PHYSICAL REVIEW RESEARCH **4**, 033185 (2022)

Functional rich clubs emerging from the diffusion geometry of complex networks

Giulia Bertagnolli [*]
Department of Mathematics, University of Trento, Via Sommarive, 14, 38123 Povo (Trento), Italy
and CoMuNe Lab, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo (Trento), Italy

Manlio De Domenico [†]
Department of Physics and Astronomy "Galileo Galilei," University of Padua, Padova, Italy
and CoMuNe Lab, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo (Trento), Italy

(Received 26 May 2022; accepted 8 August 2022; published 6 September 2022)

Real systems are characterized by complex patterns of interactions between their units, by dynamical processes on them, and by the interplay of the two. It is well known that particular structures affect dynamical processes at different scales. Sometimes richly connected units are connected by costly, long-range links. In the brain, hubs form rich clubs for integrating information between different brain regions, and many biological and social networks show this same structural organization. It remains, however, unclear whether this structural organization alone enables a rapid communication between highly connected nodes or whether a functional rich club may emerge as a combination of direct links and longer paths between rich nodes. Here, we identify functional rich clubs through the diffusion geometry, providing a perspective on rich-club phenomena in complex networks. We show that weak structural rich clubs may be functionally stronger, thanks to bridge nodes, while diffusion inside strong structural rich clubs may be damped in modular networks.

DOI: 10.1103/PhysRevResearch.4.033185

I. INTRODUCTION

Networks have long been studied from a structural perspective, investigating, for instance, how the existence probability of each edge—a microscale property of the system—affects macroscale or global properties of the network, such as the average degree and the connectedness probability. Similarly, the mesoscale structure can be characterized in terms of the number of edges in particular subgraphs, be they communities, i.e., groups of nodes that are more densely connected with each other than with the rest of the network, or a rich club, i.e., a group of high-degree nodes (called rich nodes) sharing more links with each other than are expected by chance [1,2].

Abstract as they may seem, networks represent real complex systems. Each abstract link is an interaction between the system's elements; each interaction is an exchange of information—in terms of a signal, or of a flow of energy or matter—and this pairwise communication gives rise to the system's behavior and functionality. How microscale interactions shape macroscale phenomena is the fundamental question of network and complexity science [3]: When observing a dynamical process taking place on the network's nodes and links, a natural question is how the topology of the network affects the evolution of the dynamics. For instance, do different distributions of the number of connections of each node (i.e., the degree distribution) alter the spread of information or diseases in a networked system? Of course, the answer is process dependent, and the pair, structure and dynamics, cannot—and should not—be decoupled.

In the last decades, an approach, complementing the purely structural one, has been gaining more and more attention: Through the interplay between the structure of networks and the dynamical processes on them, researchers are studying the functional shape of networks [4–8]. In this framework, a particular organization of nodes corresponds to a pattern in the functional response of the system at a certain time or during a time interval. A functional community, for instance, consists of a group of nodes displaying a more or less homogeneous behavior with respect to the considered dynamics. This approach is intrinsically multiscale [9,10] and has the twofold advantage of allowing greater fine-tuning of analyses and of overcoming common issues, such as, e.g., the resolution limit in community detection [11–13]. While community detection [14] exploiting different dynamics—random walks [15,16] and diffusion [6], synchronization [17], and spin dynamics [18,19], to cite a few—is already the state of the art, the aforementioned rich-club phenomenon has not yet been studied from a dynamical perspective in spite of the huge interest in this mesoscale organization of nodes. There is a large literature on the subject of brain networks, where the (structural) rich club is known to be relevant for the integration of the information processed by physically segregated modules [20–25], and it may also be at the base of

---

[*] Corresponding author: giulia.bertagnolli@unitn.it
[†] Corresponding author: manlio.dedomenico@unipd.it

those cognitive capacities distinguishing humans from other animal species [26]. Moreover, an abnormal rich-club ordering has been found in patients suffering from several diseases [27–32]. In a recent work [33], Moretti and Hütt studied the role played by hubs, cores, and rich clubs in self-organized wave and excitation patterns, finding that these act as organizers and facilitators for the emergence of the global, collective patterns both in network models and in the human connectome. Social networks have rich clubs too. Resources, for example, are usually not homogeneously distributed, and a subset of prominent individuals controls a large part of all the resources [34–36]; scientific collaborations, in terms of the coauthorship of papers, tend to be stronger between renowned scientists [37]; and face-to-face interactions occurring over time are more frequent and more stable than expected by chance among individuals who have more interactions [38]. Then, social interactions shape our infrastructures, and hence it is not surprising that this “the-rich-get-richer” effect and rich clubs can be found, for instance, in urban networks [39]. Among the real systems lacking a structural rich-club organization there are protein-protein interaction networks, where it is claimed [37,40] that hub proteins are mainly devoted to very specific functions.

The structural rich-club phenomenon was originally quantified through the spectrum of the rich-club coefficient [2], the density of connections among nodes with degree larger than k:

$$\phi(k) = \frac{2E_{>k}}{N_{>k}(N_{>k}-1)}, \tag{1}$$

where $N_{>k}$ and $E_{>k}$ are the number of nodes with degree larger than k and the number of edges between them, respectively. Later, it was shown [37] that a proper normalization of $\phi(k)$ is mandatory, since $\phi(k)$ alone does not allow the effect of degree-degree correlations to be filtered out. Normalized descriptors of the rich-club phenomenon were introduced in Refs. [37,41], dividing the coefficient in (1) by the average rich-club coefficient of the uncorrelated or the maximally random network with the same degree sequence of the network under study:

$$\rho_{\text{unc}}(k) = \frac{\phi(k)}{\phi_{\text{unc}}(k)}, \tag{2}$$

$$\rho_{\text{ran}}(k) = \frac{\phi(k)}{\phi_{\text{ran}}(k)}. \tag{3}$$

There are other possible formulations of the normalized structural rich club [42] and of its generalization to weighted [34,43–45], hierarchical [46], and temporal [38] networks, but the key point pooling all these descriptors together is the need to distinguish the case where nodes with a lot of (or strong) connections have more links between them just by chance from the case in which hubs have, indeed, an intense connectivity giving them, e.g., control over resources flowing in the system or facilitating the rapid exchange of information among them.

