

# Multiscale unfolding of real networks by geometric renormalization

Guillermo García-Pérez<sup>1,2</sup>, Marián Boguñá<sup>1,2</sup>  and M. Ángeles Serrano<sup>1,2,3\*</sup>

**Symmetries in physical theories denote invariance under some transformation, such as self-similarity under a change of scale. The renormalization group provides a powerful framework to study these symmetries, leading to a better understanding of the universal properties of phase transitions. However, the small-world property of complex networks complicates application of the renormalization group by introducing correlations between coexisting scales. Here, we provide a framework for the investigation of complex networks at different resolutions. The approach is based on geometric representations, which have been shown to sustain network navigability and to reveal the mechanisms that govern network structure and evolution. We define a geometric renormalization group for networks by embedding them into an underlying hidden metric space. We find that real scale-free networks show geometric scaling under this renormalization group transformation. We unfold the networks in a self-similar multilayer shell that distinguishes the coexisting scales and their interactions. This in turn offers a basis for exploring critical phenomena and universality in complex networks. It also affords us immediate practical applications, including high-fidelity smaller-scale replicas of large networks and a multiscale navigation protocol in hyperbolic space, which betters those on single layers.**

The definition of self-similarity and scale invariance<sup>1,2</sup> in complex networks has been limited by the lack of a valid source of geometric length-scale transformations. Previous efforts to study these symmetries are based on topology and include spectral coarse-graining<sup>3</sup> or box-covering procedures based on shortest path lengths between nodes<sup>4–9</sup>. However, the collection of shortest paths is a poor source of length-based scaling factors in networks due to the small-world<sup>10</sup> or even ultras-small-world<sup>11</sup> properties, and the problem remained open. Other studies have approached the multiscale structure of network models in a more geometric way<sup>12,13</sup>, but their findings cannot be directly applied to real-world networks.

Models of complex networks based on hidden metric spaces<sup>14–17</sup> open the door to a proper geometric definition of self-similarity and scale invariance and to an unfolding of the different scales present in the structure of real networks. These models are able to explain universal features shared by real networks—including the small-world property, scale-free degree distributions and clustering—and also fundamental mechanisms, such as preferential attachment in growing networks<sup>18</sup> and the emergence of communities<sup>19</sup>.

Naturally, the geometricalization of networks allows a reservoir of distance scales so that we can borrow concepts and techniques from the renormalization group in statistical physics<sup>20,21</sup>. By recursive averaging over short-distance degrees of freedom, renormalization has successfully explained, for instance, the universality properties of critical behaviour in phase transitions<sup>22</sup>. In this study, we introduce a geometric renormalization group for complex networks (RGN). The method, inspired by the block spin renormalization group devised by L. P. Kadanoff<sup>20</sup>, relies on a geometric embedding of the networks to coarse-grain neighbouring nodes into supernodes and to define a new map that progressively selects long-range connections.

## Evidence of geometric scaling in real networks

Hidden metric space network models couple the topology of a network to an underlying geometry through a universal connectivity

law depending on distances on such space, which combine popularity and similarity dimensions<sup>14,17,18</sup>, such that more popular and similar nodes have more chance to interact. Popularity is related to the degrees of the nodes<sup>14,23</sup>, and similarity stands as an aggregate of all other attributes that modulate the likelihood of interactions. These two dimensions define a hyperbolic plane as the effective geometry of networks, and their contribution to the probability of connection can be explicit or combined into an effective hyperbolic distance. This gives rise to the two isomorphic geometric models,  $S^1$  and  $H^2$ . In the  $S^1$  model<sup>14</sup>, the popularity of a node  $i$  is associated with a hidden degree  $\kappa_i$ , complemented by its angular position in the one-dimensional sphere (or circle) as a similarity measure, such that the probability of connection increases with the product of the hidden degrees and decreases with their distance along the circle (equation (5) in Methods). Reciprocally, in the equivalent  $H^2$  model<sup>16,17</sup>, the hidden degree is transformed into a radial coordinate, such that higher degree nodes are placed closer to the centre of the hyperbolic disk, while the angular coordinate remains as in the  $S^1$  circle, and the probability of connection decreases with the hyperbolic distance. In their scale-free version, both models have only three parameters  $\mu$ ,  $\gamma$  and  $\beta$ , which control the average degree  $\langle k \rangle$ , the exponent of the degree distribution  $\gamma$  and the local clustering coefficient  $\bar{c}$ , respectively. The radius  $R$  of the  $S^1$  circle is adjusted to maintain a constant density of nodes equal to one.

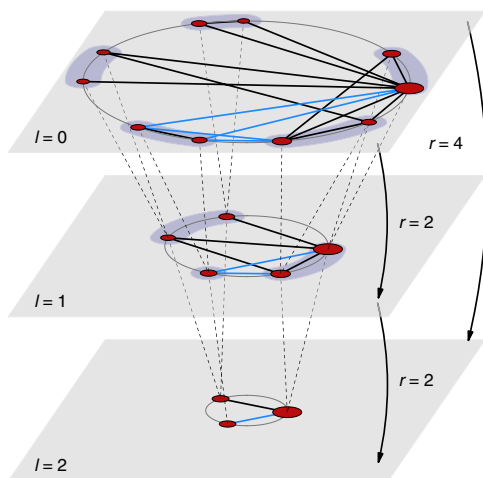
The renormalization transformation is defined on the basis of the similarity dimension represented by the angular coordinate of the nodes. We present here the formulation for the  $S^1$  model, as it makes the similarity dimension explicit and is mathematically more tractable. The transformation zooms out by changing the minimum length scale from that of the original network to a larger value. It proceeds by, first, defining non-overlapping blocks of consecutive nodes of size  $r$  along the circle and, second, by coarse-graining the blocks into supernodes. Each supernode is then placed within the angular region defined by the corresponding block so that the order of nodes is preserved. All the links between some node in

<sup>1</sup>Departament de Física de la Matèria Condensada, Universitat de Barcelona, Barcelona, Spain. <sup>2</sup>Universitat de Barcelona Institute of Complex Systems (UBICS), Universitat de Barcelona, Barcelona, Spain. <sup>3</sup>ICREA, Barcelona, Spain. \*e-mail: [marian.serrano@ub.edu](mailto:marian.serrano@ub.edu)

one supernode and some node in the other, if any, are renormalized into a single link between the two supernodes. This operation can be iterated starting from the original network at layer  $l=0$ . Finally, the set of renormalized network layers  $l$ , each  $r'$  times smaller than the original one, forms a multiscale shell of the network. Figure 1 illustrates the process.

In this study, we apply the RGN to eight different real scale-free networks from very different domains: technology (Internet<sup>24</sup>), transportation (Airports<sup>25,26</sup>), biology (Metabolic<sup>27</sup>, Proteome<sup>28</sup> and *Drosophila*<sup>29</sup>), social (Enron<sup>30,31</sup>) and languages (Music<sup>32</sup> and Words<sup>33</sup>) (section I in the Supplementary Information). First, their geometric maps are obtained by embedding the nodes in the underlying geometry using statistical inference techniques, which identify the hidden degrees and angular coordinates, maximizing the likelihood that the topology of the network is reproduced by the model<sup>15,34</sup>. Second, we apply the coarse-graining by defining blocks of size  $r=2$ , and iterate the process. In the limit  $N \rightarrow \infty$ , where  $N$  is the number of nodes, the RGN can be applied up to any desired scale of observation, whereas it is bounded to order  $\mathcal{O}(\log N)$  iterations in finite systems.

The resulting topological features of three of the renormalized networks are shown in Fig. 2 (see Supplementary Fig. 1 for the others). We observe that the degree distributions, degree-degree correlations, clustering spectra and community structures (see



**Fig. 1 | Geometric renormalization transformation for complex networks.**

Each layer is obtained after a renormalization step with resolution  $r$  starting from the original network in  $l=0$ . Each node  $i$  in red is placed at an angular position  $\theta_i^{(l)}$  on the  $\mathbb{S}^1$  circle and has a size proportional to the logarithm of its hidden degree  $\kappa_i^{(l)}$ . Straight solid lines represent the links in each layer. Coarse-graining blocks correspond to the blue shadowed areas, and dashed lines connect nodes to their supernodes in layer  $l+1$ . Two supernodes in layer  $l+1$  are connected if and only if, in layer  $l$ , some node in one supernode is connected to some node in the other (blue links give an example). The geometric renormalization transformation has Abelian semigroup structure with respect to the composition, meaning that a certain number of iterations of a given resolution are equivalent to a single transformation of higher resolution. For instance, in the figure, the same transformation with  $r=4$  goes from  $l=0$  to  $l=2$  in a single step. Whenever the number of nodes is not divisible by  $r$ , the last supernode in a layer contains less than  $r$  nodes, as in the example at  $l=1$ . The RGN transformations are valid for uneven supernode sizes as well; one could divide the circle into equally sized sectors of a certain arc length such that they contain on average a constant number of nodes. The set of transformations parametrized by  $r$  does not include an inverse element to reverse the process.

Methods) show self-similar behaviour. The last property suggests a new and efficient multiscale community detection algorithm for complex networks<sup>35–37</sup>.

### Geometric renormalization of the $\mathbb{S}^1/\mathbb{H}^2$ model

The self-similarity exhibited by real-world networks can be understood in terms of their congruency with the hidden metric space network model. As we show analytically, the  $\mathbb{S}^1$  and  $\mathbb{H}^2$  models are renormalizable in a geometric sense. This means that if a real scale-free network is compatible with the model and admits a good embedding, as it is the case for the real networks analysed in this study, the model will be able to predict its self-similarity and geometric scaling.

We demonstrate next the renormalizability of the  $\mathbb{S}^1$  model (see section II in the Supplementary Information for mathematical details and also for the definition of the RGN in hyperbolic space). The renormalized networks remain maximally congruent with the hidden metric space model by assigning a new hidden degree  $\kappa_i^{(l+1)}$  to supernode  $i$  in layer  $l+1$  as a function of the hidden degrees of the nodes it contains in layer  $l$  according to:

$$\kappa_i^{(l+1)} = \left( \sum_{j=1}^r (\kappa_j^{(l)})^\beta \right)^{1/\beta} \quad (1)$$

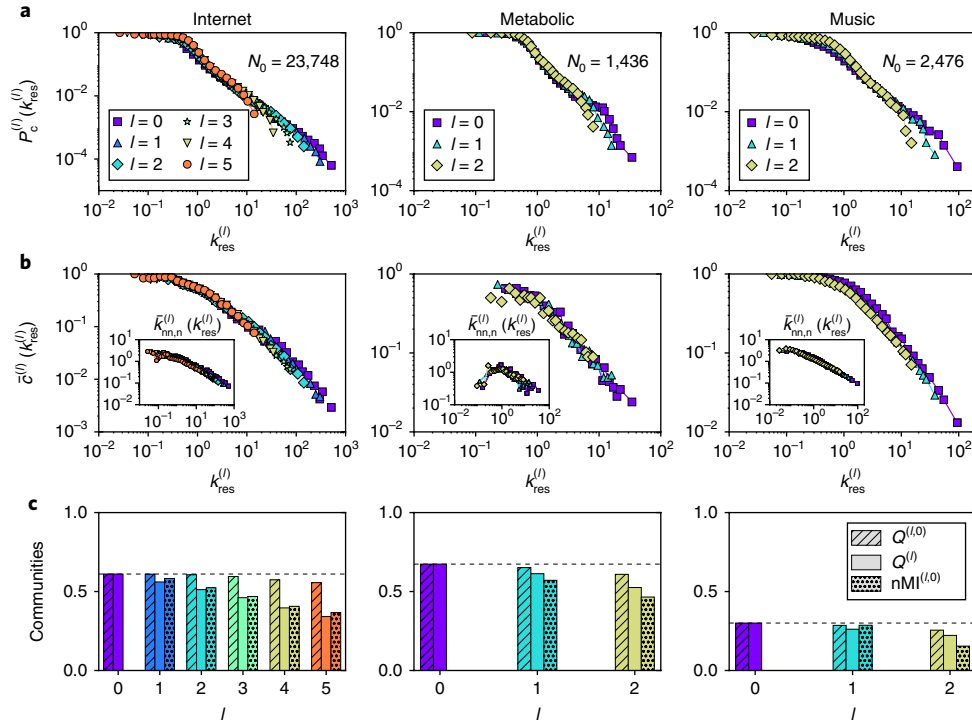
as well as an angular coordinate  $\theta_i^{(l+1)}$  given by:

$$\theta_i^{(l+1)} = \left( \frac{\sum_{j=1}^r (\theta_j^{(l)} \kappa_j^{(l)})^\beta}{\sum_{j=1}^r (\kappa_j^{(l)})^\beta} \right)^{1/\beta} \quad (2)$$

The global parameters need to be rescaled as  $\mu^{(l+1)} = \mu^{(l)}/r$ ,  $\beta^{(l+1)} = \beta^{(l)}$ , and  $R^{(l+1)} = R^{(l)}/r$ . This implies that the probability  $p_{ij}^{(l+1)}$  for two supernodes  $i$  and  $j$  to be connected in layer  $l+1$  maintains its original form (equation (5) and Fig. 3a). This applies both to the model and to real networks as long as they admit a good embedding (Supplementary Fig. 2). In addition, notice that the renormalization transformations of the geometric layout also have the Abelian semigroup structure.

As the networks remain congruent with the  $\mathbb{S}^1$  model, hidden degrees  $\kappa^{(l)}$  remain proportional to observed degrees  $k^{(l)}$ , which allows us to explore the degree distribution of the renormalized layers analytically. It can be shown that, if the original distribution of hidden degrees is a power law with characteristic exponent  $\gamma$ , the renormalized distribution is also a power law with the same exponent asymptotically, as long as  $(\gamma-1)/2 < \beta$  (section II in the Supplementary Information), with the only difference being the average degree. Interestingly, the global parameter controlling the clustering coefficient,  $\beta$ , does not change along the flow, which explains the self-similarity of the clustering spectra. Finally, the transformation for the angles (equation (2)) preserves the ordering of nodes and the heterogeneity in their angular density and, as a consequence, the community structure is preserved in the flow<sup>15,19,27,38</sup>, as shown in Fig. 3b. The model is therefore renormalizable, and RGN realizations at any scale belong to the same ensemble with a different average degree, which should be rescaled to produce self-similar replicas.

A good approximation of the behaviour of the average degree for very large networks can be calculated by taking into account the transformation of hidden degrees along the RGN flow (equation (1) and section II in the Supplementary Information). We obtain  $\langle k \rangle^{(l+1)} = r^\nu \langle k \rangle^{(l)}$ , with a scaling factor  $\nu$  depending on the



**Fig. 2 | Self-similarity of real networks along the RGN flow.** Each column shows the RGN flow with  $r=2$  of different topological features of the Internet autonomous systems (AS) network (left), the human Metabolic network (middle) and the Music network (right). **a**, Complementary cumulative distribution  $P_c$  of rescaled degrees  $k_{res}^{(l)} = k^{(l)} / \langle k^{(l)} \rangle$ . **b**, Degree-dependent clustering coefficient over rescaled-degree classes. Degree-degree correlations, as measured by the normalized average nearest-neighbour degree  $\bar{k}_{nn}^{(l)}(k_{res}^{(l)}) = \bar{k}_{nn}(k_{res}^{(l)}) \langle k^{(l)} \rangle / \langle (k^{(l)})^2 \rangle$ , are shown in the insets. **c**, RGN flow of the community structure.  $Q^{(l)}$  is the modularity in layer  $l$ ,  $Q^{(l)}$  is the modularity that the community structure of layer  $l$  induces in the original network and  $nMI^{(l,0)}$  is the normalized mutual information between the latter and the community structure detected directly in the original network (see Methods). The number of layers in each system is determined by their original size. The horizontal dashed lines indicate the modularity in the original networks.

connectivity structure of the original network. If  $0 < \frac{\gamma-1}{\beta} \leq 1$ , the flow is dominated by the exponent of the degree distribution  $\gamma$ , and the scaling factor is given by:

$$\nu = \frac{2}{\gamma-1} - 1 \tag{3}$$

whereas the flow is dominated by the strength of clustering if  $1 \leq \frac{\gamma-1}{\beta} < 2$  and

$$\nu = \frac{2}{\beta} - 1 \tag{4}$$

Therefore, if  $\gamma < 3$  or  $\beta < 2$  (phase I in Fig. 3c), then  $\nu > 0$  and the model flows towards a highly connected graph; the average degree is preserved if  $\gamma = 3$  and  $\beta \geq 2$  or  $\beta = 2$  and  $\gamma \geq 3$ , which indicates that the network is at the edge of the transition between the small-world and non-small-world phases; and  $\nu < 0$  if  $\gamma > 3$  and  $\beta > 2$ , causing the RGN flow to produce sparser networks approaching a unidimensional ring structure as a fixed point (phase II in Fig. 3c). In this case, the renormalized layers eventually lose the small-world property. Finally, if  $\beta < (\gamma-1)/2$ , the degree distribution becomes increasingly homogeneous as  $r \rightarrow \infty$  (phase III in Fig. 3c), revealing that degree heterogeneity is only present at short scales.

In Fig. 3c, several real networks are shown in the connectivity phase diagram. All of them lay in the region of small-world networks. Furthermore, all of them, except the Internet, Airports and *Drosophila* networks, belong to the  $\beta$ -dominated region. The inset shows the behaviour of the average degree of each layer  $l$ ,  $\langle k^{(l)} \rangle$ ; as predicted, it grows exponentially in all cases.

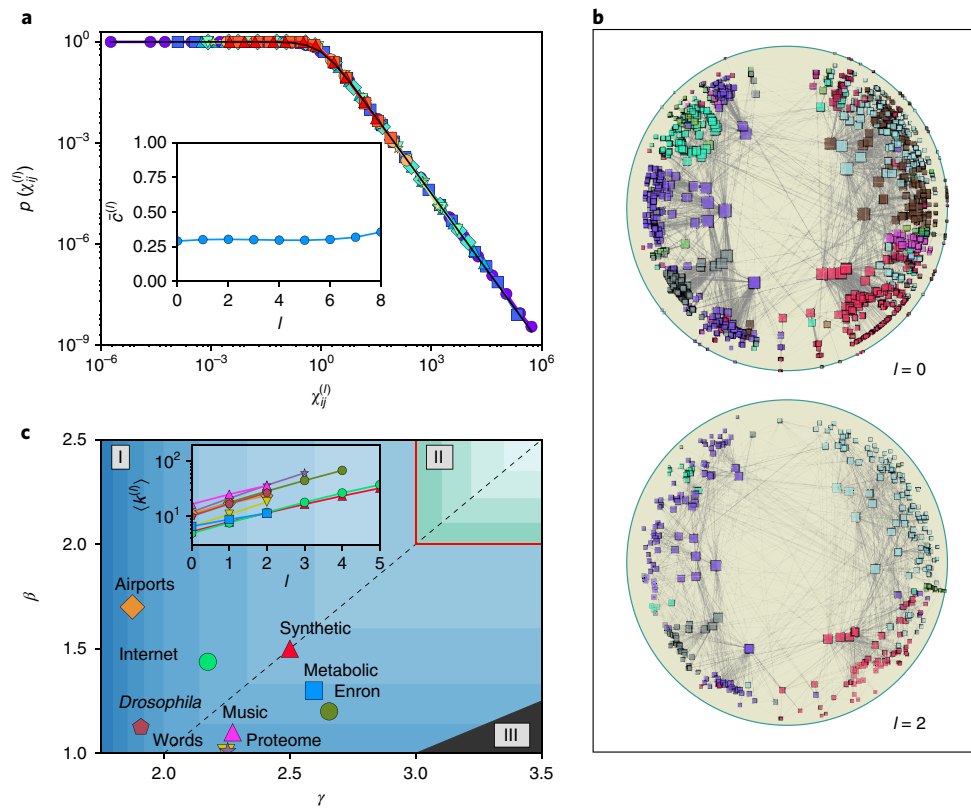
Interestingly, global properties of the model, such as those reflected in the spectrum of eigenvalues of both the adjacency and Laplacian matrices, and quantities such as the diffusion time and the restabilization time<sup>39</sup>, show a dependence on  $\gamma$  and  $\beta$ , which is in consonance with the one displayed by the RGN flow of the average degree (Supplementary Figs. 10–12). The  $S^1$  model seems to be more sensitive to small changes in degree heterogeneity in the region  $0 < \frac{\gamma-1}{\beta} \leq 1$ , whereas changes in clustering are better reflected when  $1 \leq \frac{\gamma-1}{\beta} \leq 2$ .

Finally, the RGN transformation can be reformulated for the model in  $D$  dimensions. We have recalculated the connectivity phase diagram in Fig. 3c, obtaining qualitatively the same transitions and phases, including region III. Interestingly, the high clustering coefficient observed in real networks poses an upper limit on the potential dimension of the similarity space. We have tested the renormalization transformation using the one-dimensional embedding of networks generated in higher dimensions for which the clustering is realistic, that is,  $D \lesssim 10$ , and found the same results as in the  $D=1$  case. The agreement is explained by the fact that the one-dimensional embedding provides a faithful representation for low-dimensional similarity-space networks (section II in the Supplementary Information).

### Applications

Next, we propose two specific applications. The first one, the production of downscaled network replicas, singles out a specific scale while the second one, a multiscale navigation protocol, exploits multiple scales simultaneously.

