Characterizing the Analogy Between Hyperbolic Embedding and Community Structure of Complex Networks

Ali Faqeeh, Saeed Osat, and Filippo Radicchi

Phys. Rev. Lett. **121**, 098301 — Published 29 August 2018

DOI: 10.1103/PhysRevLett.121.098301

Characterizing the analogy between hyperbolic embedding and community structure of complex networks

Ali Faqeeh,[1,2] Saeed Osat,[3] and Filippo Radicchi[2]

[1]*MACSI, Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland*
[2]*Center for Complex Networks and Systems Research, School of Informatics, Computing, and Engineering, Indiana University, Bloomington, Indiana 47408, USA*
[3]*Quantum Complexity Science Initiative, Skolkovo Institute of Science and Technology, Skoltech Building 3, Moscow, 143026, Russia*

We show that the community structure of a network can be used as a coarse version of its embedding in a hidden space with hyperbolic geometry. The finding emerges from a systematic analysis of several real-world and synthetic networks. We take advantage of the analogy for reinterpreting results originally obtained through network hyperbolic embedding in terms of community structure only. First, we show that the robustness of a multiplex network can be controlled by tuning the correlation between the community structures across different layers. Second, we deploy an efficient greedy protocol for network navigability that makes use of routing tables based on community structure.

A wealth of recent publications provides evidence of the advantages that may arise from thinking of real-world networks as instances of random network models embedded in hidden metric spaces [1, 2]. In this class of models, every node is represented by coordinates that identify its position in the underlying space, and the distance between pairs of nodes determines their likelihood of being connected. The most popular formulation of spatially embedded network models relies on hyperbolic geometry [3, 4]. Hyperbolic network geometry emerges spontaneously from models of growing simplicial complexes [5]. Hyperbolic geometry appears the natural choice for networks with broad degree distributions, under the hypothesis that the generating mechanism for edges in the network is a compromise between popularity of individual nodes and similarity among pairs of nodes [6]. Popularity is represented by the radial coordinate of nodes in the hyperbolic space, while similarity is accounted for by the difference between angular coordinates of pairs of nodes. Hyperbolic maps are useful in practical contexts, as generating efficient routing protocols in information networks [7], characterizing hierarchical organization of biochemical pathways in cellular networks [8], and monitoring the evolution of the international trade network [9]. However, thinking of networks as embedded in the hyperbolic space is important from the theoretical point of view too. Growing network models that rely on hyperbolic geometry provide a genuine explanation for the emergence of power-law degree distributions from local optimization principles only [6]. Further, recent work show that the main features of the percolation transition in multiplex networks can be predicted by simply accounting for inter-layer correlation among hyperbolic coordinates of nodes [10, 11].

Popularity and similarity are core features of models used in network hyperbolic embedding. They are, however, central in another heavily used model in network science: the degree-corrected stochastic block model (SBM) [12]. The SBM assumes a hidden cluster structure where nodes are divided into a certain number of groups. This classification accounts for similarity, as pairs of nodes have different likelihoods of being connected depending on their group memberships. The degree correction provides instead a natural way of accounting for the popularity of the individual nodes. The SBM is generally considered in the context of graph clustering, representing a generative network model with built-in mesoscopic structure [13]. The SBM is used in the formulation of principled community detection methods [14]. These methods, in turn, are equivalent to other well-established techniques for community detection, giving therefore to the SBM a central role in the graph clustering business [15].

At least superficially, the analogy between the ideas of hyperbolic embedding and community structure is apparent. In a recent paper, Wang *et al.* showed that information about community structure can be used to improve accuracy and efficiency of standard algorithms for hyperbolic embedding [16]. Also, previous work was devoted to the development of network models embedded in hyperbolic geometry with the addition of a pre-imposed community structure [17–19]. We are not aware, however, of previous attempts to investigate the theoretical and practical similarity of the two approaches when applied independently to the same network topology. This is the purpose of the present paper.