This last interpretation of the rich club raises a question: Does a topological, or structural, rich-club organization translate seamlessly into function? Vice versa, does a functional rich-club organization always need its structural counterpart? Assuming that *communication*—the exchange of information or signals—is the basic process enabling the emergence of complex system behaviors, e.g., the physiological activity of a cell or socioeconomic activity depending on trade networks, and that communication takes various forms [47,48]—from routing information along shortest paths to broadcasting—it is clear that we need other approaches complementing the structural one to understand the organization of complex systems and, eventually, their functional principles. To answer these questions and to bring us a step closer to understanding the structural-functional interplay in complex systems, we tackle the rich-club phenomenon from the functional perspective of diffusion geometry [6,8,49,50]. As a communication strategy, diffusion lies at the opposite end of the spectrum from routing, but still it cannot represent the whole communication spectrum. Nevertheless, our definition of the functional rich club has two major advantages: Firstly, random walks and diffusion are good approximations of a bunch of different dynamics, from metastable synchronization to consensus [6], so that in applications they may be “general enough.” Secondly, the fact that our approach is based on the metric induced by random walks on the network makes it easily generalizable to other process-driven geometries, when available.

II. NUMERICAL EXPERIMENTS AND DATA ANALYSIS

A. Defining the functional rich club of complex networks

We propose a functional rich club which is based on the idea that a rapid information exchange among rich nodes is facilitated not only by the presence of many, direct links between them, but also by the global connectivity of the network. For instance, bridge nodes with low degree placed between hubs may be overlooked in the structural case, whereas using an averaging process, such as the diffusion, we can account also for longer paths. Furthermore, a small set of rich nodes may form a strong structural rich club, regardless of the presence or absence of other mesoscale structures. Instead, in the functional setting these structures may be competing at some scale, leading to the lack of a functional rich-club organization.

The diffusion geometry framework [6] exploits continuous-time random walks to map the set of nodes V of a network $G = (V, E)$ to a cloud of points in space, whose distribution depends on the connectivity E and, possibly, on the information flow encoded in the edge weights and on the type of random walk dynamics [8]. The Euclidean distance between these points induces, in turn, a metric on G, called the diffusion distance D_t. Intuitively, two nodes $i, j \in V$ are near with respect to D_t if the transition probability from i to j is high; see Sec. III for more details. The time parameter t acts as a scale parameter; hence our family of diffusion distances $\{D_t\}_{t>0}$ provides a geometric, multiscale tool for uncovering the functional rich-club organization of networks.

Studying the distribution of the mutual diffusion distances, we are able to tell whether rich nodes are indeed nearer to each other than expected by chance, accounting also for the rest of the nodes in the network. The functional rich-club coefficient is defined as a Z score (see Sec. III) and measures the average diffusion distance between rich nodes, sometimes also indicated as the average core distance, in units of standard

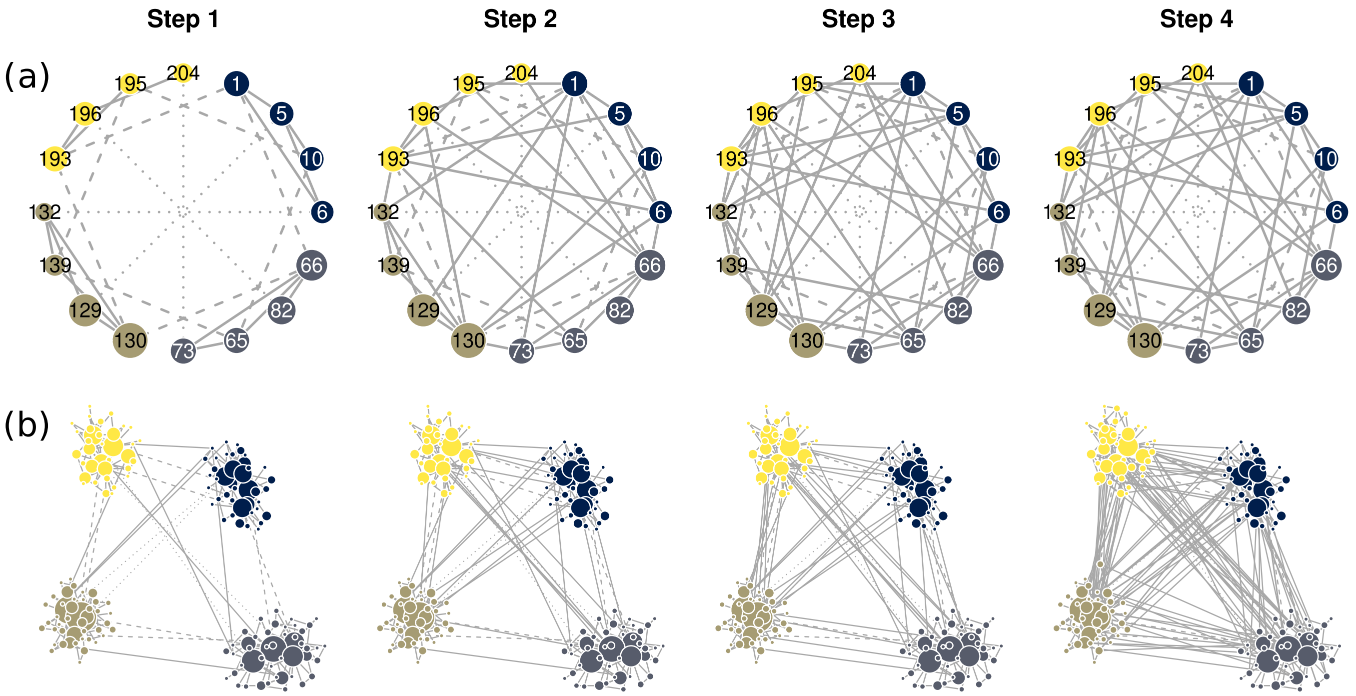

FIG. 1. Model of a functional rich-club organization: moving from weakly connected subnetworks to more densely connected rich and mildly rich nodes. A total of 256 nodes are grouped into four equally sized Barabási-Albert (BA) communities highlighted with different colors in (a)—where only the four richest nodes in each community are shown—and in (b)—where their first neighbors are added. At step 1, intercommunity links are added, at random between low-degree nodes and in "rings," dashed edges, and "crosses," dotted edges, between high-degree nodes with the same rank (see Sec. III C for more details). In steps 2 and 3 the connectivity between rich nodes is increased, adding more links between rich nodes of different rank. In step 4, mildly rich nodes with a degree $k \in [4, 9]$ are connected to nodes with a degree larger than $k = 9$. The average degree of G at the different steps varies from $\langle k \rangle \approx 4.2$ to $\langle k \rangle \approx 4.9$.

deviations from the mean of the null model. In other words, we state that a network has a functional rich club if the flow of information between nodes with high degree is significantly easier compared with the whole network and with a null model, which we choose to be a sample from the configuration model [51] with the same degree sequence as G.

Figure 1 is a summary of a simple construction of a synthetic network with a functional rich club, and Fig. 2 is a summary of its functional and structural rich-club organizations. For the structural part, we show both the Z score of the original rich-club coefficient $\phi(k)$, indicated as $\phi^Z(k)$ (see Sec. III), and the rich-club ratio $\rho(k)$ of (2). A detailed description of the model and its building steps can be found in Sec. III. The network has a very strong organization of its nodes in four equally sized ($\frac{N}{4} = 64$) communities. At the beginning these communities are very weakly connected, through random edges between low-degree nodes and through regular connection patterns between the four richest nodes in each group, referred to henceforth as the rich nodes. This is enough for observing a (weakly significant) structural rich club for $k > 15$, according to $\rho(k)$ in Fig. 2(c), while, functionally, the average distance between rich nodes is significantly larger compared with the configuration model, especially for $4 < t < \frac{N}{4}$. Adding more links between the rich nodes strengthens the structural rich club and brings the rich nodes nearer in the diffusion space at small time scales, i.e., $t < 4 \approx \langle k \rangle$, the average network degree. Finally, a functional rich club for $k \geqslant 25$ emerges when we connect rich and mildly rich nodes. It is also interesting to observe that for small time scales, say, $t < 5$, the structural and functional Z scores display a similar but not identical behavior. Thinking about the embedded chain of our continuous-time random walk, i.e., a classical discrete-time random walk on the network, one observes that for $t = 1$ a random walker can only reach the first neighbors of the initial node, and so, one could expect the two rich clubs to be the same. However, a more attentive look reveals a subtle but non-negligible difference: If the structural rich-club coefficient $\phi(k)$ is proportional to $\sum_{k_j > k} k_j$, disregarding of the rest of the network, the average distance between nodes with degree larger than k always accounts for all edges going out from those rich nodes. In particular, if two random walkers starting in two distinct rich nodes both have a nonzero probability of going into a third nonrich node, then this will decrease their reciprocal diffusion distance. On the other side, if two rich nodes are directly linked but have otherwise disjoint first neighborhoods, then the posterior distribution of the random walker being in any node after one step will be almost localized in the two neighborhoods and, consequently, the diffusion distance of the two nodes will be high.

This simple toy network suggests that the global network structure plays an important role in shaping and conditioning the information flow among hubs: If these belong to distinct communities in which they are, having high degree, deeply rooted, despite the presence of direct links between them, the diffusion is hindered, resulting in higher distances in the diffusion space with respect to the null model, where the structures

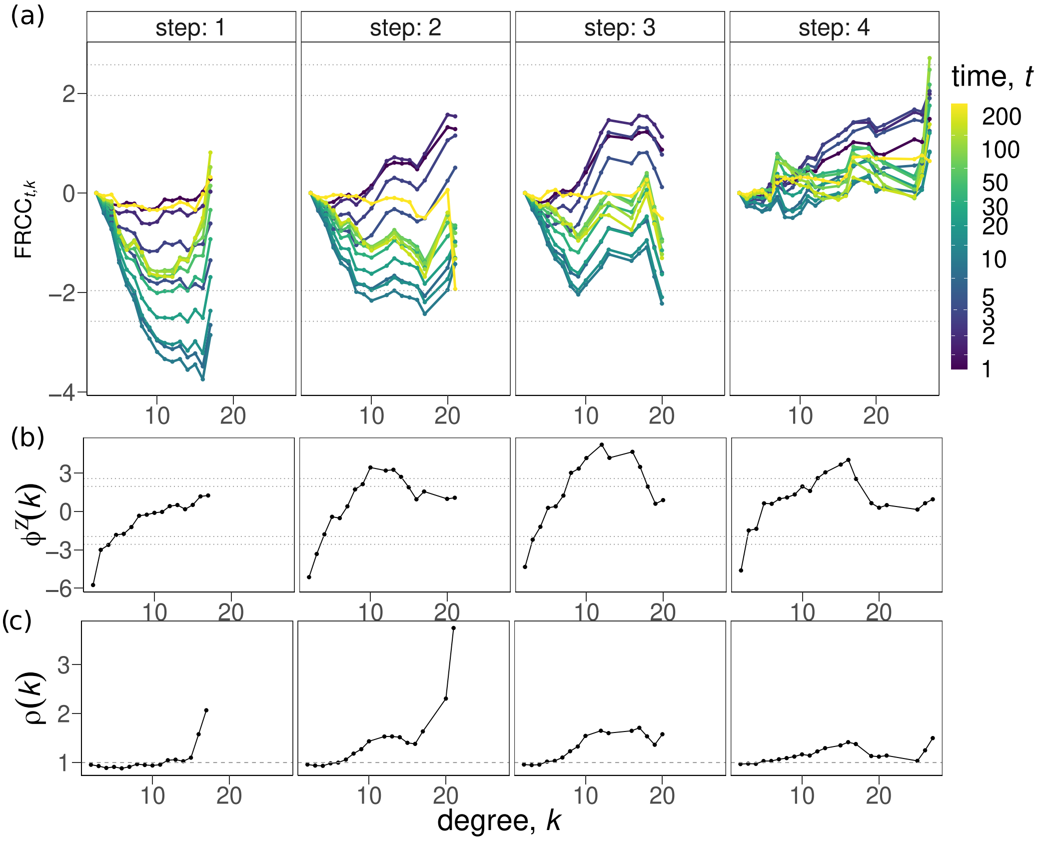

FIG. 2. The emergence of a functional rich club. (a) The functional rich-club coefficient, (b) Z score of the structural rich-club coefficient, and (c) the ratio $\rho(k)$ for the model in Fig. 1 during its building steps. The dashed lines in (a) and (b) enclose 95 and 99% of the standard normal distribution. Points outside these ranges are sufficiently extreme deviations from the mean behavior of the configuration model, which allow as to claim that G has a strong functional (or structural) rich club. Structurally, there is a clear club of mildly rich nodes, i.e., for degree $k \in [8, 12]$, at all steps (only weakly at step 1, according to ϕ^Z), whereas at first the diffusion distances between nodes of degree $k \geqslant 10$ are significantly larger than expected, due to the strong partition into communities of the model. In step 4, adding connections between rich and mildly rich nodes has a negligible effect on the structural rich club for degrees $k \in (10, 20)$, while it considerably changes the functional panorama: A significant functional rich club emerges at high values of k, and the functional shape of the network is more persistent across scales.

are destroyed. To answer our previously asked questions, we have shown that a network can have a significant structural rich club and lack a functional rich club [Fig. 2(b), steps 2 and 3]. Vice versa, in the presence of bridges and many short (even if not direct) paths between hubs, the diffusion among them may be eased sufficiently enough to form a functional rich club [Fig. 2(a), step 4].

Now that we can quantify the functional rich-club phenomenon and have identified some basic mechanisms allowing for its emergence, let us move on to real networks and more realistic network models.

B. Unraveling the functional rich-club organization of empirical networks

From previous studies [34,37], networks representing social and collaboration relationships display the (structural) rich-club phenomenon, and so we look at the results of the functional rich-club analysis on the unweighted 2010 Network Scientists network [52,53], a coauthorship network with 522 nodes, where two scholars are linked if and only if they wrote at least one paper together, which are summarized in Fig. 3. At a local scale $t = 1$—recall that the continuous-time random walk has exponential waiting times with rate 1 [8]—there is no significant functional organization of the rich nodes into a club; see Fig. 3(d), although for $k = 20, 21$ the average distance between nodes of degree of at least 20 or 21 is almost two times smaller than expected by chance (i.e., according to the average of the configuration model). Also, the structural coefficients $\rho(k)$, $\phi_Z(k)$ detect a strong rich club at $k = 19, 20$, as can be seen in Figs. 3(c) and 3(d). As time increases, the diffusion between rich nodes ($k \geqslant 18$) is significantly supported by the presence of longer paths between them, leading to strong evidence of the functional rich-club phenomenon. In Figs. 3(a) and 3(b) we show the structural and the extended functional rich clubs for $k \geqslant 20$, respectively. More specifically, in Fig. 3(b) we fix $t = 2$ and highlight all the nodes in the network which are not farther than the maximum observed distance in the rich club, i.e., $\max(D_t|_{\geqslant 20})$, from each rich node. We call this the metric neighborhood of the rich nodes. In addition to the eight rich nodes, this neighborhood contains 30 other nodes with varying degree, which may belong to small and dense communities of or connecting the rich nodes, such as, for instance, the nodes

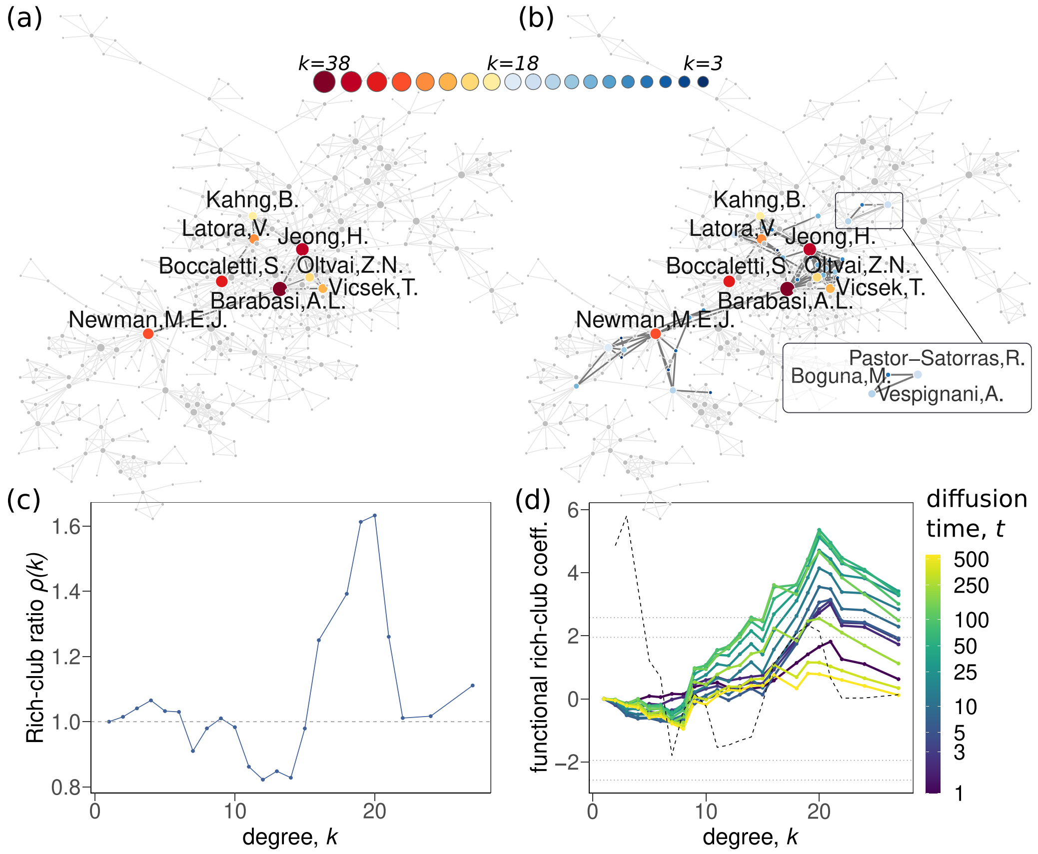

FIG. 3. The 2010 Network Scientists network. (a) Its structural rich club, where nodes with a degree equal to or larger than 20 are colored and labeled. (b) All the nodes which are in the same (metric) neighborhood as the rich nodes. These nodes are not rich themselves, but contribute to the diffusion among rich nodes. (c) The rich-club ratio $\rho(k) = \frac{\phi(k)}{\langle \phi(k)_{\mathrm{cm}} \rangle}$. For $k > 15$ the structural rich-club coefficient is larger than 1, indicating the presence of a rich club, with a peak for $k = 20$, while for larger values of the degree the rich club is not so evident. (d) The functional rich-club coefficient (coeff.) as in (6), as a function of the degree k and diffusion time t; the dashed line corresponds to the Z score of the structural rich club $\phi^Z(k)$. Dashed horizontal lines delimit the ranges $[-1.95, 1.95]$ and $[-2.575, 2.575]$, which enclose 97.5 and 99% of the standard normal distribution, respectively. There is strong evidence of a functional rich club for $k > 15$ at different time scales t. For $k = 20$ the average core diffusion distances are more than three standard deviations smaller than the distances in the cores of the configuration model. Unlike its structural counterpart, the functional rich club is clearly present also for richer nodes, i.e., $k > 20$.

around Newman or the community to which Barabási, Jeong, Oltvai, and Vicsek belong. Interestingly, there is also a small clique, consisting of Boguñá, Pastor-Satorras, and Vespignani, that is topologically, at one step, disconnected from the rich nodes (see Fig. S2 in the Supplemental Material [54] for the completely labeled neighborhood subnetwork). To summarize, in the diffusion space the rich nodes are close to each other not only because of their direct connections, but also because of the presence of different small structures (bridges, cliques, and dense communities) supporting diffusion.

The same holds for other real networks, such as the connectome in the *Drosophila*'s medulla and the worldwide airports network, while the protein-protein interaction (PPI) network of the *Plasmodium falciparum* has generally no rich club, as shown in Figs. 4(a)–4(c). These networks, details about which and references to which are reported in Table I, have been analyzed along with network models. Figures 4(d)–4(f) show the functional rich-club coefficients for the Barabási–Albert (BA) [56], the Erdős-Rényi (ER) [57], and the Lancichinetti-Fortunato-Radicchi (LFR) [58] models; others can be found in the Supplemental Material. All the synthetic networks have $N = 256$ nodes, and their parameters have been chosen to result in an average degree $\langle k \rangle \approx 12$ (see the Supplemental

TABLE I. Real networks. Size of and references for the real networks analyzed in this paper. We considered different types of biological networks: an interactome—the protein interaction network of the malaria parasite *Plasmodium falciparum* (others can be found in the Supplemental Material)—and a connectome. An instance of an infrastructural network is given by the worldwide airports network. This is the reduced version, with 965 airports out of the over 4000 airports of the complete network, proposed in Ref. [5].

	2010 Network Scientists	*Plasmodium* PPI	Connectome	Airports
Nodes (n)	552	1179	1170	965
Edges (n)	1318	2481	8905	33349
Average degree $\langle k \rangle$	4.8	4.2	10.1	69.1
Refs.	[52,53]	[59,60]	[55,61]	[3,5]

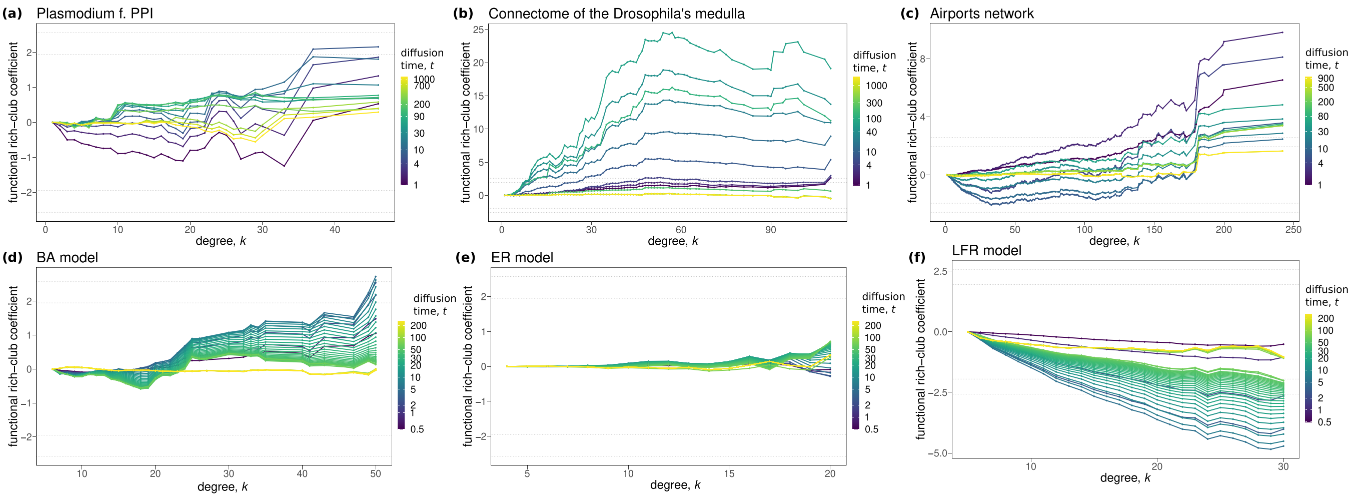

FIG. 4. The functional rich clubs of real and synthetic networks. See Table I for details of the data sets. (a) The protein-protein interaction network (PPI) of *Plasmodium falciparum* has no persistent functional rich-club organization, while (b) the connectome in the *Drosophila*'s optical medulla [55] and (c) the worldwide air transportation network [5] have a clear functional rich club at different scales, i.e., diffusion times t. Synthetic networks are generated from three models: (d) Barabási-Albert (BA), (e) Erdős-Rényi (ER), and (f) Lancichinetti-Fortunato-Radicchi (LFR). These networks have 256 nodes each and an average degree of around 5; nevertheless, their different structural features are reflected in their functional rich-club organizations: The BA network has a significant functional rich-club at the highest values of k, while the ER network is not different from the null model. The LFR has a clear community structure and lacks the functional rich-club organization; however, as nodes become richer, their average core distance grows more than expected by chance.

Material for details); notwithstanding, they have very distinctive characteristics, which are reflected in their functional organization. BA networks have a heterogeneous degree distribution with hubs, and these form a functional rich club. As expected, the nodes in ER networks are neither closer to each other nor mutually more distant than by chance. Finally, the LFR model, which is characterized by the presence of both hubs (heterogeneous degree distribution) and communities with different sizes, further confirms our conjecture that networks with a clear community structure lack a functional rich-club organization. However, we can see in Fig. 4(f) that as nodes become richer, their average core distance grows more than expected by chance.

To conclude, we have shown that not only the absence of a structural rich club, but also the presence of clear mesoscale structures, does not facilitate or even hinders the diffusion between rich nodes. Indeed, functional organizations of nodes into communities, cores, and rich-clubs are competing at some scales. Furthermore, the frequent claim that systems which need to integrate information processed in segregated subsystems display the rich-club phenomenon finds here further support.

III. THEORY

The diffusion distance [6] between two nodes $i, j \in G$ is defined as

$$D_t(i, j; \mathbf{T}) = ||\mathbf{p}(t|i) - \mathbf{p}(t|j)||_2, \tag{4}$$

where $t > 0$ is the diffusion time, $\mathbf{T}$ is the transition matrix of the (discrete-time) embedded jump chain of the particular continuous-time random walk under consideration, and $\mathbf{p}(t|i)$ indicates the probability vector whose kth component $p_k(t|i)$ is the probability of finding the random walker at node k at time t given that it started at node i at time $t = 0$ with probability 1. Defining the random walk normalized Laplacian as $\tilde{\mathbf{L}} = \mathbf{I} - \mathbf{T}$, the continuous-time random walk evolves according to the equation $\dot{\mathbf{p}}(t) = -\mathbf{p}(t)\tilde{\mathbf{L}}$ and, with the previous initial condition $p_i(0) = 1$, the transition probability at time t is given by $p_k(t|i) = (e^{-\tilde{\mathbf{L}}t})_{ik}$. In this paper, the random walker always follows a classical random walk, i.e., $\mathbf{T} = \mathbf{D}^{-1}\mathbf{A}$, so we drop the transition matrix from (4), simplifying the notation to $D_t(i, j)$.

A. Average core diffusion distance

At each $t > 0$ we have a distance matrix $\mathbf{D}_t$, whose ijth entry is given by $(\mathbf{D}_t)_{ij} = D_t(i, j)$, and we know from Ref. [8] that the pairwise distances are bounded in $[0, \sqrt{2}]$, but as $t \to \infty$, assuming the network is connected, $D_t(i, j) \to 0$ since the posterior probability vector will tend to the invariant or stationary distribution. Hence we need to remove the effect of the whole network, and we do so by standardizing the distances: Firstly, we compute the average $\mu(D_t)$ and standard deviation $\sigma(D_t)$ of the diffusion distances over the whole network and then compute the pairwise standardized distances

$$\tilde{D}_t(i, j) = \frac{D_t(i, j) - \mu(D_t)}{\sigma(D_t)}. \tag{5}$$

Now, $\tilde{D}_t(i, j)$ is zero on average with unitary variance, and may take negative values. As a matter of fact, if $\tilde{D}_t(i, j) < 0$, then i and j are closer to each other with respect to the average pairwise distance of the network. Moreover, the magnitude $|\tilde{D}_t(i, j)|$ tells us how far (i.e., how many standard deviations) this distance is from the mean network distance. Henceforth, we drop the "tilde notation" for the standardized distances, since there is no ambiguity being that all distances are standardized with respect to the network distribution.

For each node degree k, we consider $\{D_t(i, j) : i, j \in G, k_i, k_j \geqslant k\}$ and compute the average over this set, which we henceforth indicate as $\mu(D_t|_{\geqslant k})$ and refer to as the average core diffusion distance, or simply the core distance of G at level k. Observe that $\mu(D_t|_{\geqslant k})$ is not a statistic, since $\{\tilde{D}_t(i, j) : i, j \in G, k_i, k_j \geqslant k\}$ is not a random sample; instead it is a measure of the diffusion closeness of the nested subnetworks of G.

This needs to be compared with a null model, so let us take a sample $G_1, \dots, G_n$ from the configuration model [51] with the same degree sequence as the network G under study. In this paper, we mostly fixed $n = 30$, except for the toy model in Figs. 1 and 2, where $n = 50$. For each realization G_i, we compute the diffusion distance at the same time scale and repeat the same standardizing procedure. We then collect the set of core distances for each G_i, put them together in a unique large sample, and finally compute its sample mean $m(D_t^{\mathrm{cm}}|_{\geqslant k})$ and standard deviation $s(D_t^{\mathrm{cm}}|_{\geqslant k})$.

B. The functional rich-club coefficient

We define the functional rich-club coefficient, $\mathrm{FRCC}_{t,k}$, in terms of a standardized distance from the configuration model. Since we would like to observe positive values when the network G under study has smaller core distances than the null model, we take

$$\mathrm{FRCC}_{t,k} = \frac{m\big(D_t^{\mathrm{cm}}|_{\geqslant k}\big) - \mu(D_t|_{\geqslant k})}{s\big(D_t^{\mathrm{cm}}|_{\geqslant k}\big)}, \tag{6}$$

which quantifies how far the average network core distance is from the average configuration model core distance, in units of the standard deviation of the configuration model core distances. One advantage of using standardized scores is that they can be compared without worrying about scales: We can compare the $\mathrm{FRCC}_{t,k}$ at different diffusion times, and if we standardize the structural rich-club coefficient, we can also compare the two rich-club organizations. The Z score for the structural rich-club coefficient is evaluated similarly (but with inverted sign):

$$\phi^Z(k) = \frac{\phi(k) - m(\phi^{\mathrm{cm}}(k))}{s(\phi^{\mathrm{cm}}(k))}. \tag{7}$$

C. Model of a functional rich club

In Figure 1 we presented one possible model of synthetic network with a functional rich-club organization. Here, we describe in detail the four construction steps and the rationale behind them. To build a synthetic (unweighted and undirected) network with a defined and controllable functional rich-club, we need two main ingredients: hubs, i.e., nodes with high degree, and a connectivity that supports diffusion among rich nodes. So we start with four subnetworks $G_1, \dots, G_4$, with 64 nodes each, generated from the BA model (with $m = 2$), which, at step 1, are connected very loosely. Nodes inside each community are ranked in decreasing degree order, and we call the top four ranking nodes in each group *the rich nodes*.

Rich nodes with the same rank are then connected either (i) in a ring, i.e., the rth-ranking node r_i in the community G_i is connected to the rth-ranking node $r_{i+1} \in G_{i+1}$ for $i = 1, 2, 3$ and r_4 links to r_1, or (ii) in a cross, with r_1 connected to r_3 and with r_2 connected to r_4. We name this connectivity pattern "rings and crosses." Rings are placed among nodes of rank $r = 1, 3$, and crosses are placed among nodes of rank $r = 2, 4$ (see Fig. S1 in the Supplemental Material). Additionally, five random edges link the five lowest degree nodes of each pair of distinct communities. Of course, building a clique between rich nodes would add a lot of short paths between them and, consequently, decrease their pairwise diffusion distances. Unfortunately, this is not realistic and, although it guarantees a clear structural rich club, it is not enough to obtain a strong functional rich club at different scales, because of the mesoscale organization of the nodes into well-separated communities. At the second and third steps, inter-rank links between rich nodes are added in order to create paths of length larger than 1 between rich nodes. At step 2, in particular, the first-ranking node of each group G_i is linked at random with (i) two fourth-ranking nodes outside G_i and (ii) one third-ranking node of a group chosen at random in $\{G_1, \dots, G_4\} \setminus \{G_i\}$, indicated henceforth as $G \setminus G_i$, and, similarly, one second-ranking node in $G \setminus G_i$. In the same manner, at step 3, the second-ranking node of each group connects randomly to two third-ranking nodes and to one fourth-ranking node of the other groups. Finally, at step 4 we compute the 75th and 90th percentiles of the degree distribution, corresponding to $k_{0.75} = 4$ and $k_{0.9} = 8$, and we place ten edges between ten nodes with a degree $k > k_{0.9}$ chosen at random inside group G_i to ten nodes with a degree $k_{0.75} \leqslant k \leqslant k_{0.9}$ chosen at random in $G \setminus G_i$. The average degree of G at the different steps varies from $\langle k \rangle \approx 4.2$ to $\langle k \rangle \approx 4.9$.

IV. CONCLUSIONS

Many empirical networks are not static, and connections are purposefully created or destroyed to achieve a particular function, such as driving people or goods through a transportation network, processing an external stimulus in the brain, or, in general, exchanging information. Hence the interplay between structure and function has to be taken into account when ascribing a particular functional output to an observed structural pattern. Although the rich-club phenomenon is found in different types of systems and is considered responsible for the rapid exchange and integration of information, its definition is purely structural, assuming that communication follows preferentially and exclusively the shortest paths connecting rich nodes. Shortest paths and their lengths play undoubtedly a key role in the communication between the units and, ultimately, in the function of a system, where, indeed, the observed short average path lengths are achieved at a cost and are not the consequence of randomness. Similarly, shortest paths cannot be the unique reservoir for pathways of information flow and other communication strategies, and descriptors based on them should complement our network analyses.

In this paper, we fill this gap by tackling the rich-club phenomenon from the functional perspective of diffusion geometry, a framework to investigate the latent geometry induced by network-driven processes, asking whether rich nodes are closer in the diffusion space than expected by chance. From the analysis of our toy model and synthetic networks, we have shown that the structural rich-club organization does not translate seamlessly into function, because the existence of communities can hinder the diffusion between rich nodes despite the one-link paths connecting them. Remarkably, we have shown that a strong functional rich club may emerge without a structural counterpart (e.g., the airports network for large degrees), or with a weak structural counterpart (the 2010 Network Scientists network), because the diffusion geometry at not too local time scales integrates the structural information of longer paths so that these contribute to an efficient diffusion between nodes. With these results, we foresee the potential of this framework if extended to other communication mechanisms which lie in between routing and broadcasting and are important to many real systems.

ACKNOWLEDGMENT

The authors thank Dirk Brockmann for providing the worldwide air-transportation flow data.

---

[1] M. E. J. Newman, Assortative Mixing in Networks, Phys. Rev. Lett. **89**, 208701 (2002).
[2] S. Zhou and R. J. Mondragón, The rich-club phenomenon in the internet topology, IEEE Commun. Lett. **8**, 180 (2004).
[3] D. Helbing, D. Brockmann, T. Chadefaux, K. Donnay, U. Blanke, O. Woolley-Meza, M. Moussaid, A. Johansson, J. Krause, S. Schutte, and M. Perc, Saving human lives: What complexity science and information systems can contribute, J. Stat. Phys. **158**, 735 (2015).
[4] M. E. J. Newman, The structure and function of complex networks, SIAM Rev. **45**, 167 (2003).
[5] D. Brockmann and D. Helbing, The hidden geometry of complex, network-driven contagion phenomena, Science **342**, 1337 (2013).
[6] M. De Domenico, Diffusion Geometry Unravels the Emergence of Functional Clusters in Collective Phenomena, Phys. Rev. Lett. **118**, 168301 (2017).
[7] R. Peach, A. Arnaudon, and M. Barahona, Relative, local and global dimension in complex networks, Nat. Commun. **13**, 3088 (2022).
[8] G. Bertagnolli and M. De Domenico, Diffusion geometry of multiplex and interdependent systems, Phys. Rev. E **103**, 042301 (2021).
[9] A. Arnaudon, R. L. Peach, and M. Barahona, Scale-dependent measure of network centrality from diffusion dynamics, Phys. Rev. Research **2**, 033104 (2020).
[10] G. Bertagnolli, C. Agostinelli, and M. De Domenico, Network depth: identifying median and contours in complex networks, J. Complex Networks **8**, 4 (2019).
[11] S. Fortunato and M. Barthelemy, Resolution limit in community detection, Proc. Natl. Acad. Sci. USA **104**, 36 (2007).
[12] A. Arenas, A. Fernández, and S. Gómez, Analysis of the structure of complex networks at different resolution levels, New J. Phys. **10**, 053039 (2008).
[13] T. Kawamoto and M. Rosvall, Estimating the resolution limit of the map equation in community detection, Phys. Rev. E **91**, 012809 (2015).
[14] S. Fortunato and D. Hric, Community detection in networks: A user guide, Phys. Rep. **659**, 1 (2016).
[15] P. Pons and M. Latapy, Computing communities in large networks using random walks, in *Computer and Information Sciences - ISCIS 2005* (Springer, Berlin, 2005), pp. 284–293.
[16] M. Rosvall and C. T. Bergstrom, Maps of random walks on complex networks reveal community structure, Proc. Natl. Acad. Sci. USA **105**, 1118 (2008).
[17] A. Arenas, A. Díaz-Guilera, and C. J. Pérez-Vicente, Synchronization Reveals Topological Scales in Complex Networks, Phys. Rev. Lett. **96**, 114102 (2006).
[18] P. Ronhovde and Z. Nussinov, Local resolution-limit-free Potts model for community detection, Phys. Rev. E **81**, 046114 (2010).
[19] J. Reichardt and S. Bornholdt, Statistical mechanics of community detection, Phys. Rev. E **74**, 016110 (2006).
[20] G. Ball, P. Aljabar, S. Zebari, N. Tusor, T. Arichi, N. Merchant, E. C. Robinson, E. Ogundipe, D. Rueckert, A. D. Edwards, and S. J. Counsell, Rich-club organization of the newborn human brain, Proc. Natl. Acad. Sci. USA **111**, 7456 (2014).
[21] R. F. Betzel and D. S. Bassett, Multi-scale brain networks, NeuroImage **160**, 73 (2017).
[22] B. Mišić, R. F. Betzel, M. A. de Reus, M. P. van den Heuvel, M. G. Berman, A. R. McIntosh, and O. Sporns, Network-level structure-function relationships in human neocortex, Cereb. Cortex **26**, 3285 (2016).
[23] L. Harriger, M. P. van den Heuvel, and O. Sporns, Rich club organization of macaque cerebral cortex and its role in network communication, PLoS One **7**, e46497 (2012).
[24] M. Senden, G. Deco, M. A. de Reus, R. Goebel, and M. P. van den Heuvel, Rich club organization supports a diverse set of functional network configurations, NeuroImage **96**, 174 (2014).
[25] M. P. Van Den Heuvel and O. Sporns, Rich-club organization of the human connectome, J. Neurosci. **31**, 15775 (2011).
[26] A. Arnatkeviciute, B. D. Fulcher, S. Oldham, J. Tiego, C. Paquola, Z. Gerring, K. Aquino, Z. Hawi, B. Johnson, G. Ball, M. Klein, G. Deco, B. Franke, M. A. Bellgrove, and A. Fornito, Genetic influences on hub connectivity of the human connectome, Nat. Commun. **12**, 4237 (2021).
[27] M. P. van den Heuvel, O. Sporns, G. Collin, T. Scheewe, R. C. W. Mandl, W. Cahn, J. Goñi, H. E. H. Pol, and R. S. Kahn, Abnormal rich club organization and functional brain dynamics in schizophrenia, JAMA Psychiatry **70**, 783 (2013).
[28] M. Daianu, A. Mezher, M. F. Mendez, N. Jahanshad, E. E. Jimenez, and P. M. