding new links using Eq. (3) (see Methods).

We illustrate the usefulness of scaled-up network replicas through two examples. In the first, we use the upwards self-similar multiscale unfolding of a small Facebook network to detect the network size that produces the optimal response to external modulation in a noisy environment (see Fig. S24 for the self-similar statistical properties with respect to the original network). In the second, we combine the upwards self-similar multiscale unfolding of GBG with the downwards self-similar multiscale unfolding of GR to describe the critical behavior of a real network, the Internet (see the same topological properties of both scaled-up and scaled-down networks in Fig. S24), close to the transition where the global connectivity of the network disintegrates under random link failures.

**Size-dependent system response to external modulation in a noisy environment.**

Small size in real networks can be a limiting factor for the study of dynamical processes, especially when long-range dynamical correlations are non-negligible or when finite size effects play an important role in the final outcome of the dynamics. Here, we study the behavior of a model of opinion formation with nontrivial size dependence using a small Facebook network of 320 users, working for the same software company [42] (see data description in SM). The opinion formation model introduced in [43] includes imitation following a majority rule, external influence in the form of a periodic “fashion” wave, and noise. This model was shown to present a noise stochastic resonance effect in small-world networks [44], displaying an optimal response of the population to the “fashion” wave for some noise level. The system also displays a size stochastic resonance effect [45, 46], which means there is an optimal value for the number of nodes, the optimal mass, for which the average opinion best follows
FIG. 5. **Optimal mass for opinion formation in a Facebook network.** The external signal is modeled as a cosine function with amplitude $A = 0.22$ and period $T = 200$. The temporal output of each node is recorded for $2 \times 10^5$ time units, discarding the first $10^5$ time units as transient. For every point, we averaged over 100 realizations of the dynamics. The GBG multiscale unfolding was produced with $b = 2$. (A) System response, $R$, as a function of the noise intensity $\epsilon$ for different network sizes. The inset shows the maximum response as a function of $N$. (B) System response, $R$, as a function of $N$ for different values of the noise intensity $\epsilon$. The range of noise values on the $x$-axis corresponds to the gray region in plot (A).

The GBG technique provides the opportunity to study system size stochastic resonance in real networks. We produce a GBG self-similar multiscale shell with the same average degree as the original Facebook network, and simulate the dynamical process described above in each layer (see SM for details of our implementation). We model the external signal as a cosine function with amplitude $A$ and period $T$, and measure the response of the system as a function of the noise intensity $\epsilon$ for different system sizes using the spectral amplification factor $R = 4A^{-2}|\langle e^{i2\pi t/T} \rho(t) \rangle|$, where $\rho(t) = \frac{1}{N} \sum_i m_i(t)$ is the average opinion in the
evolution, $m_i(t)$ is the dynamical state of node $i$ at time $t$, and $\langle \cdots \rangle$ denotes a time average.

Our results are shown in Fig. 5A. The optimal response, $R_{\text{max}}$, is plotted in the inset as a function of $N$. For each size, $N$, there is a maximum response for some intermediate value of the noise and the optimal value occurs at some combination of noise and size. Interestingly, for sufficiently small values of noise, Fig. 5B, $R$ is enhanced by increasing the noise; and for every noise intensity, $\epsilon$, the optimal response occurs at approximately the same value as in $R_{\text{max}}$, $N = 2356$. Hence, we conclude that there is an optimal mass for which the average opinion best follows the external influence. Moreover, we also found that there is some value of $N$ for which $R$ has a minimum, that is, the average opinion follows the external influence to the least extent.

**Critical behavior of real networks under random link failures.** The random failure of links in networks leads to a percolation transition: a continuous structural change which

![Image of graphs showing critical behavior](image-url)
disaggregates the large cluster of connected network nodes into a bundle of small isolated components [40, 41], hence disabling the system. The fraction of links removed, \( p \), acts as a control parameter which can be manipulated to change the state of the system in \textit{in silico} experiments, and the transition occurs at some specific value: \( p_c \). Close to this critical point, the macroscopic properties of the network, such as the relative size of the largest connected cluster and the average cluster size, behave as power laws of the distance to the critical point, \( (p - p_c)^\delta \), with some critical exponents. One way of extracting these exponents is by observing how certain quantities vary as the size of the system changes. However, the finite size scaling technique has faced serious challenges in real networks due to the lack of data beyond single snapshots.

Next, we prove that a downwards–upwards multiscale shell of replicas produced by the joint action of the GBG and GR techniques on a real network can be used to study the finite size scaling behavior of bond percolation, Fig. 6. In each layer, we measure the average size of the largest component, \( \langle G \rangle \), and its fluctuations, i.e., susceptibility \( \chi = \frac{\langle G^2 \rangle - \langle G \rangle^2}{\langle G \rangle} \), for each combination of \( (p, N) \) in the multiscale shell using the fast algorithm of Newman and Ziff [48]. In finite systems, a peak in the susceptibility, \( \chi \), diverging with the system size indicates the presence of a continuous phase transition, and its position provides a way to estimate the percolation threshold, \( p_c \): Fig. 6A. In Fig. 6B, we show that the critical link failure probability, \( p_c \), approaches zero as a power law, \( p_c(N) \sim N^{-0.358} \), and the maximum, \( \chi_{\text{max}} \), of the susceptibility also diverges as a power law: \( \chi_{\text{max}}(N) \sim N^{0.461} \). Not only do these results suggest a vanishing percolation threshold in the real Internet graph, as usually happens in scale-free networks, but they also provide a way to estimate the corresponding critical exponents numerically, thus offering a new way to study critical phenomena in single-instance real networks.

CONCLUSIONS

Real networks are observed to evolve in a self-similar way that preserves their topology throughout the growth process over long time spans. The GBG model lays out a minimal number of simple principles that combine branching growth, one of the paradigms of evolution, and network geometry, to explain the empirical findings via a technique that generates self-similar metric expansion of a network replicating its original structure. One of the
essential assumptions in the model, the preservation of the distribution of hidden degrees as the number of nodes increases, leads to the introduction of stable distributions in the context of network modeling. Stable distributions, a rich family of probability distributions with intriguing theoretical and practical properties, are widely used to model heavy-tailed data from many types of physical and economic systems, and represent an alternative to the power law paradigm in the study of complex networks. Meanwhile, the geometric branching growth model relies on a universal connectivity law that operates at all scales, simultaneously encoding short- and long-range connections, which keeps its form over time. Our results suggest that the same principles organize network connectivity at different length scales in real networks and that these principles are also sustained over time. As a result, simplicity, as one of the rationales for self-similarity, is one of the keys to understanding and predicting network evolution.

While some limitations of our model are obvious, for instance the exclusion of the birth/death processes of links and nodes, we believe that complementary hypotheses would not affect the results and our GBG model in any fundamental way. The model captures the main mechanisms that drive and predict the self-similar evolution of real networks. In parallel, and beyond the explanatory power of the model to effectively decode the self-similar evolution of real networks, GBG is also a technique to produce scaled-up replicas of networks: an effective and versatile tool facilitating analysis of the behavior of networks at different size scales. The combination of GBG with scaled-down network replicas produced by GR provides full up-and-down self-similar multiscale unfolding of complex networks that covers both large and small scales. Potential applications that require optimization or control of system size in complex systems are countless. Apart from those explained here, we can mention the assessment of scalability issues in dynamic processes in core functions of real networks, such as in Internet routing protocols.

[1] V. Allee, The knowledge evolution (Routledge, 2012).
[2] J. Tir, P. Schafer, P. F. Diehl, G. Goertz, Territorial changes, 1816–1996: Procedures and data. Conflict Management and Peace Science 16, 89 (1998).
[3] E. Hannezo, et al., A unifying theory of branching morphogenesis. *Cell* **171**, 242 (2017).

[4] A. Terui, et al., Metapopulation stability in branching river networks. *Proceedings of the National Academy of Sciences* **115**, E5963 (2018).

[5] N. D. Sergey, F. F. M. Jose, *Handbook of Graphs and Networks: From the Genome to the Internet*, eds. S. Bornholdt and H.G. Schuster (Wiley-VCH, Berlin, 2002) (2002), pp. 318–341.

[6] A. Barabási, R. Albert, Emergence of scaling in random networks. *Science* **286**, 509 (1999).

[7] F. Papadopoulos, M. Kitsak, M. Á. Serrano, M. Boguñá, D. Krioukov, Popularity versus similarity in growing networks. *Nature* **489**, 537 (2012).

[8] R. V. Solé, R. Pastor-Satorras, E. Smith, T. B. Kepler, A model of large-scale proteome evolution. *Advances in Complex Systems* **5**, 43 (2002).

[9] R. Pastor-Satorras, E. Smith, R. V. Solé, Evolving protein interaction networks through gene duplication. *Journal of Theoretical biology* **222**, 199 (2003).

[10] A. N. Kolmogorov, N. A. Dmitriev, Branching random processes. *Dokl. Akad. Nauk SSSR* **56**, 7 (1947).

[11] G. García-Pérez, M. Boguñá, M. Á. Serrano, Multiscale unfolding of real networks by geometric renormalization. *Nature Physics* **14**, 583 (2018).

[12] K. G. Wilson, The renormalization group: Critical phenomena and the kondo problem. *Reviews of modern physics* **47**, 773 (1975).

[13] K. G. Wilson, The renormalization group and critical phenomena. *Reviews of Modern Physics* **55**, 583 (1983).

[14] L. P. Kadanoff, *Statistical physics: statics, dynamics and renormalization* (World Scientific Publishing Company, 2000).

[15] C. Song, S. Havlin, H. A. Makse, Origins of fractality in the growth of complex networks. *Nature Physics* **2**, 275 (2006).

[16] M. Á. Serrano, D. Krioukov, M. Boguñá, Self-similarity of complex networks and hidden metric spaces. *Physical review letters* **100**, 078701 (2008).

[17] D. Krioukov, F. Papadopoulos, A. Vahdat, M. Boguñá, Curvature and temperature of complex networks. *Physical Review E* **80**, 035101 (2009).

[18] G. García-Pérez, M. Á. Serrano, M. Boguñá, Soft communities in similarity space. *Journal of Statistical Physics* **173**, 775 (2018).
[19] K. Zuev, M. Boguná, G. Bianconi, D. Krioukov, Emergence of soft communities from geometric preferential attachment. *Scientific reports* **5**, 9421 (2015).

[20] D. Hric, K. Kaski, M. Kivelä, Stochastic block model reveals maps of citation patterns and their evolution in time. *Journal of Informetrics* **12**, 757 (2018).

[21] S. Fortunato, *et al.*, Science of science. *Science* **359**, eaao0185 (2018).

[22] G. García-Pérez, M. Boguñá, A. Allard, M. Á. Serrano, The hidden hyperbolic geometry of international trade: World trade atlas 1870–2013. *Scientific reports* **6**, 33441 (2016).

[23] S. Milojević, Quantifying the cognitive extent of science. *Journal of Informetrics* **9**, 962 (2015).

[24] F. Radicchi, S. Fortunato, C. Castellano, Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences* **105**, 17268 (2008).

[25] D. Wang, C. Song, A. Barabási, Quantifying long-term scientific impact. *Science* **342**, 127 (2013).

[26] Correlates of War Project, State system membership list, v2016, [http://correlatesofwar.org](http://correlatesofwar.org) (2017).

[27] V. D. Blondel, J. L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* **2008**, P10008 (2008).

[28] N. X. Vinh, J. Epps, J. Bailey, Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *Journal of Machine Learning Research* **11**, 2837 (2010).

[29] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, M. Boguñá, Hyperbolic geometry of complex networks. *Physical Review E* **82**, 036106 (2010).

[30] M. Boguñá, F. Papadopoulos, D. Krioukov, Sustaining the internet with hyperbolic mapping. *Nature communications* **1**, 62 (2010).

[31] G. García-Pérez, A. Allard, M. Á. Serrano, M. Boguñá, Mercator: uncovering faithful hyperbolic embeddings of complex networks. *arXiv:1904.10814* (2019).

[32] S. Borak, W. Härdle, R. Weron, *Statistical tools for finance and insurance* (Springer, 2005), pp. 21–44.

[33] B. Gnedenko, A. Kolmogorov, *Limit distributions for sums of independent random variables* (Cambridge, Addison-Wesley, 1954).

[34] B. B. Mandelbrot, The variation of certain speculative prices. *The Journal of Business* **36**,
E. F. Fama, The behavior of stock-market prices. *The journal of Business* **38**, 34 (1965).

J. P. Nolan, *Stable Distributions-Models for Heavy Tailed Data* (Springer Nature, Boston, 2019).

M. Kateregga, S. Mataramvura, D. Taylor, Parameter estimation for stable distributions with application to commodity futures log-returns. *Cogent Economics & Finance* **5**, 1318813 (2017).

J. P. Nolan, Numerical calculation of stable densities and distribution functions. *Communications in statistics. Stochastic models* **13**, 759 (1997).

J. Royuela-del Val, F. Simmross-Wattenberg, C. Alberola-López, libstable: Fast, parallel, and high-precision computation of $\alpha$-stable distributions in r, c/c++, and matlab. *Journal of Statistical Software* **78** (2017).

M. Á. Serrano, D. Krioukov, M. Boguñá, Percolation in self-similar networks. *Physical review letters* **106**, 048701 (2011).

S. N. Dorogovtsev, A. V. Goltsev, J. F. Mendes, Critical phenomena in complex networks. *Reviews of Modern Physics* **80**, 1275 (2008).

M. Fire, R. Puzis, Organization mining using online social networks. *Networks and Spatial Economics* **16**, 545 (2016).

M. Kuperman, D. Zanette, Stochastic resonance in a model of opinion formation on small-world networks. *The European Physical Journal B-Condensed Matter and Complex Systems* **26**, 387 (2002).

L. Gammaitoni, P. Hänggi, P. Jung, F. Marchesoni, Stochastic resonance. *Reviews of modern physics* **70**, 223 (1998).

A. Pikovsky, A. Zaikin, M. A. de La Casa, System size resonance in coupled noisy systems and in the ising model. *Physical review letters* **88**, 050601 (2002).

R. Toral, C. J. Tessone, Finite size effects in the dynamics of opinion formation. *Communications in Computational Physics* **1**, 1 (2006).

P. Jung, P. Hänggi, Stochastic nonlinear dynamics modulated by external periodic forces. *EPL (Europhysics Letters)* **8**, 505 (1989).

M. Newman, R. Ziff, Efficient monte carlo algorithm and high-precision results for percolation. *Physical Review Letters* **85**, 4104 (2000).

M. Á. Serrano, M. Boguñá, A. Vespignani, Extracting the multiscale backbone of complex
weighted networks. *Proceedings of the national academy of sciences* **106**, 6483 (2009).

[50] K. Claffy, Y. Hyun, K. Keys, M. Fomenkov, D. Krioukov, *2009 Cybersecurity Applications & Technology Conference for Homeland Security* (IEEE, 2009), pp. 205–211.

[51] M. Á. Serrano, M. Boguñá, F. Sagués, Uncovering the hidden geometry behind metabolic networks. *Molecular biosystems* **8**, 843 (2012).

[52] Openflights network dataset, [http://konect.uni-koblenz.de/networks/openflights](http://konect.uni-koblenz.de/networks/openflights) (2016).

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

[54] S.-Y. Takemura, *et al.*, A visual motion detection circuit suggested by drosophila connectomics. *Nature* **500**, 175 (2013).

[55] B. Klimt, Y. Yang, Introducing the enron corpus. In *CEAS’04: Proceedings of the 1st Conference on Email and Anti-Spam* (2004).

[56] J. Leskovec, K. J. Lang, A. Dasgupta, M. W. Mahoney, Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics* **6**, 29 (2009).

[57] J. Serrà, Á. Corral, M. Boguñá, M. Haro, J. L. Arcos, Measuring the evolution of contemporary western popular music. *Scientific reports* **2**, 521 (2012).

[58] T. Rolland, *et al.*, A proteome-scale map of the human interactome network. *Cell* **159**, 1212 (2014).

[59] R. Milo, *et al.*, Superfamilies of evolved and designed networks. *Science* **303**, 1538 (2004).

[60] C. J. Tessone, R. Toral, System size stochastic resonance in a model for opinion formation. *Physica A: Statistical Mechanics and its Applications* **351**, 106 (2005).

[61] P. Lévy, Calcul des probabilités (gauthier-villars) (1925).

**ACKNOWLEDGMENTS**

**Funding:** We acknowledge support from: a James S. McDonnell Foundation Scholar Award in Complex Systems; the ICREA Academia award, funded by the *Generalitat de Catalunya*; the Spanish *Ministerio de Ciencia, Innovación y Universidades* project no. FIS2016-76830-C2-2-P (AEI/FEDER, UE); project *Mapping Big Data Systems: embed-
ding large complex networks in low-dimensional hidden metric spaces, Ayudas Fundación BBVA a Equipos de Investigación Científica 2017, and Generalitat de Catalunya grant No. 2017SGR1064. Furthermore, G. G.-P. acknowledges financial support from the Academy of Finland via the Centre of Excellence program (Project no. 312058 as well as Project no. 287750), and from the emmy.network foundation under the aegis of the Fondation de Luxembourg.

**Authors contributions:** All the authors participated in the design of the research, the implementation of the analysis and the writing of the paper.

**Competing interests:** The authors declare no competing interests.

**Data and materials availability:** All the data are available in the manuscript and the supplementary materials, or they will be provided on request.

**SUPPLEMENTARY MATERIALS**

Materials and Methods
Supplementary Text
Figs. S1 to S24
Tables S1 to S3
References (1-61)
Supplementary Materials for

Geometric origins of self-similarity in the evolution of real networks

Muhua Zheng$^{1,2}$, Guillermo García-Pérez$^{3,4}$, Marián Boguñá$^{1,2}$ & M. Ángeles Serrano$^{1,2,5}$

1. Departament de Física de la Matèria Condensada,
   Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain
2. Universitat de Barcelona Institute of Complex Systems (UBICS),
   Universitat de Barcelona, Barcelona, Spain
3. QTF Centre of Excellence, Turku Centre for Quantum Physics,
   Department of Physics and Astronomy,
   University of Turku, FI-20014 Turun Yliopisto, Finland
4. Complex Systems Research Group,
   Department of Mathematics and Statistics,
   University of Turku, FI-20014 Turun Yliopisto, Finland
5. ICREA, Pg. Lluís Companys 23, E-08010 Barcelona, Spain

*Correspondence and requests for materials should be addressed to M.A.S.
(marian.serrano@ub.edu)

This PDF file includes:

• Materials and Methods
• Supplementary Text
• Figures S1 to S24
• Tables S1 to S3
• References
## CONTENTS

Introduction  2

Self-similar evolution of real networks  4

The geometric branching growth model  6

Scaled-up real network replicas  14

Conclusions  17

References  18

Acknowledgments  22

Supplementary materials  23

I. Supplementary Materials and Methods  4
   A. Data description  4
   B. Network embedding to produce geometric network maps  6
   C. Adjustment of stable distributions and generation of $z^\pm$  7
   D. Connecting descendants of connected ancestors  8
   E. Estimation of parameter $a$ in real networks  10
   F. Upscaled real network replicas  11
   G. Simulation of the MEIN model  11

II. Supplementary Text  12
   A. The $S^1/H^2$ models  12
   B. Geometric renormalization  13
   C. Stable distributions  15
   D. Downscaled network replicas  16

III. Supplementary Figures S1 to S25  17
   A. Self-similar evolution of real networks  17
   B. Stable distributions in real networks  21
C. Up and down in real networks

D. Predicting the evolution of the JCN and the WTW

IV. Supplementary Tables S1 to S3

References
I. SUPPLEMENTARY MATERIALS AND METHODS

A. Data description

Journal Citation Network (JCN). The citation networks from 1900 to 2013 were reconstructed from data on citations between scientific articles extracted from the Thomson Reuters Citation Index [20]. Years were grouped in time windows of ten years before 1970 and of five years afterwards. A node corresponds to a journal with publications in the given time period. A directed edge is connected from journal \( i \) to journal \( j \) if an article in journal \( i \) cites an article in journal \( j \), and the weight of this link is taken to be the number of such citations. Time differences between the citing and the cited articles are shorter than the length of the corresponding time window.

In this work, we use undirected and unweighted versions of the original networks. We first discard the directions for any link and preserve the weight \( \omega_{ij} \) with the sum of the citations, i.e., \( \omega_{ij} = \omega_{i \rightarrow j} + \omega_{j \rightarrow i} \). The resulting networks are weighted and undirected but very dense. Next, we extract the backbone by preserving the edges that represent statistically significant deviations with respect to a null model for the local assignment of weights to edges [49]. Eventually, we obtain the giant connected component of the undirected and unweighted networks from the backbone.

World Trade Atlas. We used networks in the World Trade Atlas [22], a collection of annual world trade network maps in hyperbolic space that provides information for the long-term evolution of the international trade system from 1870 to 2013 (world war periods, 1914-1919 and 1939-1947, were not available due to the lack of reported information). The networks were reconstructed using historical aggregate import/export data. In every network, nodes are countries, undirected links represent bilateral trade relationships, and link weights correspond to the value of goods exchanged in a given year in current US millions of dollars. The original networks were extremely dense, and in this work we consider the unweighted versions of the disparity backbones [22, 49].

Internet. The AS Internet topology in June 2009 was extracted from data collected by the archipelago active measurement infrastructure developed by Cooperative Association for Internet Data Analysis [50]. The network contains 23748 ASs represented as nodes, and it has average AS degree \( \langle k \rangle = 4.92 \), and average clustering (measured over ASs of degree
larger than 1) \( \bar{c} = 0.61 \). The exponent of the power-law degree distribution is \( \gamma \approx 2.1 \).

**Metabolic.** This network is the one-mode projection of metabolites of the bipartite human metabolic network at the cell level [51]. The network has \( N = 1436 \) nodes, average degree \( \langle k \rangle = 6.57 \), and average clustering \( \bar{c} = 0.54 \). The exponent of the power-law degree distribution is \( \gamma \approx 2.6 \).

**Facebook.** The Facebook dataset [42] contains 320 nodes and 2369 links, where nodes correspond to Facebook users who stated that they worked for a specific software development company in their Facebook profiles, and links represent the friendship relationships. The network has \( N = 320 \) nodes, average degree \( \langle k \rangle = 14.81 \), and average clustering \( \bar{c} = 0.49 \). The exponent of the power-law degree distribution is \( \gamma \approx 3.8 \).

We also analyzed the following networks (giant connected component of the undirected and unweighted versions):

- **Airports.** The airports dataset is obtained from Ref. [52, 53]. Directed links represent flights by airlines. A undirected version is used by keeping bidirectional edges only.

- **Drosophila.** The dataset is from Ref. [54]. It is a functional connectome within the Drosophila melanogaster optic medulla related to the motion detection circuit.

- **Enron.** The network of email communication within the Enron company is obtained from Refs. [55, 56].

- **Music.** The dataset is obtained from Ref. [11, 57], where nodes are chords-sets of musical notes played in a single beat-and connections represent observed transitions among them in a set of songs. Instead of the weighted, directed and dense network of original dataset, we considered the undirected, unweighted and sparser version by applied the disparity filter [49].

- **Proteome.** This proteome network is obtained from the human HI-II-14 interactome in Ref. [58] and is removed self-loops in this work.

- **Words.** The Words network is the network of adjacency between words in Darwin’s book “The Origin of Species”, from Ref. [59].

In Tables S1 and S2 we give year by year statistics for the JCN and the WTW, respectively. Table S3 summarizes the main topological features and fitting parameters of stable
distribution for the rest of networks.

B. Network embedding to produce geometric network maps

We embed each considered network (except WTW networks) into hyperbolic space using the algorithm introduced in Ref. [31], named Mercator. Mercator takes the network adjacency matrix $A_{ij}$ ($A_{ij} = A_{ji} = 1$ if there is a link between nodes $i$ and $j$, and $A_{ij} = A_{ji} = 0$ otherwise) as input and then returns inferred hidden degrees, angular positions of nodes and global model parameters. More precisely, the hyperbolic maps were inferred by finding the hidden degree and angular position of each node, $\{\kappa_i\}$ and $\{\theta_i\}$, that maximize the likelihood $L$ that the structure of the network was generated by the $S^1$ model, where

$$L = \prod_{i<j} [p_{ij}^{A_{ij}} [1 - p_{ij}]^{1 - A_{ij}}, \quad (5)$$

and $p_{ij}$ is the connected probability in $S^1$ model

$$p_{ij} = \frac{1}{1 + \chi_{ij}^{\beta}} = \frac{1}{1 + \left(\frac{d_{ij}}{\mu \kappa_i \kappa_j}\right)^\beta} \quad (6)$$

For the WTW, we used the coordinates in the hyperbolic maps from our previous work [22], in which we used a Metropolis-Hastings algorithm to embed the sequence of networks starting from the most recent one. The embedding of each network was based on the $H^2$ model, taking as initial coordinates the ones obtained in the embedding of the posterior layer. To give a more accurate estimation of parameters $\mu$ and $\beta$, we have adjusted them so that the ensemble of synthetic networks generated by Eq. (6) with the set of coordinates $\{\kappa_i, \theta_i\}$ in [22] have the same average degree and clustering coefficient as the empirical network on average. The detailed procedure to adjusted $\mu$ and $\beta$ is as follows:

1. Take the coordinates $\{\kappa_i, \theta_i\}$ and parameters $\mu$ and $\beta$ from the embeddings in [22].

2. To obtain a synthetic network with the average degree $\langle k \rangle_{real}$ of a real network, we correct the value of $\mu$ as $\mu_{new} = \xi \langle k \rangle_{real} \langle k \rangle_{syn} \mu$, where the initial value of parameter $\mu$ is taken from the embedding in [22], and we generate a synthetic network by connecting pairs of nodes using Eq. (6). If $\langle k \rangle_{syn} > \langle k \rangle_{real}$, we set $\xi - 0.1u \rightarrow \xi$, where $u$ is a random variable uniformly distributed between (0,1). Similarly, if $\langle k \rangle_{syn} < \langle k \rangle_{real}$, $\xi + 0.1u \rightarrow \xi$. The process is iterated and ends when $|\langle k \rangle_{syn} - \langle k \rangle_{real}| < 0.1$. 

6
3. To obtain a synthetic network with the average clustering coefficient \( \langle c \rangle_{\text{real}} \) of the real network, we set \( \beta_{\text{max}} = 3 \) and \( \beta_{\text{min}} = 1 \) and \( \beta_{\text{new}} = (\beta_{\text{max}} + \beta_{\text{min}})/2 \), where the initial value of parameter \( \beta \) is taken from the embedding in [22], and we generate a synthetic network connecting pairs of nodes using Eq. (6). If \( \langle c \rangle_{\text{syn}} > \langle c \rangle_{\text{real}} \), we set \( \beta_{\text{max}} = \beta_{\text{new}} \), else \( \beta_{\text{min}} = \beta_{\text{new}} \). The process is iterated and ends when \( |\langle c \rangle_{\text{syn}} - \langle c \rangle_{\text{real}}| < 0.01 \).

We generated 1000 synthetic networks for each WTW snapshot and obtained the final parameters \( \mu \) and \( \beta \) as their average values over the generated ensembles.

C. Adjustment of stable distributions and generation of \( z^{\pm} \)

The hidden degree \( z^{+} \) of the descendant of an ancestor with hidden degree \( z \) is obtained numerically by solving

\[
\int_{z_{\text{cut}}^{+}}^{z^{+}} \rho(z'|z)_{\text{nor}} \, dz' = u,
\]

where \( u \) is a random variable uniformly distributed between 0 and 1, and

\[
\rho(z'|z)_{\text{nor}} = C(z) \rho(z'|z) = C(z) \frac{f(z^{+}; \alpha^{+}, \eta^{+}, c^{+}, d^{+}) f(z-z^{+}; \alpha^{+}, \eta^{+}, c^{+}, d^{+})}{f(z; \alpha, \eta, c, d)}.
\]

The stable distribution for descendants \( f(z^{\pm}; \alpha^{\pm}, \eta^{\pm}, c^{\pm}, d^{\pm}) \) follows immediately from one of the basic properties of stable distributions in the parametrization used in this work [39] (see subsection C in “Supplementary Text”). If \( z_1 \sim f(z_1; \alpha_1, \eta_1, c_1, d_1) \) and \( z_2 \sim f(z_2; \alpha_2, \eta_2, c_2, d_2) \) are independent variables, then \( z_1 + z_2 \sim f(z; \alpha, \eta, c, d) \) with

\[
\alpha_1 = \alpha_2 = \alpha, \quad \eta = \frac{\eta_1 c_1^\alpha + \eta_2 c_2^\alpha}{c_1^\alpha + c_2^\alpha}, \quad c^{\alpha} = c_1^\alpha + c_2^\alpha, \quad d = d_1 + d_2;
\]

and then \( f(z^{\pm}; \alpha^{\pm}, \eta^{\pm}, c^{\pm}, d^{\pm}) = f(z^{\pm}; \alpha, \eta, c/2^{1/\alpha}, d/2) \).

The purpose of factor \( C(z) \) in Eq. (8) is two-fold. On the one hand, it acts as a normalization that ensures that the hidden degrees of descendants are non-negative, given that, in general, the stable distribution can have negative support. On the other hand, as we are considering real-world networks, finite-size effects play an important role. To control for these effects, we introduce the normalization factor \( C(z) \) in Eq. (8) defined as

\[
C(z) = \left[ \int_{z_{\text{cut}}^{+}}^{z^{+}} \frac{f(z^{+}; \alpha^{+}, \eta^{+}, c^{+}, d^{+}) f(z-z^{+}; \alpha^{+}, \eta^{+}, c^{+}, d^{+})}{f(z; \alpha, \eta, c, d)} \right]^{-1}.
\]
To ensure that $z^\pm$ are non-negative, we impose a minimum hidden degree cut-off to descendants that is related to the minimum hidden degree $z_0$ in the distribution of ancestors, $z_{\text{cut}}^\pm = z_0 / 2^{1/\alpha}$.

Let us consider the following aspects:

I. Locally, and therefore also globally, $z^+$ and $z^-$ follow the same distribution, i.e., $f(z^+; \alpha, \eta, c^+, d^+) = f(z^-; \alpha, \eta, c^-, d^-)$. This can be shown straightforwardly, since

$$
\rho(z^-|z) = \int dz^+ \rho(z^+|z) \delta(z^+ - (z - z^-)) = \frac{f(z - z^-; \alpha, \eta, c^+, d^+) f(z^-; \alpha, \eta, c^+, d^+)}{f(z; \alpha, \eta, c, d)}.
$$

II. Globally, both variables $z^\pm$ are distributed as $f(z^\pm; \alpha, \eta, c^\pm, d^\pm)$. To see this, we compute the resulting distribution as

$$
\rho(z^\pm) = \int dz^\pm \rho(z^\pm|z) \rho(z) \\
= \int dz^\pm \frac{f(z^\pm; \alpha, \eta, c^\pm, d^\pm) f(z - z^\pm; \alpha, \eta, c^\pm, d^\pm)}{f(z; \alpha, \eta, c, d)} f(z; \alpha, \eta, c, d) \\
= f(z^\pm; \alpha, \eta, c^\pm, d^\pm) \int dz^\pm f(z - z^\pm; \alpha, \eta, c^\pm, d^\pm) \\
= f(z^\pm; \alpha, \eta, c^\pm, d^\pm) \int_0^\infty dz f(z; \alpha, \eta, c, d) \\
= f(z^\pm; \alpha, \eta, c^\pm, d^\pm).
$$

Eq. (11) proves that the distribution for descendants, $f(z^\pm; \alpha^\pm, \eta^\pm, c^\pm, d^\pm)$, has the basic properties of the stable distribution, with the same shape parameters and adjusted scale and location as the ancestor layer. In Fig. 2 in the main text and in Fig. S5, we show that the shape of the distribution of hidden variables $z^\pm$ for the descendants is invariant as the ancestor layer in different empirical networks, which confirms the results in Eq. (11) very well.

D. Connecting descendants of connected ancestors

In the non-inflationary model, after the branching process, we establish links between descendants of connected ancestors with probability $p_{ij}(\mu')$, with $\mu' = b\mu$, but fulfilling the
condition that there is at least one link between them. Given a pair on connected ancestors, we characterize the state of links between descendants with different ancestor using binary variables \( \sigma_e, e \in \{1, \ldots, E\} \), with \( \sigma_e = 1 \) meaning that the link exists and \( \sigma_e = 0 \) that the link is absent and variable \( E \) giving the number of all possible links. We have changed from the \( ij \) notation to a one-index notation for the sake of simplicity. If we assume that nodes either remain or split in at most two descendants, \( E \) can only take the values 4, 2, or 1. The number of possible configurations combining absence and presence of links between descendants is \( 2^E - 1 \), and the possible states are encoded in vector \( V = [\sigma_1, \sigma_2, \ldots, \sigma_E] \).

Any possible configuration occurs with unbiased probability

\[
P(V) = \frac{\prod_{e=1}^{E} p_e(\mu')^{\sigma_e} [1 - p_e(\mu')]^{1-\sigma_e}}{1 - q},
\]

where \( q = \prod_{e=1}^{E} (1 - p_e(\mu')) \) is the probability that the descendants of connected ancestors are not connected.

We use the following simple algorithm to generate one of the possible configurations of connections between descendants satisfying the previous equation but ensuring that at least one connection is formed:

1. Compute probability \( q = \prod_{e=1}^{E} (1 - p_e(\mu')) \).

2. For each possible link \( e = 1, \ldots, E \):
   - If at least one of the \( E \) potential links has been already formed, connect the current link \( e \) with probability \( p_e(\mu') \).
   - Else, connect the current link \( e \) with probability \( p_e(\mu')/(1 - q) \) and then update \( q \) as \( q \rightarrow q/(1 - p_e(\mu')) \).

The advantage of this algorithm is that it allows the random generation of any possible configuration by assigning links sequentially only once and without discarding any history.

For the sake of completeness, let us show that the above algorithm indeed generates any sequence of links with the unbiased probability given by Eq. (12). To do so, consider a general configuration in which the \( i \)-th link is the first to be assigned, that is, \( \sigma_n = 0 \) for \( n < i \) and \( \sigma_i = 1 \). Note that, at any time \( n \leq i \), the value of \( q \) (to which we will now refer
as \( q_n \) is

\[
q_n = \prod_{e=n}^{E} (1 - p_e(\mu')).
\]  

(13)

Therefore, the probability for link \( n < i \) not to be assigned is

\[
P(\sigma_n = 0) = 1 - \frac{p_n}{1 - q_n} = \frac{1 - p_n(\mu') - \prod_{e=n}^{E} (1 - p_e(\mu'))}{1 - \prod_{e=n}^{E} (1 - p_e(\mu'))}
\]

\[
= \frac{(1 - p_n(\mu')) \left( 1 - \prod_{e=n+1}^{E} (1 - p_e(\mu')) \right)}{1 - \prod_{e=n}^{E} (1 - p_e(\mu'))}
\]

\[
= \frac{(1 - p_n)(1 - q_{n+1})}{1 - q_n}.
\]  

(14)

Now, let us take a look at the resulting probability with which the whole configuration is generated if we follow the algorithm:

\[
P(V) = \left( \prod_{e=1}^{i-1} \frac{(1 - p_n(\mu')) (1 - q_{n+1})}{1 - q_n} \right) \frac{p_i}{1 - q_i} \prod_{e=i+1}^{E} p_e(\mu')^{\sigma_e} (1 - p_e(\mu'))^{1 - \sigma_e}
\]

\[
= \frac{1}{1 - q_i} \left( \prod_{e=1}^{i-1} (1 - p_e(\mu')) \right) p_i \prod_{e=i+1}^{E} p_e(\mu')^{\sigma_e} (1 - p_e(\mu'))^{1 - \sigma_e}
\]

\[
= \frac{1}{1 - \prod_{e=1}^{E} (1 - p_e(\mu'))} \left( \prod_{e=1}^{i-1} (1 - p_e(\mu')) \right) p_i \prod_{e=i+1}^{E} p_e(\mu')^{\sigma_e} (1 - p_e(\mu'))^{1 - \sigma_e},
\]  

(15)

which is Eq. (12) in this particular case. Notice that, if \( i = E \), \( q_i = 1 - p_i \) and \( P(\sigma_i = 1) = 1 \), so there is always one link at least.

E. **Estimation of parameter \( a \) in real networks**

Parameter \( a \) is used to adjust the growth of the average degree as the network evolves.

To estimate its value numerically, we perform the following steps

1. Based on Eq. (4) in main text, the slope \( s \) of \( \langle k \rangle \) as a function of \( N \) in log-log scale is

\[
s = \frac{\ln b - \nu + \ln a}{\ln b} = \frac{\ln a \varphi}{\ln b}.
\]  

Parameter \( a \) can be calculated for a certain branching ratio \( b \) once \( s \) and \( \nu \), or alternatively \( \varphi \), are estimated.
(2) We estimate $s$ by least squares fitting the empirical data in a given time period, see Fig. 4A and B in main text.

(3) We take a network snapshot and apply to it one step of non-inflationary GBG ($a = 1$), i.e., without adding extra links. Then, we measure the average degree $\langle k\rangle_{a=1}^{(1)}$ and $\langle k\rangle^{(0)}$ in the descendant and original layers and calculate $\varphi = b^{-\nu} = \langle k\rangle_{a=1}^{(1)}/\langle k\rangle^{(0)}$ using Eq. (4) in main text.

(4) Go to (1) and calculate the value of $a$.

We repeat the process for different time snapshots to get a set of values of $a$. In Fig. S15, we show the values of $a$ in the JCN and WTW. In both systems, the variances of $a$ are very small.

F. Upscaled real network replicas

Depending on the desired final size, $b$ and the number of layers $l$ are fixed to some specific values. To adjust the average degree, we set $a = \xi \frac{\langle k^{(0)} \rangle}{\langle k^{(l)} \rangle}$, where $\langle k^{(0)} \rangle$ is the target average degree and $\langle k^{(l)} \rangle$ is the obtained average degree in non-inflationary GBG layer $l$. We start an iterative process with $\xi = 1$ and $\langle k^{(l)} \rangle$ the initial average degree of layer $l$. In each round, we add links to the network using Eq. (3) in the main text and calculate the average degree $\langle k^{(l)}_{\text{new}} \rangle$ of the resulting network. If $\langle k^{(l)}_{\text{new}} \rangle > \langle k^{(0)} \rangle$, we discard the whole realization and start the process again giving $\xi$ a new value $\xi - 0.1u$, where $u$ is a random variable uniformly distributed in the interval $(0, 1)$. Similarly, if $\langle k^{(l)}_{\text{new}} \rangle < \langle k^{(0)} \rangle$, $\xi$ is updated as $\xi + 0.1u$. The process is repeated until $|\langle k^{(l)}_{\text{new}} \rangle - \langle k^{(0)} \rangle|$ is below a given threshold, that we set to 0.1.

G. Simulation of the MEIN model

Stochastic resonance was found in a very simple model for opinion formation in social systems [43, 46, 60]. In the model, the opinion of an individual can change due to three basic ingredients: (i) the influence of connected neighbors, modeled by a simple majority rule; (ii) the influence of fashion, modeled as some external time varying signal and (iii) random events.
We implement the three effects mentioned above as follows. At time $t = 0$, we assign random values $m_i = \pm 1$ to each individual; then, at a given time $t$, the next three steps are applied consecutively:

i. One individual $i$ is randomly selected and it adopts the majority opinion among its connected neighbors, i.e., $m_i(t) = \text{sign} [\sum_{j \in n(i)} m_j(t)]$. If $\sum_{j \in n(i)} m_j(t) = 0$, in case of a tie, a random value for $m_i(t)$ is selected.

ii. With probability $A|\cos(2\pi t/T)|$, set $m_i(t) = \text{sign}[\cos(2\pi t/T)]$ to follow the fashion. Parameter $A(0 \leq A \leq 1)$ measures the strength of the fashion and $T$ its period.

iii. With probability $\epsilon$, let $m_i(t)$ adopt randomly a new value (independently of its present state), where $\epsilon$ represents the noise intensity in the dynamics.

After the three steps have been performed, time increases by a fixed amount $t = t + 1/N$, so that after one unit of time every individual has been updated once on average.

To measure of the response of the system to external modulation, we compute the spectral amplification factor [47]:

$$R = 4A^{-2}|\langle e^{i2\pi t/T} \rho(t) \rangle|,$$

(16)

where $\rho(t) = \frac{1}{N} \sum_i m_i(t)$ is the average opinion in the network at time $t$, and $\langle \cdots \rangle$ denotes a time average.

II. SUPPLEMENTARY TEXT

A. The $S^1/H^2$ models

The $S^1$ model is the simplest geometric model able to match real world topologies [16]. Each node $i$ has a pair of hidden variables $(\kappa_i, \theta_i)$ representing the node’s expected degree (as a proxy for its popularity) and its angular position (as an abstraction of a similarity space) in a one-dimensional sphere (or circle). $N$ nodes are distributed with a fixed density, set to one without loss of generality, so that the radius of the circle is $R = N/2\pi$. The connection probability between any pair of nodes takes the form of a gravity law, whose magnitude increases with the product of the hidden degrees (i.e., their combined popularities), and decreases with the angular distance between the two nodes. In other words, similar nodes
are angularly closer and are therefore more likely to be connected. Specifically, nodes $i$ and $j$ are connected with probability

$$p_{ij} = \frac{1}{1 + \chi_{ij}^\beta} = \frac{1}{1 + \left( \frac{d_{ij}}{\mu \kappa_i \kappa_j} \right)^\beta}, \quad (17)$$

where $d_{ij} = R \Delta \theta_{ij}$ is the arc length of the circle between nodes $i$ and $j$ separated by an angular distance $\Delta \theta_{ij} = \pi - |\pi - |\theta_i - \theta_j||$. Parameters $\mu$ and $\beta$ control the average degree of the network and the level of clustering, respectively. Apart from minor corrections due to the finite size of the network, the hidden degrees $\kappa_i$ and $\kappa_j$ are proportional to the degrees of nodes $i$ and $j$, respectively.

The $S^1$ model is equivalent to a purely geometric version—the $H^2$ model [17]. By transforming the expected node degrees $\kappa_i$ into radial coordinates $r_i$ via $r_i = R_{H^2} - 2 \ln \frac{\kappa_i}{\kappa_0}$, where $R_{H^2} = 2 \ln \frac{N \pi \mu \kappa_0^2}{\pi}$ is the radius of the finite hyperbolic disk containing the nodes and $\kappa_0 = \min(\{\kappa_i\})$ is the smallest hidden degree in the system. Higher degree nodes are therefore located closer to the center of the hyperbolic disk. The connection probability becomes $p_{ij} = 1/(1 + e^{\frac{\beta}{2}(x_{ij} - R_{H^2})})$, where $x_{ij} = r_i + r_j + 2 \ln \frac{\Delta \theta_{ij}}{2}$ is a very good approximation of the hyperbolic distance between two points in the hyperbolic disk.

### B. Geometric renormalization

The geometric renormalization (GR) technique for real networks was first introduced in Ref. [11]. The approach relies on geometric network maps to coarse-grain neighbouring nodes into supernodes and rescale connections. The renormalization transformation zooms out by changing the minimum length scale from that of the original network to a larger value. We apply the GR process as follows:

1. First, given an original network topology corresponding to $l = 0$, the hidden degrees $\kappa_{i}^{(0)}$ and angular coordinates $\theta_{i}^{(0)}$ are obtained by embedding the nodes in the hyperbolic space the algorithm Mercator [31].

2. Second, we define non-overlapping blocks of consecutive nodes of size $r$ along the similarity circle and apply the coarse-graining by merging the nodes in the blocks to form supernodes with probability $p$. As a consequence, the population in layer $l + 1$
decreases as

\[ N^{(l+1)} = N^{(l)} \frac{p + r(1 - p)}{r} \]  

(18)

In accordance with the GBG technique, we have \( N^{(l)} / N^{(l+1)} = b \) (notice that here the superindex \( l \) increases in the GR flow and decreases in the GBG flow), and then

\[ p = \frac{(b - 1)r}{(r - 1)b} \]  

(19)

In particular, when \( p = 1 \), \( b = r \) and \( N^{(l+1)} = N^{(l)}/r \) is the standard geometric renormalization in Ref. [11].

Once a supernode has been created, it is placed within the angular region defined by the corresponding block so that the original ordering of nodes is preserved. The renormalized networks remain maximally congruent with the hidden metric space model by assigning a hidden degree \( \kappa_i^{(l+1)} \) to supernode \( i \) in layer \( l + 1 \) given by

\[ \kappa_i^{(l+1)} = \left[ \sum_{j=1}^{r} (\kappa_j^{(l)})^\beta \right]^{1/\beta}, \]  

(20)

as well as an angular coordinate \( \theta_i^{(l+1)} \) given by

\[ \theta_i^{(l+1)} = \left[ \frac{\sum_{j=1}^{r} (\theta_j^{(l)})^\beta \sum_{j=1}^{r} (\kappa_j^{(l)})^\beta}{\sum_{j=1}^{r} (\kappa_j^{(l)})^\beta} \right]^{1/\beta} \]  

(21)

Global parameters are rescaled as \( \mu^{(l+1)} = \mu^{(l)}/b \), \( \beta^{(l+1)} = \beta^{(l)} \), and \( R^{(l+1)} = R^{(l)}/b \). This implies that the probability for two supernodes \( i \) and \( j \) to be connected in layer \( l + 1 \) maintains its original form independently of the layer.

3. Third, we connect two supernodes \( i \) and \( j \) in layer \( l + 1 \) as long as some node in \( i \) is connected to some node in \( j \) in layer \( l \).

This procedure is iterated starting from the original network at layer \( l = 0 \). Finally, the set of renormalized network layers \( l \) forms a multiscale shell of the network.
C. Stable distributions

Stable distributions are a rich class of probability distributions that allow skewness and heavy tails and have many intriguing mathematical properties. In probability theory, a distribution is said to be stable if a linear combination of two independent random variables with this distribution has the same distribution, up to location and scale parameters \[36, 61\]. As mentioned in the main text, a stable distribution requires four parameters \(\{\alpha, \eta, c, d\}\) to be fully characterized. Apart from the four parameters, there are multiple parameterizations for stable laws depending on the research purpose \[36, 38, 39\]. This variety of parameterizations is caused by a combination of historical evolution, plus the numerous problems that have been analyzed using specialized forms of stable distributions. The probability density function for a general stable distribution does not have an analytic form, but its characteristic function \(\Phi(t)\) (its Fourier transform) does. In the most common parametrization, the one that we use in this work with parametrization subindex 1 as in \[39\], \(\Phi(t) = \exp[\Psi(t)]\) where

\[
\Psi(t) = \begin{cases} 
-|ct|^\alpha[1 - i\eta \tan(\frac{\pi \alpha}{2}) \text{sign}(t)] + idt, & \alpha \neq 1 \\
-|ct|[1 + i\eta \frac{2}{\pi} \text{sign}(t) \ln(|t|)] + idt, & \alpha = 1 
\end{cases}
\]

\[\text{sign}(t) = \begin{cases} 
1, & t > 0 \\
0, & t = 0 \\
-1, & t < 0 
\end{cases}\]

and \(i\) denotes the imaginary part of a complex numbers.

To obtain parameters \(\{\alpha, \eta, c, d\}\) for the distribution of hidden degrees \(z\) in the original layer we use the software “libstable” in Ref. \[39\]. In particular, we use the modified maximum likelihood method, where the maximization search is only performed in the 2D \(\alpha-\eta\) space, such that the procedure produces more accurate estimates with much shorter execution times. In Fig. 2 in the main text and Fig. S5, we show that the stable distribution offers a very good fit to the hidden variables \(z\) in real networks.
D. Downscaled network replicas

Once a renormalized network layer is obtained using GR, we need to reduce the resulting average degree to that of the original network in order to obtain a downscaled network replica. We follow the technique presented in [11]. To obtain a pruned network congruent with the underlying metric space model, we only need to find the right value of $\mu_{\text{new}}^{(l)}$ so that $\langle k^{(l)}_{\text{new}} \rangle = \langle k^{(0)} \rangle$. Similar to the upscaled network replicas procedure, we correct the value $\mu_{\text{new}}^{(l)} = \xi \frac{\langle k^{(0)} \rangle}{\langle k^{(l)} \rangle} \mu^{(l)}$. Initially, we set $\xi$ to 1. For every value of $\xi$, we prune the network. In particular, if $\langle k^{(l)}_{\text{new}} \rangle > \langle k^{(0)} \rangle$, we set $\xi$ as $\xi - 0.1 u \rightarrow \xi$, where $u$ is a random variable uniformly distributed between $0, 1)$. Similarly, if $\langle k^{(l)}_{\text{new}} \rangle < \langle k^{(0)} \rangle$, $\xi + 0.1 u \rightarrow \xi$. The process stops when $|\langle k^{(l)}_{\text{new}} \rangle - \langle k^{(0)} \rangle|$ is below a given threshold, that we set to 0.1.
III. SUPPLEMENTARY FIGURES S1 TO S25

A. Self-similar evolution of real networks

![Graphs showing evolution of structural properties](image)

FIG. S1. Evolution of structural properties. (a) and (b), the number of nodes, (c) and (d), the average degree, (e) and (f), the average clustering coefficient in the journal citation network (left column) and in the world trade web (right column). World war periods, 1914-1919 and 1939-1947, were not available in the WTW due to lack of reported information.
FIG. S2. (Color online). **Self-similar evolution of the journal citation network.** In (a)-(c), we show the complementary cumulative degree distributions, clustering spectra and the degree-dependent average nearest neighbors degree distributions in all the time-windows from 1900-1910 to 2008-2013. In (d)-(f), we only include snapshots from 1950-1960 to 2008-2013. Fluctuations are smaller after World War II.
FIG. S3. (Color online). **Self-similar evolution of the world trade web.** In (a)-(c), we show the complementary cumulative degree distributions, clustering spectra and the degree-dependent average nearest neighbors degree distributions in all the networks from 1870 to 2013. In (d)-(f), we only include snapshots from 1950 to 2013. Data in the world war periods, 1914-1919 and 1939-1947, was not available due to lack of reported information.
FIG. S4. **Evolution of community structure.** Modularity $Q$ and adjusted mutual information AMI between the community partitions of two consecutive snapshots by considering the nodes that exist in both snapshots. The JCN in (a) and the WTW in (b).
B. Stable distributions in real networks

FIG. S5. Performance of the stable distribution in real networks under GBG. The left column shows the complementary cumulative distributions of $z$ in the original networks and the stable distribution fitting. Middle and right columns show the distribution $P_{c}(z^{\pm})$ of different descendant layers ($l = 0$ indicates the original network) with $b = 1.2$ and $b = 2.0$, respectively. Each row shows the results for a real network.
C. Up and down in real networks

FIG. S6. GR vs GBG on real networks with $b = 1.2$. Each column from left to right represents the mean clustering coefficient $\langle c \rangle$, average degree $\langle k \rangle$, network size $N$, modularity $Q$ and the connection probability on the GR and the GBG flows. Each row shows the results for a real network.
FIG. S7. **GR vs GBG on the real network with** $b = 1.2$. Each column from left to right shows the complementary cumulative degree distribution, degree dependent clustering coefficient and degree-degree correlations of rescaled degrees $k_{res}^{(l)} = k^{(l)}/(k^{(l)})$. Each row shows the results for a real network.
FIG. S8. **GR vs GBG on real networks with** \( b = 2 \). Each column from left to right represents the mean clustering coefficient \( \langle c \rangle \), average degree \( \langle k \rangle \), network size \( N \), modularity \( Q \) and the connection probability on the GR and the GBG flows. Each row shows the results for a real network.
FIG. S9. GR vs GBG on real networks with $b = 2$. Each column from left to right shows the complementary cumulative degree distribution, degree dependent clustering coefficient and degree-degree correlations of rescaled degrees $k_{res}^{(l)} = k^{(l)}/\langle k^{(l)} \rangle$. Each row shows the results for a real network.
FIG. S10. GBG vs GR on real networks with $b = 1.2$. Each column from left to right represents the mean clustering coefficient $\langle c \rangle$, average degree $\langle k \rangle$, network size $N$, modularity $Q$ and the connection probability on the GBG and GR flows. Each row shows the results for a real network.
FIG. S11. **GBG vs GR on real networks with $b = 1.2$.** Each column from left to right shows the complementary cumulative degree distribution, degree dependent clustering coefficient and degree-degree correlations of rescaled degrees $k_{\text{res}}^{(l)} = k^{(l)}/\langle k^{(l)} \rangle$. Each row shows the results for a real network.
FIG. S12. GBG vs GR on real networks with $b = 2$. Each column from left to right represents the mean clustering coefficient $\langle c \rangle$, average degree $\langle k \rangle$, network size $N$, modularity $Q$ and the connection probability on the GBG and GR flows. Each row shows the results for a real network.
FIG. S13. **GBG vs GR on real networks with** $b = 2$. Each column from left to right shows the complementary cumulative degree distribution, degree dependent clustering coefficient and degree-degree correlations of rescaled degrees $k^{(l)}_{res} = k^{(l)}/(k^{(l)})$. Each row shows the results for a real network.
D. Predicting the evolution of the JCN and the WTW

FIG. S14. **Scaling of the average degree with network size.** Empirical data is compared with simulation results and the theoretical prediction.

FIG. S15. **Parameter \( a \) estimated from empirical snapshots.** The slope \( s = \frac{\ln b - \ln c + \ln a}{\ln b} \) of \( \langle k \rangle \) as a function of \( N \) in log-log scale is \( s = 0.3344 \pm 0.0380 \) for the JCN from 1950-1960 to 1965-1975, and \( s = 0.4113 \pm 0.2339 \) for the WTW from 1950 to 1960. The error bars show one standard deviation around the expected values of \( s \).
FIG. S16. Local rich-club and nested self-similarity effects in the nested hierarchy of subgraphs produced by degree thresholding. Each column from left to right shows the complementary cumulative distribution $P_c^{(l)}(k_{\text{res}}^{(l)})$ of rescaled degrees $k_{\text{res}}^{(l)} = k^{(l)}/\langle k^{(l)} \rangle$, degree dependent clustering coefficient over rescaled-degree classes, degree-degree correlations measured by the normalized average nearest-neighbour degree $k_{\text{nn},n}(k_{\text{res}}^{(l)}) = k_{\text{nn}}(k_{\text{res}}^{(l)})/\langle k^{(l)} \rangle^2$, the modularity $Q$, and the ratio of the subgraph average degree $\langle k_T \rangle$ to the original average degree $\langle k \rangle$ as a function of the inverse relative subgraph size $N/N_T$. The subgraphs are obtained by removing nodes with degrees below threshold $k_T$ from the original network.
FIG. S17. Ratio of the subgraph average degree $\langle k_T \rangle$ to the original average degree $\langle k \rangle$ as a function of the inverse relative subgraph size $N/N_T$ for the different snapshots of the JCN and the WTW. The subgraphs are obtained by removing nodes with degrees below threshold $k_T$ from the original network. To reduce the effects of finite-size, the data are shown only for the subgraph average degree $\langle k_T \rangle$ smaller than 0.05$N_T$ in JCN and 0.1$N_T$ in WTW.
FIG. S18. Local rich-club effect in different growing network models. (a) Behavior of the average degree $\langle k \rangle$ as a function of network size $N$. In particular, for the Barabási-Albert model (BA) [6], we set $m = 20$, and for the Popularity Similarity Optimization model (PSO) [7] we use parameters $\gamma = 2.76$, $T = 0.4$, $m = 20$ so that the generated networks eventually reach the same average degree, clustering and degree distribution than a targeted network, i.e., the empirical snapshot 1994-1999 in JCN. Notice that the average degree does not change with network size in the standard BA and PSO models. To make the average degree increasing with network size, we modify the standard models by changing $m$ as $m = 1 + \text{Int}(t/t_m)$, where $t$ denotes the current network size in the node addition process so that $m$ only increases after remaining stable for periods of lengths $t_m$. We set $t_m = 300$ such that the average degree of the targeted network can be reached. We name the new versions BA modified and PSO modified. (b) Ratio of subgraphs average degree $\langle k_T \rangle$ to the original average degree $\langle k \rangle$ as a function of the inverse relative subgraph size $N/N_T$ in the final graphs reached by the different growing models. Even if the average degree in the BA modified and the PSO modified grows with the system size as shown in (a), the behavior of the relative average degree in the final graphs in (b) shows that the observed behavior in real networks cannot be reproduced.
The GBG model predicts the self-similar evolution of the JCN and the WTW with different $b$ and starting time points. Similar to Fig. 3 in the main paper, but we take $b = 1.2$ for the JCN and 1.15 for the WTW (the corresponding values of $a$ are 1.203 and 1.075, respectively). We grow the networks with the GBG transformation with starting time points in 1990-1995 for the JCN and in 1965 for the WTW. The topological properties of simulated and empirical networks are compared. The results show that the prediction of the GBG method is robust, and that the evolution of the two systems is reproduced (at a statistical level) even when using different values of $b$ and different starting time points.
FIG. S20. Comparing average clustering, average degree, and network size of the empirical JCN snapshots and of layers evolved by the GBG. Orange symbols represent empirical snapshots and blue pentagons correspond to simulated GBG networks with approximately the same size. The starting time point is 1965-1975, and $b = 1.5$.

FIG. S21. Comparing average clustering, average degree, and network size of the empirical WTW snapshots and of layers evolved by the GBG. Orange symbols represent empirical snapshots and blue pentagons correspond to simulated GBG networks with approximately the same size. The starting time point is 1960, and $b = 1.2$. 
FIG. S22. Comparing self-similarity in the hierarchy of nested subgraphs in empirical JCN snapshots and in layers evolved by the GBG. From left to right, each column shows the complementary cumulative degree distribution $P_{c}^{(l)}(k_{res})$ of rescaled degrees $k_{res}^{(l)} = k^{(l)}/(k^{(l)})$, the degree dependent clustering coefficient $c^{(l)}(k_{res})$ over rescaled degrees, and the degree-degree correlations measured by the normalized average nearest-neighbour degree $\bar{k}_{nn,n}(k_{res}) = \bar{k}_{nn}(k_{res})/(k^{(l)})/((k^{(l)})^{2})$. Each row shows results for a different time snapshot. The subgraphs are obtained by removing nodes with degree below a threshold $k_{T}$. 

JCN
FIG. S23. Comparing self-similarity in the hierarchy of nested subgraphs in empirical WTW snapshots and in layers evolved by the GBG. From left to right, each column shows the complementary cumulative degree distribution $P_c^{(l)}(k_{res}^{(l)})$ of rescaled degrees $k_{res}^{(l)} = k^{(l)}/\langle k^{(l)} \rangle$, the degree dependent clustering coefficient $c^{(l)}(k_{res}^{(l)})$ over rescaled degrees, and the degree-degree correlations measured by the normalized average nearest-neighbour degree $\bar{k}_{nn,n}(k_{res}^{(l)}) = \bar{k}_{nn}(k^{(l)})\langle k^{(l)} \rangle / \langle (k^{(l)})^2 \rangle$. Each row shows results for a different time snapshot. The subgraphs are obtained by removing nodes with degree below a threshold $k_T$. 
FIG. S24. Network properties of upscaled and downscaled network replicas of real networks. (a) and (b), the complementary cumulative degree distribution, (c) and (d), degree dependent clustering coefficient, (e) and (f), degree-degree correlations over rescaled degrees $k_{res}^{(l)} = k^{(l)}/\langle k^{(l)} \rangle$, and (g) and (h) average clustering coefficient $\langle c \rangle$. Results for upscaled Facebook network replicas are shown on the left column, and results for downscaled Internet replicas are on the right column. The green symbol in (g) and (h) indicates the value for the real network.
IV. SUPPLEMENTARY TABLES S1 TO S3

TABLE S1. Year by year statistics for the Journal Citation Network. Columns are: the time windows of each network (Years), the number of nodes ($N$), the average degree ($\langle k \rangle$), and the average local clustering coefficient ($\langle c \rangle$).

| No. | Years     | $N$ | $\langle k \rangle$ | $\langle c \rangle$ | No. | Years     | $N$ | $\langle k \rangle$ | $\langle c \rangle$ |
|-----|-----------|-----|---------------------|--------------------|-----|-----------|-----|---------------------|--------------------|
| 0   | 1900-1910 | 118 | 4.068               | 0.403              | 26  | 1926-1936 | 340 | 8.506               | 0.550              |
| 1   | 1901-1911 | 127 | 4.315               | 0.409              | 27  | 1927-1937 | 362 | 10.331              | 0.554              |
| 2   | 1902-1912 | 151 | 5.682               | 0.387              | 28  | 1928-1938 | 364 | 9.610               | 0.541              |
| 3   | 1903-1913 | 147 | 4.571               | 0.433              | 29  | 1929-1939 | 372 | 9.435               | 0.538              |
| 4   | 1904-1914 | 159 | 5.761               | 0.463              | 30  | 1930-1940 | 380 | 10.053              | 0.553              |
| 5   | 1905-1915 | 179 | 6.715               | 0.461              | 31  | 1931-1941 | 399 | 10.832              | 0.559              |
| 6   | 1906-1916 | 181 | 7.348               | 0.494              | 32  | 1932-1942 | 408 | 11.770              | 0.554              |
| 7   | 1907-1917 | 176 | 6.091               | 0.489              | 33  | 1933-1943 | 399 | 11.599              | 0.555              |
| 8   | 1908-1918 | 185 | 6.151               | 0.476              | 34  | 1934-1944 | 381 | 11.470              | 0.555              |
| 9   | 1909-1919 | 178 | 5.011               | 0.501              | 35  | 1935-1945 | 400 | 11.725              | 0.546              |
| 10  | 1910-1920 | 177 | 4.757               | 0.498              | 36  | 1936-1946 | 578 | 11.626              | 0.613              |
| 11  | 1911-1921 | 204 | 6.245               | 0.525              | 37  | 1937-1947 | 596 | 12.651              | 0.608              |
| 12  | 1912-1922 | 240 | 7.083               | 0.527              | 38  | 1938-1948 | 653 | 11.525              | 0.606              |
| 13  | 1913-1923 | 253 | 7.455               | 0.522              | 39  | 1939-1949 | 701 | 14.382              | 0.598              |
| 14  | 1914-1924 | 228 | 6.167               | 0.523              | 40  | 1940-1950 | 741 | 16.159              | 0.597              |
| 15  | 1915-1925 | 236 | 6.373               | 0.526              | 41  | 1941-1951 | 777 | 15.483              | 0.591              |
| 16  | 1916-1926 | 245 | 6.580               | 0.534              | 42  | 1942-1952 | 794 | 15.657              | 0.591              |
| 17  | 1917-1927 | 250 | 7.080               | 0.523              | 43  | 1943-1953 | 800 | 15.650              | 0.586              |
| 18  | 1918-1928 | 279 | 7.290               | 0.534              | 44  | 1944-1954 | 827 | 14.989              | 0.578              |
| 19  | 1919-1929 | 304 | 9.033               | 0.539              | 45  | 1945-1955 | 781 | 15.636              | 0.586              |
| 20  | 1920-1930 | 296 | 7.797               | 0.560              | 46  | 1946-1956 | 846 | 18.234              | 0.597              |
| 21  | 1921-1931 | 287 | 8.000               | 0.562              | 47  | 1947-1957 | 943 | 18.554              | 0.590              |
| 22  | 1922-1932 | 290 | 8.234               | 0.567              | 48  | 1948-1958 | 970 | 16.097              | 0.586              |
| 23  | 1923-1933 | 332 | 9.440               | 0.545              | 49  | 1949-1959 | 1053 | 17.673             | 0.585               |
| 24  | 1924-1934 | 334 | 9.246               | 0.538              | 50  | 1950-1960 | 1106 | 18.029             | 0.594               |
| 25  | 1925-1935 | 339 | 10.124              | 0.559              | 51  | 1951-1961 | 1139 | 18.586             | 0.593               |
| No. | Years          | N   | ⟨k⟩   | ⟨c⟩   | No. | Years          | N   | ⟨k⟩   | ⟨c⟩   |
|-----|----------------|-----|-------|-------|-----|----------------|-----|-------|-------|
| 52  | 1952-1962      | 1381| 17.586| 0.632 | 78  | 1983-1988      | 7360| 27.677| 0.487 |
| 53  | 1953-1963      | 1474| 19.228| 0.630 | 79  | 1984-1989      | 7334| 27.087| 0.489 |
| 54  | 1954-1964      | 1543| 20.139| 0.629 | 80  | 1985-1990      | 7379| 28.461| 0.484 |
| 55  | 1955-1965      | 1724| 21.831| 0.631 | 81  | 1986-1991      | 7260| 28.621| 0.486 |
| 56  | 1956-1966      | 1945| 19.788| 0.645 | 82  | 1987-1992      | 8016| 37.887| 0.468 |
| 57  | 1957-1967      | 2313| 23.739| 0.637 | 83  | 1988-1993      | 8945| 39.171| 0.478 |
| 58  | 1958-1968      | 2549| 26.846| 0.622 | 84  | 1989-1994      | 9816| 37.690| 0.505 |
| 59  | 1959-1969      | 2821| 26.747| 0.620 | 85  | 1990-1995      | 10788|37.057| 0.516 |
| 60  | 1960-1970      | 3061| 24.989| 0.625 | 86  | 1991-1996      | 11433|38.225| 0.521 |
| 61  | 1961-1971      | 3281| 25.367| 0.616 | 87  | 1992-1997      | 11831|40.454| 0.523 |
| 62  | 1962-1972      | 3468| 25.950| 0.615 | 88  | 1993-1998      | 12225|42.360| 0.520 |
| 63  | 1963-1973      | 3805| 27.710| 0.614 | 89  | 1994-1999      | 12874|39.697| 0.530 |
| 64  | 1964-1974      | 3925| 25.381| 0.615 | 90  | 1995-2000      | 14675|38.525| 0.542 |
| 65  | 1965-1975      | 4168| 26.844| 0.608 | 91  | 1996-2001      | 16114|38.276| 0.562 |
| 66  | 1966-1976      | 4657| 33.135| 0.589 | 92  | 1997-2002      | 17174|37.238| 0.582 |
| 67  | 1967-1977      | 4994| 33.486| 0.580 | 93  | 1998-2003      | 18106|37.874| 0.597 |
| 68  | 1968-1978      | 5634| 44.415| 0.540 | 94  | 1999-2004      | 18879|38.501| 0.607 |
| 69  | 1969-1979      | 6154| 45.717| 0.527 | 95  | 2000-2005      | 19249|40.333| 0.607 |
| 70  | 1970-1980      | 6727| 43.615| 0.528 | 96  | 2001-2006      | 19389|42.643| 0.607 |
| 71  | 1972-1981      | 7026| 42.295| 0.515 | 97  | 2002-2007      | 19750|44.656| 0.606 |
| 72  | 1974-1982      | 7071| 39.770| 0.503 | 98  | 2003-2008      | 20703|44.987| 0.609 |
| 73  | 1976-1983      | 6971| 29.752| 0.518 | 99  | 2004-2009      | 21744|48.983| 0.600 |
| 74  | 1978-1984      | 7060| 27.995| 0.505 |100  | 2005-2010      | 22151|52.850| 0.594 |
| 75  | 1980-1985      | 7026| 25.625| 0.491 |101  | 2006-2011      | 22627|56.650| 0.586 |
| 76  | 1981-1986      | 7275| 26.329| 0.490 |102  | 2007-2012      | 22556|61.182| 0.576 |
| 77  | 1982-1987      | 7342| 26.996| 0.490 |103  | 2008-2013      | 21460|49.790| 0.594 |
TABLE S2. Year by year statistics for the World Trade Web. Columns are: the time of each network (Year), the number of nodes ($N$), the average degree ($\langle k \rangle$), and the average local clustering coefficient ($\langle c \rangle$).

| No. | Year | $N$ | $\langle k \rangle$ | $\langle c \rangle$ | No. | Year | $N$ | $\langle k \rangle$ | $\langle c \rangle$ |
|-----|------|-----|---------------------|---------------------|-----|------|-----|---------------------|---------------------|
| 0   | 1870 | 24  | 3.667               | 0.709               | 36  | 1906 | 32  | 3.938               | 0.707               |
| 1   | 1871 | 24  | 4.167               | 0.704               | 37  | 1907 | 31  | 3.484               | 0.785               |
| 2   | 1872 | 23  | 4.348               | 0.709               | 38  | 1908 | 31  | 4.839               | 0.741               |
| 3   | 1873 | 26  | 3.154               | 0.700               | 39  | 1909 | 32  | 4.875               | 0.801               |
| 4   | 1874 | 26  | 3.769               | 0.713               | 40  | 1910 | 37  | 3.405               | 0.711               |
| 5   | 1875 | 25  | 2.800               | 0.420               | 41  | 1911 | 38  | 3.158               | 0.641               |
| 6   | 1876 | 27  | 3.333               | 0.759               | 42  | 1912 | 41  | 3.463               | 0.685               |
| 7   | 1877 | 27  | 3.630               | 0.709               | 43  | 1913 | 41  | 3.805               | 0.772               |
| 8   | 1878 | 31  | 3.806               | 0.774               | 44  | 1914 | -    | -                   | -                   |
| 9   | 1879 | 28  | 3.786               | 0.664               | 45  | 1915 | -    | -                   | -                   |
| 10  | 1880 | 29  | 3.793               | 0.707               | 46  | 1916 | -    | -                   | -                   |
| 11  | 1881 | 28  | 4.000               | 0.700               | 47  | 1917 | -    | -                   | -                   |
| 12  | 1882 | 27  | 3.185               | 0.614               | 48  | 1918 | -    | -                   | -                   |
| 13  | 1883 | 29  | 4.483               | 0.693               | 49  | 1919 | -    | -                   | -                   |
| 14  | 1884 | 28  | 3.357               | 0.690               | 50  | 1920 | 50  | 3.640               | 0.737               |
| 15  | 1885 | 28  | 3.929               | 0.718               | 51  | 1921 | 52  | 3.346               | 0.597               |
| 16  | 1886 | 29  | 3.655               | 0.719               | 52  | 1922 | 53  | 3.321               | 0.704               |
| 17  | 1887 | 29  | 4.276               | 0.761               | 53  | 1923 | 54  | 3.667               | 0.686               |
| 18  | 1888 | 29  | 4.069               | 0.710               | 54  | 1924 | 54  | 3.630               | 0.819               |
| 19  | 1889 | 29  | 4.069               | 0.702               | 55  | 1925 | 51  | 3.490               | 0.767               |
| 20  | 1890 | 31  | 3.613               | 0.677               | 56  | 1926 | 53  | 3.547               | 0.661               |
| 21  | 1891 | 31  | 4.194               | 0.732               | 57  | 1927 | 55  | 3.091               | 0.723               |
| 22  | 1892 | 30  | 3.933               | 0.673               | 58  | 1928 | 56  | 3.643               | 0.798               |
| 23  | 1893 | 31  | 3.871               | 0.682               | 59  | 1929 | 56  | 3.607               | 0.758               |
| 24  | 1894 | 25  | 4.720               | 0.732               | 60  | 1930 | 46  | 2.739               | 0.618               |
| 25  | 1895 | 28  | 6.214               | 0.766               | 61  | 1931 | 57  | 3.825               | 0.832               |
| 26  | 1896 | 31  | 5.161               | 0.698               | 62  | 1932 | 57  | 3.439               | 0.723               |
| 27  | 1897 | 32  | 4.125               | 0.665               | 63  | 1933 | 56  | 3.321               | 0.701               |
| 28  | 1898 | 34  | 4.294               | 0.662               | 64  | 1934 | 57  | 3.193               | 0.642               |
| 29  | 1899 | 34  | 5.000               | 0.695               | 65  | 1935 | 53  | 3.396               | 0.642               |
| 30  | 1900 | 29  | 4.000               | 0.740               | 66  | 1936 | 59  | 3.085               | 0.618               |
| 31  | 1901 | 28  | 4.643               | 0.722               | 67  | 1937 | 60  | 3.067               | 0.313               |
| 32  | 1902 | 28  | 4.071               | 0.721               | 68  | 1938 | 56  | 3.000               | 0.248               |
| 33  | 1903 | 29  | 4.552               | 0.707               | 69  | 1939 | -    | -                   | -                   |
| 34  | 1904 | 29  | 4.345               | 0.700               | 70  | 1940 | -    | -                   | -                   |
| 35  | 1905 | 34  | 4.118               | 0.621               | 71  | 1941 | -    | -                   | -                   |
| No. | Year | N  | $\langle k \rangle$ | $\langle c \rangle$ |
|-----|------|----|---------------------|---------------------|
| 72  | 1942 |    | -                   | -                   |
| 73  | 1943 |    | -                   | -                   |
| 74  | 1944 |    | -                   | -                   |
| 75  | 1945 |    | -                   | -                   |
| 76  | 1946 |    | -                   | -                   |
| 77  | 1947 |    | -                   | -                   |
| 78  | 1948 | 61 | 2.426               | 0.236               |
| 79  | 1949 | 62 | 3.097               | 0.826               |
| 80  | 1950 | 64 | 3.562               | 0.828               |
| 81  | 1951 | 65 | 3.908               | 0.793               |
| 82  | 1952 | 67 | 4.328               | 0.726               |
| 83  | 1953 | 69 | 3.594               | 0.679               |
| 84  | 1954 | 70 | 4.514               | 0.639               |
| 85  | 1955 | 78 | 4.538               | 0.698               |
| 86  | 1956 | 80 | 4.100               | 0.640               |
| 87  | 1957 | 82 | 4.000               | 0.644               |
| 88  | 1958 | 81 | 4.469               | 0.648               |
| 89  | 1959 | 83 | 4.771               | 0.639               |
| 90  | 1960 | 98 | 4.429               | 0.639               |
| 91  | 1961 | 106| 5.038               | 0.727               |
| 92  | 1962 | 108| 3.759               | 0.613               |
| 93  | 1963 | 108| 4.056               | 0.641               |
| 94  | 1964 | 118| 5.068               | 0.725               |
| 95  | 1965 | 121| 5.074               | 0.748               |
| 96  | 1966 | 125| 6.112               | 0.716               |
| 97  | 1967 | 125| 6.288               | 0.725               |
| 98  | 1968 | 127| 5.654               | 0.690               |
| 99  | 1969 | 131| 5.847               | 0.668               |
| 100 | 1970 | 131| 4.718               | 0.629               |
| 101 | 1971 | 136| 5.397               | 0.686               |
| 102 | 1972 | 137| 5.635               | 0.667               |
| 103 | 1973 | 138| 5.275               | 0.634               |
| 104 | 1974 | 141| 6.383               | 0.670               |
| 105 | 1975 | 147| 6.925               | 0.684               |
| 106 | 1976 | 149| 6.456               | 0.682               |
| 107 | 1977 | 150| 6.427               | 0.697               |
TABLE S3. Overview of the considered real-world networks. Columns are: the name of each network (Name), the number of nodes ($N$), the average degree ($\langle k \rangle$), the average local clustering coefficient ($\langle c \rangle$), the exponent of the power-law degree distribution ($\gamma$), the hyperbolic embedding parameter $\beta$ and $\mu$, and the stable distribution fitting parameters $[\alpha, \eta, c, d]$.

| No. | Name   | $N$ | $\langle k \rangle$ | $\langle c \rangle$ | $\gamma$ | $\beta$ | $\mu$ | $[\alpha, \eta, c, d]$ |
|-----|--------|-----|---------------------|---------------------|----------|---------|------|------------------------|
| 0   | Airports | 3397 | 11.32               | 0.63                | 1.88     | 1.955   | 0.0274598 | [0.500, 1.000, 3.228, -1.081] |
| 1   | Drosophila | 1770 | 10.01               | 0.34                | 1.91     | 1.085   | 0.00419572 | [0.787, 0.999, 3.744, -8.593] |
| 2   | Enron | 33696 | 10.73               | 0.70                | 2.66     | 2.663   | 0.0365131 | [0.709, 1.000, 2.049, -2.252] |
| 3   | Internet | 23748 | 4.92                | 0.61                | 2.17     | 1.979   | 0.0640209 | [0.781, 0.962, 0.840, -0.585] |
| 4   | Metabolic | 1436 | 6.57                | 0.54                | 2.6      | 2.104   | 0.0507968 | [0.918, 0.759, 2.337, -8.581] |
| 5   | Music | 2476 | 16.66               | 0.82                | 2.27     | 2.192   | 0.0207461 | [0.767, 1.000, 4.117, -9.034] |
| 6   | Proteome | 4100 | 6.52                | 0.09                | 2.25     | 1.005   | 0.000827578 | [0.783, 0.999, 10.322, -19.597] |
| 7   | Words | 7377 | 11.99               | 0.47                | 2.25     | 1.006   | 0.000512119 | [0.759, 1.000 3.905, -7.023] |
| 8   | Facebook | 320  | 14.81               | 0.49                | 3.80     | 1.868   | 0.0199591 | [0.923, 0.999, 77.544, -620.528] |

43
[1] V. Allee, The knowledge evolution (Routledge, 2012).
[2] J. Tir, P. Schafer, P. F. Diehl, G. Goertz, Territorial changes, 1816–1996: Procedures and data. Conflict Management and Peace Science 16, 89 (1998).
[3] E. Hannezo, et al., A unifying theory of branching morphogenesis. Cell 171, 242 (2017).
[4] A. Terui, et al., Metapopulation stability in branching river networks. Proceedings of the National Academy of Sciences 115, E5963 (2018).
[5] N. D. Sergey, F. F. M. Jose, Handbook of Graphs and Networks: From the Genome to the Internet, eds. S. Bornholdt and H.G. Schuster (Wiley-VCH, Berlin, 2002) (2002), pp. 318–341.
[6] A. Barabási, R. Albert, Emergence of scaling in random networks. science 286, 509 (1999).
[7] F. Papadopoulos, M. Kitsak, M. Á. Serrano, M. Boguñá, D. Krioukov, Popularity versus similarity in growing networks. Nature 489, 537 (2012).
[8] R. V. Solé, R. Pastor-Satorras, E. Smith, T. B. Kepler, A model of large-scale proteome evolution. Advances in Complex Systems 5, 43 (2002).
[9] R. Pastor-Satorras, E. Smith, R. V. Solé, Evolving protein interaction networks through gene duplication. Journal of Theoretical biology 222, 199 (2003).
[10] A. N. Kolmogorov, N. A. Dmitriev, Branching random processes. Dokl. Akad. Nauk SSSR 56, 7 (1947).
[11] G. García-Pérez, M. Boguñá, M. Á. Serrano, Multiscale unfolding of real networks by geometric renormalization. Nature Physics 14, 583 (2018).
[12] K. G. Wilson, The renormalization group: Critical phenomena and the kondo problem. Reviews of modern physics 47, 773 (1975).
[13] K. G. Wilson, The renormalization group and critical phenomena. Reviews of Modern Physics 55, 583 (1983).
[14] L. P. Kadanoff, Statistical physics: statics, dynamics and renormalization (World Scientific Publishing Company, 2000).
[15] C. Song, S. Havlin, H. A. Makse, Origins of fractality in the growth of complex networks. Nature Physics 2, 275 (2006).
[16] M. Á. Serrano, D. Krioukov, M. Boguñá, Self-similarity of complex networks and hidden metric spaces. *Physical review letters* **100**, 078701 (2008).

[17] D. Krioukov, F. Papadopoulos, A. Vahdat, M. Boguñá, Curvature and temperature of complex networks. *Physical Review E* **80**, 035101 (2009).

[18] G. García-Pérez, M. Á. Serrano, M. Boguñá, Soft communities in similarity space. *Journal of Statistical Physics* **173**, 775 (2018).

[19] K. Zuev, M. Boguná, G. Bianconi, D. Krioukov, Emergence of soft communities from geometric preferential attachment. *Scientific reports* **5**, 9421 (2015).

[20] D. Hric, K. Kaski, M. Kivelä, Stochastic block model reveals maps of citation patterns and their evolution in time. *Journal of Informetrics* **12**, 757 (2018).

[21] S. Fortunato, *et al.*, Science of science. *Science* **359**, eaao0185 (2018).

[22] G. García-Pérez, M. Boguñá, A. Allard, M. Á. Serrano, The hidden hyperbolic geometry of international trade: World trade atlas 1870–2013. *Scientific reports* **6**, 33441 (2016).

[23] S. Milojević, Quantifying the cognitive extent of science. *Journal of Informetrics* **9**, 962 (2015).

[24] F. Radicchi, S. Fortunato, C. Castellano, Universality of citation distributions: Toward an objective measure of scientific impact. *Proceedings of the National Academy of Sciences* **105**, 17268 (2008).

[25] D. Wang, C. Song, A. Barabási, Quantifying long-term scientific impact. *Science* **342**, 127 (2013).

[26] Correlates of War Project, State system membership list, v2016, [http://correlatesofwar.org](http://correlatesofwar.org) (2017).

[27] V. D. Blondel, J. L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment* **2008**, P10008 (2008).

[28] N. X. Vinh, J. Epps, J. Bailey, Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance. *Journal of Machine Learning Research* **11**, 2837 (2010).

[29] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, M. Boguñá, Hyperbolic geometry of complex networks. *Physical Review E* **82**, 036106 (2010).

[30] M. Boguñá, F. Papadopoulos, D. Krioukov, Sustaining the internet with hyperbolic mapping. *Nature communications* **1**, 62 (2010).
[31] G. García-Pérez, A. Allard, M. Á. Serrano, M. Boguñá, Mercator: uncovering faithful hyperbolic embeddings of complex networks. *arXiv:1904.10814* (2019).

[32] S. Borak, W. Härdle, R. Weron, *Statistical tools for finance and insurance* (Springer, 2005), pp. 21–44.

[33] B. Gnedenko, A. Kolmogorov, *Limit distributions for sums of independent random variables* (Cambridge, Addison-Wesley, 1954).

[34] B. B. Mandelbrot, The variation of certain speculative prices. *The Journal of Business* **36**, 394 (1963).

[35] E. F. Fama, The behavior of stock-market prices. *The journal of Business* **38**, 34 (1965).

[36] J. P. Nolan, *Stable Distributions-Models for Heavy Tailed Data* (Springer Nature, Boston, 2019).

[37] M. Kateregga, S. Mataramvura, D. Taylor, Parameter estimation for stable distributions with application to commodity futures log-returns. *Cogent Economics & Finance* **5**, 1318813 (2017).

[38] J. P. Nolan, Numerical calculation of stable densities and distribution functions. *Communications in statistics. Stochastic models* **13**, 759 (1997).

[39] J. Royuela-del Val, F. Simmross-Wattenberg, C. Alberola-López, libstable: Fast, parallel, and high-precision computation of α-stable distributions in r, c/c++, and matlab. *Journal of Statistical Software* **78** (2017).

[40] M. Á. Serrano, D. Krioukov, M. Boguñá, Percolation in self-similar networks. *Physical review letters* **106**, 048701 (2011).

[41] S. N. Dorogovtsev, A. V. Goltsev, J. F. Mendes, Critical phenomena in complex networks. *Reviews of Modern Physics* **80**, 1275 (2008).

[42] M. Fire, R. Puzis, Organization mining using online social networks. *Networks and Spatial Economics* **16**, 545 (2016).

[43] M. Kuperman, D. Zanette, Stochastic resonance in a model of opinion formation on small-world networks. *The European Physical Journal B-Condensed Matter and Complex Systems* **26**, 387 (2002).

[44] L. Gammaitoni, P. Hänggi, P. Jung, F. Marchesoni, Stochastic resonance. *Reviews of modern physics* **70**, 223 (1998).

[45] A. Pikovsky, A. Zaikin, M. A. de La Casa, System size resonance in coupled noisy systems and in the ising model. *Physical review letters* **88**, 050601 (2002).
[46] R. Toral, C. J. Tessone, Finite size effects in the dynamics of opinion formation. *Communications in Computational Physics* **1**, 1 (2006).

[47] P. Jung, P. Hänggi, Stochastic nonlinear dynamics modulated by external periodic forces. *EPL (Europhysics Letters)* **8**, 505 (1989).

[48] M. Newman, R. Ziff, Efficient monte carlo algorithm and high-precision results for percolation. *Physical Review Letters* **85**, 4104 (2000).

[49] M. Á. Serrano, M. Boguñá, A. Vespignani, Extracting the multiscale backbone of complex weighted networks. *Proceedings of the national academy of sciences* **106**, 6483 (2009).

[50] K. Claffy, Y. Hyun, K. Keys, M. Fomenkov, D. Krioukov, 2009 Cybersecurity Applications & Technology Conference for Homeland Security (IEEE, 2009), pp. 205–211.

[51] M. Á. Serrano, M. Boguñá, F. Sagüés, Uncovering the hidden geometry behind metabolic networks. *Molecular biosystems* **8**, 843 (2012).

[52] Openflights network dataset, [http://konect.uni-koblenz.de/networks/openflights](http://konect.uni-koblenz.de/networks/openflights) (2016).

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

[54] S.-Y. Takemura, et al., A visual motion detection circuit suggested by drosophila connectomics. *Nature* **500**, 175 (2013).

[55] B. Klimt, Y. Yang, Introducing the enron corpus. *In CEAS’ 04: Proceedings of the 1st Conference on Email and Anti-Spam* (2004).

[56] J. Leskovec, K. J. Lang, A. Dasgupta, M. W. Mahoney, Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters. *Internet Mathematics* **6**, 29 (2009).

[57] J. Serrà, Á. Corral, M. Boguñá, M. Haro, J. L. Arcos, Measuring the evolution of contemporary western popular music. *Scientific reports* **2**, 521 (2012).

[58] T. Rolland, et al., A proteome-scale map of the human interactome network. *Cell* **159**, 1212 (2014).

[59] R. Milo, et al., Superfamilies of evolved and designed networks. *Science* **303**, 1538 (2004).

[60] C. J. Tessone, R. Toral, System size stochastic resonance in a model for opinion formation. *Physica A: Statistical Mechanics and its Applications* **351**, 106 (2005).

[61] P. Lévy, Calcul des probabilités (gauthier-villars) (1925).