We assume that the topology of an undirected and unweighted network G with N nodes is fully specified by its adjacency matrix A, whose element $A_{i,j} = A_{j,i} = 1$ if a connection between nodes i and j is present, or $A_{i,j} = A_{j,i} = 0$, otherwise. The hyperbolic embedding of the network G consists in a pair of coordinates (r_i, θ_i) for every node $i \in G$. The quantity r_i is the radial coordinate of node i; θ_i is its angular coordinate. We assume that this information is at our disposal. The way we acquire such a knowledge depends on whether the network analyzed is synthetic or real. For synthetic graphs, we consider single instances of the popularity-similarity optimization model (PSOM) [6], so that hyperbolic coordinates correspond to ground-truth values of the model. We analyze also several real networks, where coordinates of nodes are obtained by fitting graphs against the PSOM. In this second scenario, we either rely on embeddings publicly available [10, 20] or we apply publicly available algorithms to the graphs [20]. Details are provided in [21]. We remark that the PSOM is the model of reference in most of the hyperbolic embedding techniques. It assumes the existence of an underlying hyperbolic space, and consists in a random growing network model where nodes are connected depending on their distance, and the value of other model parameters, such as average degree $\langle k \rangle$, exponent γ of the power-law degree distribution $P(k) \sim k^{-\gamma}$, and temperature T. When a real network is fitted against the PSOM, the parameters $\langle k \rangle$ and γ of the model are determined on the basis of the observed network, while T is treated as a free parameter [20]. Its value may be set to the one that yields the best match between theoretical and numerical results for the distance dependent connection probability [38]; when hyperbolic embedding is used in greedy routing, one may look for the T value that results in the highest success rate [20]. The radial coordinate r_i of every node i is uniquely identified by its degree k_i, hence r_i doesn't require to be truly learned. The angular coordinate θ_i for every node $i \in G$ is instead treated as a fitting parameter. There are various techniques to perform the fit, including approximated optimization algorithms [20, 38], and *ad-hoc* heuristic methods [39, 40].

In our analysis, we further assume to know the community structure of the graph G, consisting in a flat partition of the network into C total communities, where every node $i \in G$ is associated with a discrete-valued coordinate $\sigma_i = 1, \ldots, C$. Algorithms for community detection are numerous [13]. Here, we rely on results obtained by three popular methods: the Louvain algorithm [41], Infomap [42], and the algorithm by Ronhovde and Nussinov [43]. We remark that, in the degree-corrected SBM, the probability for nodes i and j to be connected is a function of σ_i, σ_j, k_i and k_j. Hence, the graph G can be thought as embedded into a community structure, where every node i is *de facto* represented by the coordinates (k_i, σ_i).

A direct comparison between the hyperbolic embedding and the community structure of the graph G consists in a comparison between the coordinates of the individual nodes in the two representations. Further, as the degree of the nodes trivially matches in both representations, the comparison reduces only in matching angular coordinates θs and group memberships σs. From the numerous empirical tests we conducted on both real and synthetic networks, two main conclusions emerge. First, networks usually considered in hyperbolic embedding applications are highly modular, in the sense that partitions found by community detection algorithms correspond to very large values of the modularity function Q [44] (see Figure 1 and [21]). Second, nodes within the same communities are likely to have similar angular coordinates. We note that this second finding is in line with what already shown in Ref. [16]. To quantify coherence among angular coordinates of nodes within the same community g, we first define the variables ξ_g and ϕ_g with

$$\xi_g \, e^{\mathrm{i}\,\phi_g} = \frac{1}{n_g} \sum_{j=1}^{N} \delta_{\sigma_j, g} \, e^{\mathrm{i}\,\theta_j} \; . \tag{1}$$

$\delta_{x,y} = 1$ if $x = y$ and $\delta_{x,y} = 0$, otherwise. The r.h.s. of Eq. (1) stands for the sums of vectors in the complex plane of the type $e^{\mathrm{i}\,\theta} = \cos(\theta) + \mathrm{i}\sin(\theta)$ of all nodes in group g. The vectorial sum is divided by the community size n_g to obtain an average vector for the community. ϕ_g is the angular coordinate of community g. The module $0 \leq \xi_g \leq 1$ indicates how coherent are the angular coordinates of the nodes within group g. Note that the definition of Eq. (1) resembles the one used for the order parameter of the Kuramoto model [45]. We finally measure the angular coherence of a partition as the weighted average

$$\bar{\xi} = \frac{1}{N} \sum_{g=1}^{C} n_g \xi_g \; . \tag{2}$$

By definition, we have that $0 \leq \bar{\xi} \leq 1$. For all networks considered in our analysis (see Figure 1 and [21]), angular coherence is typically large.