The downscaling of the topology of large real-world complex networks to produce smaller high-fidelity replicas can find useful



**Fig. 3 | RGN flow of synthetic and real networks.** **a**, Empirical connection probability ( $p$ ) in a synthetic  $\mathbb{S}^1$  network. Fraction of connected pairs of nodes as a function of  $x_{ij}^{(l)} = R^{(l)} \Delta \theta_{ij}^{(l)} / (\mu^{(l)} \kappa_i^{(l)} \kappa_j^{(l)})$  in the renormalized layers, from  $l=0$  to  $l=8$ , and  $r=2$ . The original synthetic network has  $N \cong 225,000$  nodes,  $\gamma=2.5$  and  $\beta=1.5$ . The black dashed line shows the theoretical curve (equation (5)). The inset shows the invariance of the mean local clustering along the RGN flow. **b**, Hyperbolic embedding of the Metabolic network (top) and its renormalized layer  $l=2$  (bottom). The colours of the nodes correspond to the community structure detected by the Louvain algorithm. Notice how the renormalized network preserves the original community structure despite being four times smaller. **c**, Real networks in the connectivity phase diagram. The synthetic network above is also shown. Darker blue (green) in the shaded areas represent higher values of the exponent  $\nu$ . The dashed line separates the  $\gamma$ -dominated region from the  $\beta$ -dominated region. In phase I,  $\nu > 0$  and the network flows towards a fully connected graph. In phase II,  $\nu < 0$  and the network flows towards a one-dimensional ring. The red thick line indicates  $\nu = 0$  and, hence, the transition between the small-world and non-small-world phases. In region III, the degree distribution loses its scale-freeness along the flow. The inset shows the exponential increase of the average degree of the renormalized real networks  $\langle k^{(l)} \rangle$  with respect to  $l$ .

applications, for instance, as an alternative or guidance to sampling methods in large-scale simulations and, in networked communication systems like the Internet, as a reduced testbed to analyse the performance of new routing protocols<sup>40–43</sup>. Downscaled network replicas can also be used to perform finite size scaling of critical phenomena in real networks, so that critical exponents could be evaluated starting from a single size instance network. However, the success of such programmes is based on the quality of the downscaled versions, which should reproduce not only local properties of the original network but also its mesoscopic structure. We now present a method for their construction that exploits the fact that, under renormalization, a scale-free network remains self-similar and congruent with the underlying geometric model in the whole self-similarity range of the multilayer shell.

The idea is to single out a specific scale after a certain number of renormalization steps. Typically, the renormalized average degree of real networks increases in the flow (see inset in Fig. 3c), so we apply a pruning of links to reduce the density to the level of the original network (see Methods). Basically, we re-adjust the average-degree parameter  $\mu^{(l)}$  in the  $\mathbb{S}^1$  model and then keep only the renormalized links that are consistent with the readjusted connection probabilities to obtain a statistically equivalent but reduced version.

To illustrate the high-fidelity that downscaled network replicas can achieve, we use them to reproduce the behaviour of

dynamical processes in real networks. We selected three different dynamical processes, the classic ferromagnetic Ising model, the susceptible–infected–susceptible (SIS) epidemic spreading model and the Kuramoto model of synchronization (see Methods). We test these dynamics in all the self-similar network layers of the real networks analysed in this study. Results are shown in Fig. 4 and Supplementary Fig. 13. Quite remarkably, for all dynamics and all networks, we observe very similar results between the original and downscaled replicas at all scales. This is particularly interesting as these dynamics have a strong dependence on the mesoscale structure of the underlying networks.

Next, we introduce a new multiscale navigation protocol for networks embedded in hyperbolic space, which improves single-layer results<sup>15</sup>. To this end, we exploit the quasi-isomorphism between the  $\mathbb{S}^1$  model and the  $\mathbb{H}^2$  model in hyperbolic space<sup>16,17</sup> to produce a purely geometric representation of the multiscale shell (see Methods).

The multiscale navigation protocol is based on greedy routing, in which a source node transmitting a packet to a target node sends it to its neighbour closest to the destination in the metric space. As performance metrics, we consider the success rate (fraction of successful greedy paths), and the stretch of successful path (ratio between the number of hops in the greedy path and the topological shortest path). Notice that greedy routing cannot guarantee the

existence of a successful path among all pairs of nodes; the packet can get trapped into a loop if sent to an already visited node. In this case, the multiscale navigation protocol can find alternative paths by taking advantage of the increased efficiency of greedy forwarding in the coarse-grained layers. When node  $i$  needs to send a packet to some destination node  $j$ , node  $i$  performs a virtual greedy forwarding step in the highest possible layer to find which supernode should be next in the greedy path. Based on this, node  $i$  then forwards the packet to its physical neighbour in the real network, which guarantees that it will eventually reach such supernode. The process is depicted in Fig. 5a (full details in Methods). To guarantee navigation inside supernodes, we require an extra condition in the renormalization process and only consider blocks of connected consecutive nodes (a single node can be left alone forming a supernode by itself). Notice that the new requirement does not alter the self-similarity of the renormalized networks forming the multiscale shell nor the congruency with the hidden metric space (section IV in the Supplementary Information).

Figure 5b shows the increase of the success rate as the number of renormalized layers  $L$  used in the navigation process is increased for the different real networks considered in this study. Interestingly, the success rate always increases, even in systems with very high navigability in  $l=0$  like the Internet, and this improvement increases the stretch of successful paths only mildly (Fig. 5c). Counterintuitively, the slight increase of the stretch reduces the burden on highly connected nodes (Fig. 5d). As the number of renormalized layers  $L$  increases, the average fraction of successful paths passing through the most connected hubs in the network decreases. The improvements come at the expense of adding information about the supernodes to the knowledge needed for standard greedy routing in single-layered networks. However, the trade-off between improvement and information overload is advantageous, as for many sys-

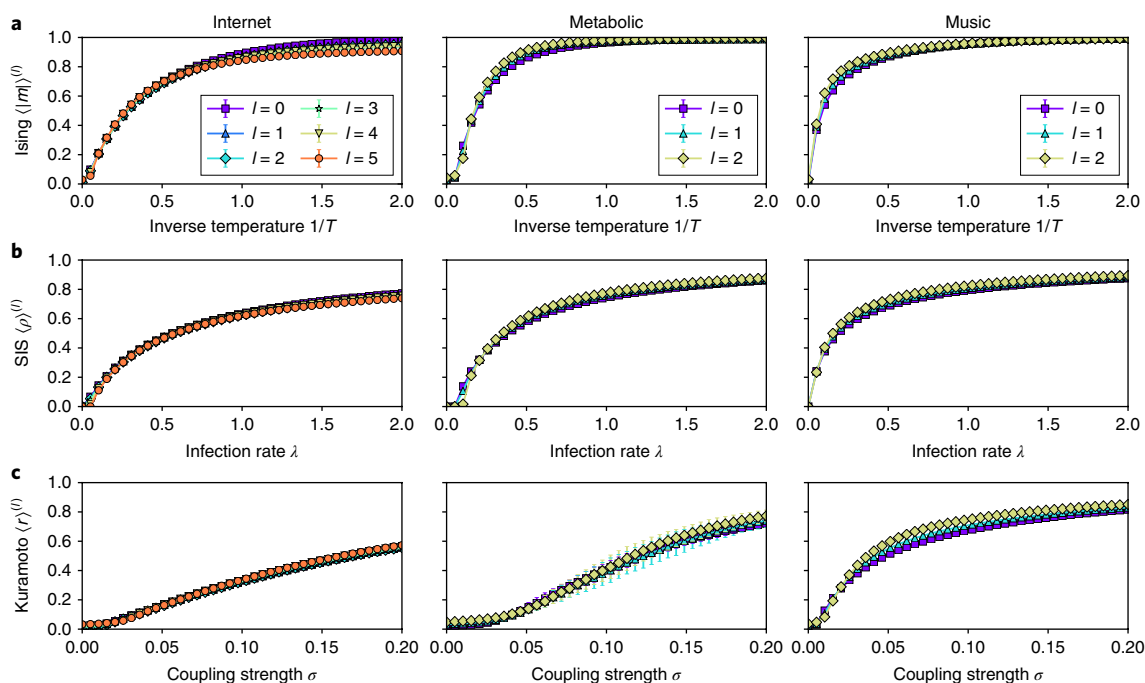
tems the addition of just one or two renormalized layer produces already a notable effect.

## Discussion

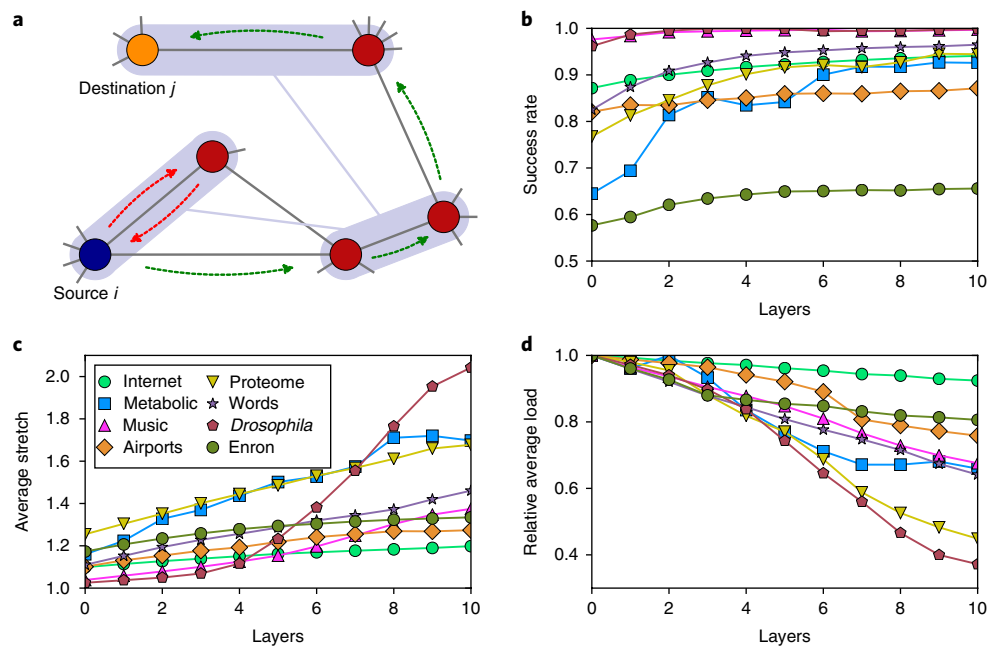
It is a very well established empirical fact that most real complex networks share a very special set of universal features. Among the most relevant ones, networks have heterogeneous degree distributions, strong clustering and are small world. Our hidden metric space network model<sup>14,17,18</sup>, independent of its  $S^1$  or  $\mathbb{H}^2$  formulation, provides a very natural explanation of these properties with a very limited number of parameters and using an effective hyperbolic geometry of two dimensions. Even if the model and the renormalization group can be formulated in  $D$  dimensions, the high clustering coefficient observed in real networks poses an upper limit on the potential dimension of the similarity space so that networks can be faithfully embedded in the one-dimensional representation. This is in line with the accumulated empirical evidence, which unambiguously supports the one-dimensional similarity plus degrees as an extremely good proxy for the geometry of real networks.

Interestingly, the existence of a metric space underlying complex networks allows us to define a geometric renormalization group that reveals their multiscale nature. Our geometric models of scale-free networks are shown to be self-similar under the RGN transformation. Even more important is the finding that self-similarity under geometric renormalization is a ubiquitous symmetry of real-world scale-free networks, which provides new evidence supporting the hypothesis that hidden metric spaces underlie real networks.

The renormalization group presented in this work is similar in spirit to the topological renormalization studied in refs<sup>4–9</sup> and should be taken as complementary. Instead of using shortest paths as a source of length scales to explore the fractality of networks, we use a continuum geometric framework that allows us to



**Fig. 4 | Dynamics on the downscaled replicas.** Each column shows the order parameters versus the control parameters of different dynamical processes on the original and downscaled replicas of the Internet AS network (left), the human Metabolic network (middle) and the Music network (right) with  $r=2$ , that is, every value of  $l$  identifies a network  $2^l$  times smaller than the original one. All points show the results averaged over 100 simulations. Error bars indicate the standard deviations of the order parameters. **a**, Magnetization  $\langle |m| \rangle^{(l)}$  of the Ising model as a function of the inverse temperature  $1/T$ . **b**, Prevalence  $\langle \rho \rangle^{(l)}$  of the SIS model as a function of the infection rate  $\lambda$ . **c**, Coherence  $\langle r \rangle^{(l)}$  of the Kuramoto model as a function of the coupling strength  $\sigma$ . In all cases, the curves of the smaller-scale replicas are extremely similar to the results obtained on the original networks.



**Fig. 5 | Performance of the multiscale navigation protocol in real networks.** **a**, Illustration of the multiscale navigation protocol. Red arrows show the unsuccessful greedy path in the original layer of a message attempting to reach the target yellow node. Green arrows show the successful greedy path from the same source using both layers. **b**, Success rate as a function of the number of layers used in the process, computed for  $10^5$  randomly selected pairs. **c**, Average stretch ( $l_g/l_s$ ), where  $l_g$  is the topological length of a path found by the algorithm and  $l_s$  is the actual shortest path length in the network. **d**, Relative average load  $\langle \lambda_i^l \rangle_{\text{Hubs}} / \langle \lambda_0^l \rangle_{\text{Hubs}}$  of hubs, where  $\lambda_i^l$  is the fraction of successful greedy paths that pass through node  $i$  in the multiscale navigation protocol with  $L$  renormalized layers. The averages are computed over the 20 highest-degree nodes in the network.

unveil the role of degree heterogeneity and clustering in the self-similarity properties of networks. In our model, a crucial point is the explicit contribution of degrees to the probability of connection, allows us to produce both short-range and long-range connections using a single mechanism captured in a universal connectivity law. The combination of similarity with degrees is a necessary condition to make the model predictive of the multiscale properties of real networks.

From a fundamental point of view, the geometric renormalization group introduced here has proven to be an exceptional tool to unravel the global organization of complex networks across scales and promises to become a standard methodology to analyse real complex networks. It can also help in areas such as the study of metapopulation models, in which transportation fluxes or population movements occur on both local and global scales<sup>44</sup>. From a practical point of view, we envision many applications. In large-scale simulations, downscaled network replicas could serve as an alternative or guidance to sampling methods, or for fast-track exploration of rough parameter spaces in the search of relevant regions. Downscaled versions of real networks could also be applied to perform finite size scaling, which would allow for the determination of critical exponents from single snapshots of their topology. Other possibilities include the development of a new multilevel community detection method<sup>45–47</sup> that would exploit the mesoscopic information encoded in the different observation scales.

## Methods

Methods, including statements of data availability and any associated accession codes and references, are available at <https://doi.org/10.1038/s41567-018-0072-5>.

Received: 12 June 2017; Accepted: 1 February 2018;  
Published online: 19 March 2018

## References

- Mandelbrot, B. *The Fractal Geometry of Nature* (W. H. Freeman and Company, San Francisco, CA, 1982).
- Stanley, H. E. *Introduction to Phase Transitions and Critical Phenomena* (Oxford Univ. Press, Oxford, 1971).
- Gfeller, D. & De Los Rios, P. Spectral coarse graining of complex networks. *Phys. Rev. Lett.* **99**, 038701 (2007).
- Song, C., Havlin, S. & Makse, H. A. Self-similarity of complex networks. *Nature* **433**, 392–395 (2005).
- Goh, K. I., Salvi, G., Kahng, B. & Kim, D. Skeleton and fractal scaling in complex networks. *Phys. Rev. Lett.* **96**, 018701 (2006).
- Song, C., Havlin, S. & Makse, H. A. Origins of fractality in the growth of complex networks. *Nat. Phys.* **2**, 275–281 (2006).
- Kim, J. S., Goh, K. I., Hahng, B. & Kim, D. Fractality and self-similarity in scale-free networks. *New J. Phys.* **9**, 177 (2007).
- Radicchi, F., Ramasco, J. J., Barrat, A. & Fortunato, S. Complex networks renormalization: flows and fixed points. *Phys. Rev. Lett.* **101**, 148701 (2008).
- Rozenfeld, H. D., Song, C. & Makse, H. A. Small-world to fractal transition in complex networks: a renormalization group approach. *Phys. Rev. Lett.* **104**, 025701 (2010).
- Watts, D. J. & Strogatz, S. H. Collective dynamics of ‘small-world’ networks. *Nature* **393**, 440–442 (1998).
- Cohen, R. & Havlin, S. Scale-free networks are ultrasmall. *Phys. Rev. Lett.* **90**, 058701 (2003).
- Newman, M. & Watts, D. Renormalization group analysis of the small-world network model. *Phys. Lett. A* **263**, 341–346 (1999).
- Boettcher, S. & Brunson, C. Renormalization group for critical phenomena in complex networks. *Front. Physiol.* **2**, 102 (2011).
- Serrano, M. Á., Krioukov, D. & Boguñá, M. Self-similarity of complex networks and hidden metric spaces. *Phys. Rev. Lett.* **100**, 078701 (2008).
- Boguñá, M., Papadopoulos, F. & Krioukov, D. Sustaining the Internet with hyperbolic mapping. *Nat. Commun.* **1**, 62 (2010).
- Krioukov, D., Papadopoulos, F., Vahdat, A. & Boguñá, M. Curvature and temperature of complex networks. *Phys. Rev. E* **80**, 035101 (2009).
- Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A. & Boguñá, M. Hyperbolic geometry of complex networks. *Phys. Rev. E* **82**, 036106 (2010).
- Papadopoulos, F., Kitsak, M., Serrano, M. A., Boguñá, M. & Krioukov, D. Popularity versus similarity in growing networks. *Nature* **489**, 537–540 (2012).

19. Zuev, K., Boguñá, M., Bianconi, G. & Krioukov, D. Emergence of soft communities from geometric preferential attachment. *Sci. Rep.* **5**, 9421 (2015).
20. Kadanoff, L. P. *Statistical Physics: Statics, Dynamics and Renormalization* (World Scientific, Singapore, 2000).
21. Wilson, K. G. The renormalization group: critical phenomena and the Kondo problem. *Rev. Mod. Phys.* **47**, 773–840 (1975).
22. Wilson, K. G. The renormalization group and critical phenomena. *Rev. Mod. Phys.* **55**, 583–600 (1983).
23. Boguñá, M. & Pastor-Satorras, R. Class of correlated random networks with hidden variables. *Phys. Rev. E* **68**, 036112 (2003).
24. Claffy, K., Hyun, Y., Keys, K., Fomenkov, M. & Krioukov, D. Internet mapping: from art to science. In *2009 Cybersecurity Applications Technology Conf. for Homeland Security* 205–211 (IEEE, New York, NY, 2009).
25. *Openflights Network Dataset* (The Koblenz Network Collection, 2016); <http://konect.uni-koblenz.de/networks/openflights>
26. Kunegis, J. KONECT—The Koblenz Network Collection. In *Proc. Int. Conf. on World Wide Web Companion* (eds Scwhabe, D., Almeida, V. & Glaser, H.) 1343–1350 (ACM, New York, NY, 2013).
27. Serrano, M. A., Boguna, M. & Sagues, F. Uncovering the hidden geometry behind metabolic networks. *Mol. BioSyst.* **8**, 843–850 (2012).
28. Rolland, T. et al. A proteome-scale map of the human interactome network. *Cell* **159**, 1212–1226 (2014).
29. Takemura, S.-y et al. A visual motion detection circuit suggested by drosophila connectomics. *Nature* **500**, 175–181 (2013).
30. Klimt, B. & Yang, Y. The Enron Corpus: A new dataset for email classification research. In *Machine Learning: ECML 2004* (eds Boulicaut, J. F. et al.) 217–226 (Springer, Berlin, Heidelberg, 2004).
31. Leskovec, J., Lang, K. J., Dasgupta, A. & Mahoney, M. W. Community structure in large networks: natural cluster sizes and the absence of large well-defined clusters. *Internet Math.* **6**, 29–123 (2009).
32. Serrà, J., Corral, A., Boguñá, M., Haro, M. & Arcos, J. L. Measuring the evolution of contemporary western popular music. *Sci. Rep.* **2**, 521 (2012).
33. Milo, R. et al. Superfamilies of evolved and designed networks. *Science* **303**, 1538–1542 (2004).
34. Papadopoulos, F., Aldecoa, R. & Krioukov, D. Network geometry inference using common neighbors. *Phys. Rev. E* **92**, 022807 (2015).
35. Arenas, A., Fernández, A. & Gómez, S. Analysis of the structure of complex networks at different resolution levels. *New J. Phys.* **10**, 053039 (2008).
36. Ronhovde, P. & Nussinov, Z. Multiresolution community detection for megascale networks by information-based replica correlations. *Phys. Rev. E* **80**, 016109 (2009).
37. Ahn, Y.-Y., Bagrow, J. P. & Lehmann, S. Link communities reveal multiscale complexity in networks. *Nature* **466**, 761–764 (2010).
38. Garca-Pérez, G., Boguñá, M., Allard, A. & Serrano, M. A. The hidden hyperbolic geometry of international trade: World Trade Atlas 1870–2013. *Sci. Rep.* **6**, 33441 (2016).
39. Miegheem, P. V. *Graph Spectra for Complex Networks* (Cambridge Univ. Press, New York, NY, 2011).
40. Papadopoulos, F., Psounis, K. & Govindan, R. Performance preserving topological downscaling of Internet-like networks. *IEEE J. Sel. Area Commun.* **24**, 2313–2326 (2006).
41. Papadopoulos, F. & Psounis, K. Efficient identification of uncongested internet links for topology downscaling. *SIGCOMM Comput. Commun. Rev.* **37**, 39–52 (2007).
42. Yao, W. M. & Fahmy, S. Downscaling network scenarios with denial of service (dos) attacks. In *2008 IEEE Sarnoff Symp.* 1–6 (IEEE, New York, NY, 2008).
43. Yao, W. M. & Fahmy, S. Partitioning network testbed experiments. In *2011 31st Int. Conf. on Distributed Computing Systems* 299–309 (IEEE, New York, NY, 2011).
44. Colizza, V., Pastor-Satorras, R. & Vespignani, A. Reaction-diffusion processes and metapopulation models in heterogeneous networks. *Nat. Phys.* **3**, 276–282 (2007).
45. Karypis, G. & Kumar, V. A fast and high quality multilevel scheme for partitioning irregular graphs. In *Int. Conf. on Parallel Processing* (eds Banerjee, P., Polychronopoulos, C. D. & Gallivan, K. A.) 113–122 (CRC, 1995).
46. Karypis, G. & Kumar, V. A fast and highly quality multilevel scheme for partitioning irregular graphs. *SIAM J. Sci. Comput.* **20**, 359–392 (1999).
47. Abou-Rjeili, A. & Karypis, G. Multilevel algorithms for partitioning power-law graphs. In *IEEE Int. Parallel and Distributed Processing Symp. (IPDPS)* 124–124 (IEEE, New York, NY, 2006).

### Acknowledgements

We acknowledge support from a James S. McDonnell Foundation Scholar Award in Complex Systems, the ICREA Academia prize, funded by the Generalitat de Catalunya, and Ministerio de Economía y Competitividad of Spain projects no. FIS2013-47282-C2-1-P and no. FIS2016-76830-C2-2-P (AEI/FEDER, UE).

### Author contributions

G.G.-P., M.B. and M.Á.S. contributed to the design and implementation of the research, the analysis of the results and the writing of the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41567-018-0072-5>.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Correspondence and requests for materials** should be addressed to M.Á.S.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Methods

**Renormalization flow of the community structure.** To assess how the community structure of the network changes with the RGN flow, we obtained a partition into communities of every layer  $l$ ,  $P^{(l)}$ , using the Louvain method<sup>48</sup>; Fig. 2c shows their modularities  $Q^{(l)}$ . We also defined the partition induced by  $P^{(0)}$  on the original network,  $P^{(l,0)}$ , obtained by considering that two nodes  $i$  and  $j$  of the original network are in the same community in  $P^{(l,0)}$  if and only if the supernodes of  $i$  and  $j$  in layer  $l$  belong to the same community in  $P^{(0)}$ . Both the modularity  $Q^{(l,0)}$  of  $P^{(l,0)}$  and the normalized mutual information  $nMI^{(l,0)}$  between both partitions  $P^{(0)}$  and  $P^{(l,0)}$  are shown in Fig. 2c.

**Connection probability in the  $S^1 / \mathbb{H}^2$  geometric model.** The  $S^1$  model<sup>14</sup> places the nodes of a network into a one-dimensional sphere of radius  $R$  and connects every pair  $i, j$  with probability

$$P_{ij} = \frac{1}{1 + \chi_{ij}^\beta} = \frac{1}{1 + \left(\frac{d_{ij}}{\mu \kappa_i \kappa_j}\right)^\beta} \quad (5)$$

where  $\mu$  controls the average degree of the network,  $\beta$  its clustering, and  $d_{ij} = R\Delta\theta_{ij}$  is the distance between the nodes separated by an angle  $\Delta\theta_{ij}$ ;  $R$  is, without loss of generality, always set to  $N/2\pi$ , where  $N$  is the number of nodes, so that the density of nodes along the circle is equal to 1. The hidden degrees  $\kappa_i$  and  $\kappa_j$  are proportional to the degrees of nodes  $i$  and  $j$ , respectively.

The  $S^1$  model is isomorphic to a purely geometric model, the  $\mathbb{H}^2$  model<sup>17</sup>, in which nodes are placed in a two-dimensional hyperbolic disk of radius:

$$R_{\mathbb{H}^2} = 2\ln\left(\frac{2R}{\mu\kappa_0^2}\right) \quad (6)$$

where  $\kappa_0 = \min\{\kappa_i\}$ . By mapping every mass  $\kappa_i$  to a radial coordinate  $r_i$  according to:

$$r_i = R_{\mathbb{H}^2} - 2\ln\frac{\kappa_i}{\kappa_0} \quad (7)$$

the connection probability, equation (5), becomes

$$P_{ij} = \frac{1}{1 + e^{\frac{\beta}{2}(x_{ij} - R_{\mathbb{H}^2})}} \quad (8)$$

where  $x_{ij} = r_i + r_j + 2\ln\frac{\Delta\theta_{ij}}{2}$  is a good approximation to the hyperbolic distance between two points with coordinates  $(r_i, \theta_i)$  and  $(r_j, \theta_j)$  in the native representation of hyperbolic space. The exact hyperbolic distance  $d_{\mathbb{H}^2}$  is given by the hyperbolic law of cosines:

$$d_{\mathbb{H}^2} = \text{acosh}(\cosh r_i \cosh r_j - \sinh r_i \sinh r_j \cos \Delta\theta_{ij}) \quad (9)$$

**Adjusting the average degree of downscaled network replicas.** To reduce the average degree in a renormalized network to the level of the original network, we apply a pruning of links using the underlying metric model with which the networks in all layers are congruent. The procedure is detailed in this section.

The renormalized network in layer  $l$  has an average degree  $\langle k^{(l)} \rangle$  generally larger (in phase I) from the original network's  $\langle k^{(0)} \rangle$ . Moreover, the new network is congruent with the underlying hidden metric space with a parameter  $\mu^{(l)} = \mu^{(0)}/r^l$  controlling its average degree. The main idea is to decrease the value of  $\mu^{(l)}$  to a new value  $\mu_{\text{new}}^{(l)}$ —which implies that the connection probability of every pair of nodes  $(i, j)$ ,  $P_{ij}^{(l)}$ , decreases to  $P_{ij,\text{new}}^{(l)}$ . We then prune the existing links by keeping them with probability

$$q_{ij}^{(l)} = \frac{P_{ij,\text{new}}^{(l)}}{P_{ij}^{(l)}} \quad (10)$$

Therefore, the probability for a link to exist in the pruned network reads:

$$P\{a_{ij,\text{new}}^{(l)} = 1\} = p_{ij}^{(l)} q_{ij}^{(l)} = P_{ij,\text{new}}^{(l)} \quad (11)$$

whereas the probability for it not to exist is:

$$P\{a_{ij,\text{new}}^{(l)} = 0\} = 1 - p_{ij}^{(l)} + p_{ij}^{(l)}(1 - q_{ij}^{(l)}) = 1 - P_{ij,\text{new}}^{(l)} \quad (12)$$

that is, the pruned network has a lower average degree and is also congruent with the underlying metric space model with the new value of  $\mu_{\text{new}}^{(l)}$ . Hence, we only need to find the right value of  $\mu_{\text{new}}^{(l)}$  so that  $\langle k_{\text{new}}^{(l)} \rangle = \langle k^{(0)} \rangle$ . In the thermodynamic limit, the average degree of an  $S^1$  network is proportional to  $\mu$ , so we could simply set

$$\mu_{\text{new}}^{(l)} = \frac{\langle k^{(0)} \rangle}{\langle k^{(l)} \rangle} \mu^{(l)} \quad (13)$$

However, as we consider real-world networks, finite-size effects play an important role. Indeed, we need to correct the value of  $\mu_{\text{new}}^{(l)}$  in equation (13). To this end, we use a correcting factor  $c$ , initially set to  $c = 1$ , and use  $\mu_{\text{new}}^{(l)} = c \frac{\langle k^{(0)} \rangle}{\langle k^{(l)} \rangle} \mu^{(l)}$ ; for every value of  $c$ , we prune the network. If  $\langle k_{\text{new}}^{(l)} \rangle > \langle k^{(0)} \rangle$ , we give  $c$  the new value  $c - 0.1\mu \rightarrow c$ , where  $\mu$  is a random variable uniformly distributed between 0 and 1. Similarly, if  $\langle k_{\text{new}}^{(l)} \rangle < \langle k^{(0)} \rangle$ ,  $c + 0.1\mu \rightarrow c$ . The process ends when  $|\langle k_{\text{new}}^{(l)} \rangle - \langle k^{(0)} \rangle|$  is below a given threshold (in our case, we set it to 0.1).

**Simulation of dynamical processes.** The Ising model is an equilibrium model of interacting spins<sup>49</sup>. Every node  $i$  is assigned a variable  $s_i$  with two possible values  $s_i = \pm 1$ , and the energy of the system is, in the absence of external field, given by the Hamiltonian

$$\mathcal{H} = - \sum_{i < j} J_{ij} a_{ij} s_i s_j \quad (14)$$

where  $a_{ij}$  are the elements of the adjacency matrix and  $J_{ij}$  are coupling constants, which we set to one. We start from an initial condition with  $s_i = 1$  for all  $i$  and explore the ensemble of configurations using the Metropolis-Hastings algorithm: we randomly select one node and propose a change in its spin,  $-s_i \rightarrow s_i$ . If  $\Delta\mathcal{H} \leq 0$ , we accept the change; otherwise, we accept it with probability  $e^{-\Delta\mathcal{H}/T}$ , where  $T$  is the temperature acting as a control parameter. The order parameter is the absolute magnetization per spin  $|m|$ , where  $m = \frac{1}{N} \sum_i s_i$ ; if all spins point in the same direction,  $|m| = 1$ , whereas  $|m| = 0$  if half the spins point in each direction.

In the SIS dynamical model of epidemic spreading<sup>50</sup>, every node  $i$  can present two states at a given time  $t$ , susceptible ( $n_i(t) = 0$ ) or infected ( $n_i(t) = 1$ ). Both infection and recovery are Poisson processes. An infected node recovers with rate  $\lambda$ , whereas infected nodes infect their susceptible neighbours at rate  $\lambda$ . We simulate this process using the continuous-time Gillespie algorithm with all nodes initially infected. The order parameter is the prevalence or fraction of infected nodes  $\rho(t) = \frac{1}{N} \sum_i n_i(t)$ .

The Kuramoto model is a dynamical model for coupled oscillators. Every node  $i$  is described by a natural frequency  $\omega_i$  and a time-dependent phase  $\theta_i(t)$ . A node's phase evolves according to:

$$\dot{\theta}_i = \omega_i + \sigma \sum_{i < j} a_{ij} \sin(\theta_j(t) - \theta_i(t)) \quad (15)$$

where  $a_{ij}$  are the adjacency matrix elements and  $\sigma$  is the coupling strength. We integrate the equations of motion using Heun's method. Initially, the phases  $\theta_i(0)$  and the frequencies  $\omega_i$  are randomly drawn from the uniform distributions  $U(-\pi, \pi)$  and  $U(-1/2, 1/2)$  respectively, as in ref. <sup>51</sup>. The order parameter  $r(t) = \frac{1}{N} \left| \sum_i e^{i\theta_i(t)} \right|$  measures the phase coherence of the set of nodes; if all nodes oscillate in phase,  $r(t) = 1$ , whereas  $r(t) \rightarrow 0$  if nodes oscillate in a disordered manner.

In every realization, we compute an average of the order parameter in the stationary state. In the case of the SIS model, the single-realization mean of prevalence values is weighted by time. The curves presented in this work correspond to statistics over 100 realizations.

**Multiscale navigation.** Given a network and its embedding (layer 0), we merge pairs of consecutive nodes only if they are connected, which guarantees navigation inside supernodes; this process generates layer 1. We repeat the process to generate  $L$  renormalized layers. The multiscale navigation protocol requires every node  $i$  to be provided with the following local information:

1. The coordinates  $(r_i^{(l)}, \theta_i^{(l)})$  of node  $i$  in every layer  $l$ .
2. The list of (super)neighbours of node  $i$  in every layer as well as their coordinates.
3. Let  $\text{SuperN}(i, l)$  be the supernode to which  $i$  belongs in layer  $l$ . If  $\text{SuperN}(i, l)$  is connected to  $\text{SuperN}(k, l)$  in layer  $l$ , at least one of the (super) nodes in layer  $l-1$  belonging to  $\text{SuperN}(i, l)$  must be connected to at least one of the (super)nodes in layer  $l-1$  belonging to  $\text{SuperN}(k, l)$ ; such node is called 'gateway'. For every superneighbour of node  $\text{SuperN}(i, l)$  in layer  $l$ , node  $i$  knows which (super)node or (super)nodes in layer  $l-1$  are gateways reaching it. Notice that both the gateways and  $\text{SuperN}(i, l-1)$  belong to  $\text{SuperN}(i, l)$  in layer  $l$  so, in layer  $l-1$ , they must either be the same (super)node or different but connected (super)nodes.

4. If  $\text{SuperN}(i, l-1)$  is a gateway reaching some supernode  $s$ , at least one of its (super)neighbours in layer  $l-1$  belongs to  $s$ ; node  $i$  knows which.

This information allows us to navigate the network as follows. Let  $j$  be the destination node to which  $i$  wants to forward a message, and let node  $i$  know  $j$ 's coordinates in all  $L$  layers  $(r_j^{(l)}, \theta_j^{(l)})$ . To decide which of its physical neighbours (that is, in layer 0) should be next in the message-forwarding process, node  $i$  must first check if it is connected to  $j$ ; in that case, the decision is clear. If it is not, it must:

1. Find the highest layer  $l_{\text{max}}$  in which  $\text{SuperN}(i, l_{\text{max}})$  and  $\text{SuperN}(j, l_{\text{max}})$  still have different coordinates. Set  $l = l_{\text{max}}$ .
2. Perform a standard step of greedy routing in layer  $l$ : find the closest neighbour of  $\text{SuperN}(i, l)$  to  $\text{SuperN}(j, l)$ . This is the current target  $\text{SuperT}(l)$ .



3. While  $l > 0$ , look into layer  $l - 1$ :

Set  $l = l - 1$ .

If SuperN( $i, l$ ) is a gateway connecting to some (super)node within SuperT( $l + 1$ ), node  $i$  sets as new current target SuperT( $l$ ) its (super)neighbour belonging to SuperT( $l + 1$ ) closest to SuperN( $j, l$ ).

Else node  $i$  sets as new target SuperT( $l$ ) the gateway in SuperN( $i, l + 1$ ) connecting to SuperT( $l + 1$ ) (its (super)neighbour belonging to SuperN( $i, l + 1$ )).

4. In layer  $l = 0$ , SuperT(0) belongs to the real network and she is a neighbour of  $i$ , so node  $i$  forwards the message to SuperT(0).

**Data availability.** The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

## References

48. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Étienne, L. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, P10008 (2008).
49. Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. F. Ising model on networks with an arbitrary distribution of connections. *Phys. Rev. E* **66**, 016104 (2002).
50. Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200–3203 (2001).
51. Moreno, Y. & Pacheco, A. F. Synchronization of kuramoto oscillators in scale-free networks. *Europhys. Lett.* **68**, 603 (2004).