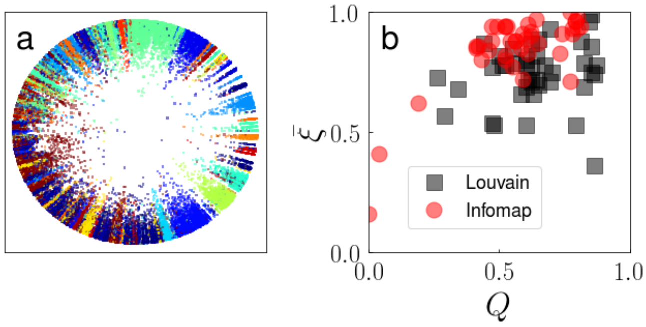

Figure 1. Hyperbolic embedding and community structure for real and synthetic networks. (a) We compare the hyperbolic embedding of the IPv4 Internet with its community structure. Every point represents a node in the largest connected component of the graph. Positions are determined by the radial and angular coordinates of the nodes in the hyperbolic embedding of the network [10]. We use the best partition found by the Louvain algorithm to determine the community structure of the graph [41]. The partition consists of $C = 31$ communities. Colors of the points identify community memberships. The value of the modularity is $Q = 0.61$, while angular coherence is $\bar{\xi} = 0.72$. (b) We consider 39 real-world networks and 2 instances of the PSOM, and compare their community structure and hyperbolic embedding (see details in [21]). The plot displays each network on the $(Q, \bar{\xi})$-plane. We show results obtained using Louvain (black squares) and Infomap (red circles) [42].

Our empirical tests demonstrate that strong angular coherence within communities of strongly modular networks is a quite robust feature of both synthetic and real systems. This finding tells us that the analogy between community structure and hyperbolic embedding may extend beyond the mere similarity among their ingredients. The following examples show that the analogy is useful also in the interpretation of physical properties of networks and the design of practical algorithms on networks.

Our first example regards the rephrasing, in terms of community structure only, of a result obtained by analyzing the hyperbolic embedding of multiplex networks. In two recent papers [10, 11], Kleineberg and collaborators found that inter-layer correlation between hyperbolic coordinates of nodes in multiplex networks is a good predictor for the robustness of a system under targeted attack. Specifically, they found that, when correlation among angular coordinates is high, the percolation transition is smooth. Instead, multiplex networks characterized by a small value of inter-layer correlation exhibit abrupt percolation transitions. The finding was initially obtained for real-world multiplex networks. A theoretical explanation was then given in terms of a synthetic network model [11]. To further support the analogy between hyperbolic embedding and community structure that we are arguing for in this paper, we replicated all results of Ref. [11] using community structure only. First, we analyzed the same real-world multiplex networks considered in Ref. [11]. We found that their robustness can be predicted very well by the level of correlation among the community structures of the layers [21]. Then, we provided a theoretical explanation. We replaced the network model by Kleineberg *et al.* with a variant of the SBM known in the literature as the Lancichinetti-Fortunato-Radicchi (LFR) benchmark graph [46]. The LFR model mostly differs from the standard SBM for relying on heterogeneous distributions of node degrees and community sizes. In our model for multiplex networks [21], we first generate a single LFR graph that is used as the topology for both layers. We then exchange the node labels in one layer to destroy edge overlap and degree-degree correlation. We consider two distinct scenarios. In the first case, we exchange the label of every node with the one of a randomly chosen node from the same community. This allows us to maintain perfect correlation between the community structure of the two layers. In the second case, we exchange the labels of a number of randomly sampled nodes such that the edge overlap between the layers equals the value obtained in the first randomization scheme. This second recipe completely destroys correlation between the community structures of the two layers. In Fig. 2a, we show the phase diagrams for instances of the multiplex model when relabeling uses information about the community structure of the graph. Here, the community structure is strong, in the sense that the fraction of external connections per node is only $\mu = 0.1$. The transition appears smooth, and becomes smoother as the size of the model increases. This is an indication that, in the limit of infinitely large LFR graphs, the percolation transition is likely continuous. In Fig. 2b, we consider the second relabeling scheme that doesn't account for community structure. The resulting diagrams indicate that the percolation transition is abrupt. The level of correlation among community structure of the two layers can be decreased by increasing μ, so that community structure itself becomes less neat. This is done in Figs. 2c and d, where the transition appear abrupt no matter how the labels of the nodes are relabeled. In [21], we report results for different parameter values of the LFR model. Results confirm our claim that the extent of correlation between the community structure of the layers of a multiplex can be used to explain

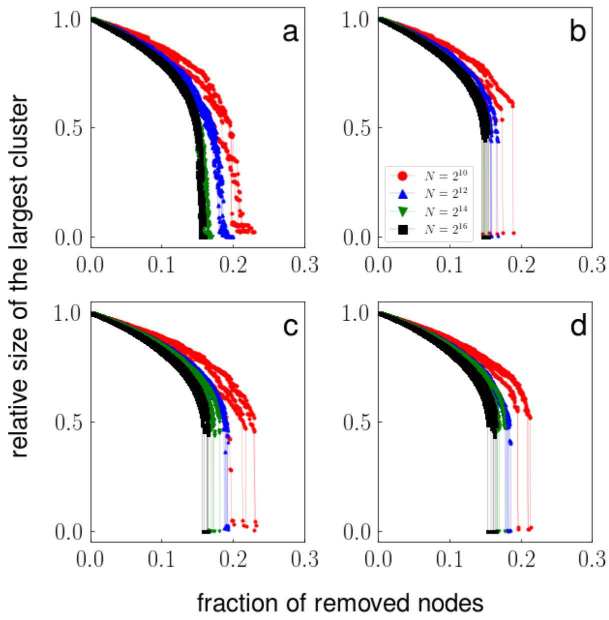

Figure 2. Robustness of multiplex networks with correlated community structure. We measure the relative size of the largest mutually connected cluster as a function of the fraction of nodes removed from the system. The synthetic multiplex graphs are obtained using the recipe described in the text, where two Lancichinetti-Fortunato-Radicchi (LFR) networks with size N are coupled together. The LFR models are such that: the average degree is $\langle k \rangle = 6$; the maximum degree is $k_{max} = \sqrt{N}$; node degrees k obey a power-law distribution $P(k) \sim k^{-\gamma}$ with exponent $\gamma = 2.6$; there are $C = \sqrt{N}$ communities of identical size $S = \sqrt{N}$. For every N, we show the results for five distinct instances of the model. (a) LFR graphs are generated with $\mu = 0.1$. Labels are exchanged only among nodes within the same clusters. All nodes are considered for relabeling at least once. (b) Same as in panel a. However, relabeling of nodes is not constrained by community structure. The number of nodes that are relabeled is such that the edge overlap among layers is the same as in panel a [21]. (c and d) Same as in panels a and b, respectively, but for LFR graphs constructed using $\mu = 0.3$.

robustness properties of the system under targeted attack.

Our second example focuses on greedy routing [2, 7]. To be brief, the scenario considered is the following. A packet originated by node s must be delivered to node t. The packet can navigate the network by walking at each step on an edge. The packet moves on the network till it reaches its destination t, or it visits twice the same node. In the first case, the packet is correctly delivered. In the second case, the packet is considered lost, and it is discarded. The goal of a good routing strategy is to deliver packets with high probability and with a small number of steps, for any randomly chosen pair of source and target nodes s and t. Hyperbolic embedding turns out to be very useful in the formulation of a greedy strategy, where individual steps are determined on the basis of the distance among nodes in the hyperbolic space. Specifically, if a message is at node i, then the next move will be on the node

$$j^{(i)}_{(best)} = \arg \min_{j \in \mathcal{N}_i} d(j,t) \;, \tag{3}$$

where $\mathcal{N}_i$ is the set of neighbors of i, and $d(j,t)$ is the distance between nodes j and t. The greedy technique is computationally feasible as every node needs to know only the identity and the geometric coordinates of its neighbors. The regimes of effectiveness of the routing method have been systematically studied in artificial network models [2]. The technique has been proven to be extremely effective on some real-world topologies [2, 7]. We devised a new greedy routing protocol that makes use of the cluster structure of a network instead of its hyperbolic embedding. Specifically, we replaced the definition of distance in the hyperbolic space between nodes with the fitness

$$d(j,t) = \beta D_{\sigma_j,\sigma_t} - (1-\beta) \ln k_j \;, \tag{4}$$

where k_j is the degree of node j, and σ_j and σ_t are the indices of the communities of nodes j and t, respectively. D_{σ_j,σ_t} is the length of the shortest path between communities σ_j and σ_t calculated on a weighted supernetwork in which supernodes are communities of the original network. Each pair of supernodes g and q is connected with a superedge with weight $1 - \ln \rho_{g,q}$; here $\rho_{g,q}$ is the probability that, in the original network, a randomly chosen node in community g has an edge to community q [21]. The term $\ln k_j$ in Eq. (4) serves to perform degree correction. The factor $0 \leq \beta \leq 1$ serves to control the relative importance of one factor over the other. β plays a similar role as of the temperature T in hyperbolic routing protocols [20], and its value may be appropriately chosen with the goal of optimizing the success rate in the delivery of messages [21]. The routing protocol based on Eq. (4) is still computationally efficient as long as the total number of communities C grows sub-linearly with the size of the graph N. In the extreme case, where every community is formed by a single node, so that $C = N$, the method will be 100% accurate in delivering packets, but also computationally expensive. In Figure 3, we display the performance of community-based greedy routing as a function of the mean size of the communities. We study the performance on both synthetic and real-world networks. The number of communities is tuned by changing the resolution parameter in the algorithm by Ronhovde and Nussinov [43]. Success rates of the community-based greedy protocol are always very good, as long as communities are not too large.

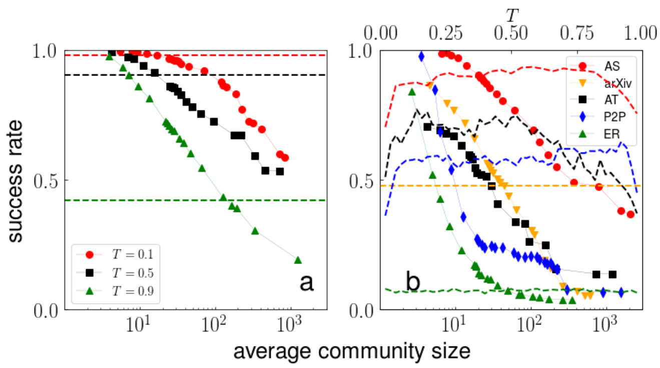

Figure 3. Performance of community-based routing. (a) We consider single instances of the growing network model of Ref. [38] with $N = 5{,}000$ nodes, $\langle k \rangle = 5$, and degree exponent $\gamma = 2.1$. Different symbols and colors refer to different values of the temperature T. The plot shows how success rate of the community-based greedy routing strategy changes as a function the average size of the communities. Communities are identified using the algorithm by Ronhovde and Nussinov [43]. Their number can be varied by changing the resolution level of the algorithm. Dashed lines are obtained on the same networks but using hyperbolic greedy routing. (b) Same as in panel a, but for real networks. We consider the following networks: the Internet at the level of autonomous systems (AS) [47]; the worldwide air transportation network (AT) [48]; the European road network (ER) [49]; the peer-to-peer network (P2P) [50]; the arXiv collaboration network [51]. For all the networks (except arXiv) the dashed lines are obtained by varying the temperature T in the algorithm for hyperbolic embedding introduced in Ref. [20]; for the arXiv network the dashed line shows the result for the optimum hyperbolic coordinates whose data was available in [10]. Details can be found in [21].

In summary, we showed that looking at a network as embedded in a hyperbolic geometry is similar, both in theory and practice, to pretending that the network is organized into communities, provided that community structure is detected by a method that accounts for the degree of the nodes. Our finding provides evidence that the inter-community structure in networks may have geometric organization, meaning that at the global level, geometry dominates, while at the local scale, community memberships prevail. Thus, real networks may be modeled by a graphon [52] consisting of a mixture of latent-spatial and block-like structures. This fundamental model has the potential to generate further understanding of physical processes, such as spreading and synchronization, in real networks.

The authors thank G. Bianconi, C. V. Cannistraci, D. Krioukov , and M.Á. Serrano for comments on the manuscript. A.F. and F.R. acknowledge support from the U.S. Army Research Office (W911NF-16-1-0104). F.R. acknowledges support from the National Science Foundation (CMMI-1552487).

[1] M. Á. Serrano, D. Krioukov, and M. Boguná, Phys. Rev. Lett. **100**, 078701 (2008).
[2] M. Boguna, D. Krioukov, and K. C. Claffy, Nat. Phys. **5**, 74 (2009).
[3] D. Krioukov, F. Papadopoulos, A. Vahdat, and M. Boguná, Phys. Rev. E **80**, 035101 (2009).
[4] D. Krioukov, F. Papadopoulos, M. Kitsak, A. Vahdat, and M. Boguná, Phys. Rev. E **82**, 036106 (2010).
[5] G. Bianconi and C. Rahmede, Sci. Rep. **7** (2017).
[6] F. Papadopoulos, M. Kitsak, M. Á. Serrano, M. Boguná, and D. Krioukov, Nature **489**, 537 (2012).
[7] M. Boguná, F. Papadopoulos, and D. Krioukov, Nat. Commun. **1**, 62 (2010).

[8] M. Á. Serrano, M. Boguñá, and F. Sagués, Mol. BioSyst. **8**, 843 (2012).
[9] G. García-Pérez, M. Boguñá, A. Allard, and M. Á. Serrano, Sci. Rep. **6**, 33441 (2016).
[10] K.-K. Kleineberg, M. Boguná, M. Á. Serrano, and F. Papadopoulos, Nat. Phys. **12**, 1076 (2016).
[11] K.-K. Kleineberg, L. Buzna, F. Papadopoulos, M. Boguñá, and M. Á. Serrano, Phys. Rev. Lett. **118**, 218301 (2017).
[12] B. Karrer and M. E. Newman, Phys Rev. E **83**, 016107 (2011).
[13] S. Fortunato, Phys. Rep. **486**, 75 (2010).
[14] T. P. Peixoto, arXiv preprint arXiv:1705.10225 (2017).
[15] M. Newman, Phys. Rev. E **94**, 052315 (2016).
[16] Z. Wang, Q. Li, F. Jin, W. Xiong, and Y. Wu, Physica A **455**, 104 (2016).
[17] K. Zuev, M. Boguñá, G. Bianconi, and D. Krioukov, Sci. Rep. **5**, 9421 (2015).
[18] G. García-Pérez, M.Á. Serrano, and M. Boguñá, arXiv preprint arXiv:1707.09610 (2017).
[19] A. Muscoloni and C. V. Cannistraci, arXiv preprint arXiv:1707.07325 (2017).
[20] F. Papadopoulos, R. Aldecoa, and D. Krioukov, Phys. Rev. E **92**, 022807 (2015).
[21] See Supplemental Material [url] for the analysis of additional real and synthetic networks, which includes Refs. [22-37].
[22] G. García-Pérez, M. Boguñá, and M. Á. Serrano, Nat. Phys. **14**, 583 (2018).
[23] B. L. Chen, D. H. Hall, and D. B. Chklovskii, Proc. Natl. Acad. Sci. USA **103**, 4723 (2006).
[24] C. Stark, B.-J. Breitkreutz, T. Reguly, L. Boucher, A. Breitkreutz, and M. Tyers, Nucleic Acids Res. **34**, D535 (2006).
[25] M. De Domenico, V. Nicosia, A. Arenas, and V. Latora, Nat. Commun. **6**, 6864 (2015).
[26] M. De Domenico, A. Lancichinetti, A. Arenas, and M. Rosvall, Phys. Rev. X **5**, 011027 (2015).
[27] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney, Internet Math. **6**, 29 (2009).
[28] J.Serrà, Á. Corral, M. Boguñá, M.Haro, and J.L.Arcos, Sci. Rep. **2**, 521 (2012).
[29] J. Kunegis, in *Proceedings of the 22nd International Conference on World Wide Web* (ACM, 2013) pp. 13431350.
[30] T. Rolland, M. Tas an, B. Charloteaux, S. J. Pevzner, Q. Zhong, N. Sahni, S. Yi, I. Lemmens, C. Fontanillo, R. Mosca, *et al.* Cell **159**, 1212 (2014).
[31] K. Claffy, Y. Hyun, K. Keys, M. Fomenkov, and D. Krioukov, in *2009 Cybersecurity Applications Technology Conference for Homeland Security* (2009) pp. 205–211.
[32] L. Danon, A. Díaz-Guilera, J. Duch, and A. Arenas, J. Stat. Mech., P09008 (2005).
[33] F. Radicchi, Phys. Rev. E **97**, 022316 (2018).
[34] F. Radicchi, Nat. Phys. **11**, 597 (2015).
[35] S. Osat, A. Faqeeh, and F. Radicchi, Nat. Commun. **8**, 1540 (2017).
[36] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, Nature **464**, 1025 (2010).
[37] D. B. Johnson, J. ACM **24**, 1 (1977).
[38] F. Papadopoulos, C. Psomas, and D. Krioukov, IEEE/ACM Trans. Netw. **23**, 198 (2015).
[39] G. Alanis-Lobato, P. Mier, and M. A. Andrade-Navarro, Sci. Rep. **6**, 30108 (2016).
[40] A. Muscoloni, J. M. Thomas, S. Ciucci, G. Bianconi, and C. V. Cannistraci, Nat. Commun. **8**, 1615 (2017).
[41] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, J. Stat. Mech. , P10008 (2008).
[42] M. Rosvall and C. T. Bergstrom, Proc. Natl. Acad. Sci. USA **105**, 1118 (2008).
[43] P. Ronhovde and Z. Nussinov, Phys. Rev. E **81**, 046114 (2010).
[44] M. E. Newman and M. Girvan, Phys. Rev. E **69**, 026113 (2004).
[45] Y. Kuramoto, *Chemical oscillations, waves, and turbulence* (Dover Publications, New York, 1984).
[46] A. Lancichinetti, S. Fortunato, and F. Radicchi, Phys. Rev. E **78**, 046110 (2008).
[47] J. Leskovec, J. Kleinberg, and C. Faloutsos, in *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining* (ACM, 2005) pp. 177–187.
[48] R. Guimerà, S. Mossa, A. Turtschi, and L. A. N. Amaral, Proc. Natl. Acad. Sci. USA **102**, 7794 (2005).
[49] L. Šubelj and M. Bajec, The Eur. Phys. J. B **81**, 353 (2011).
[50] M. Ripeanu and I. Foster, in *Peer-to-Peer Systems*, edited by P. Druschel, F. Kaashoek, and A. Rowstron (Springer Berlin Heidelberg, Berlin, Heidelberg, 2002) pp. 85–93.
[51] M. De Domenico, M. A. Porter, and A. Arenas, J. Compl. Net. **3**, 159 (2015).
[52] L. Lovász, *Large networks and graph limits*, Vol. 60 (American Mathematical Soc., Rhode Island, USA, 2012)
.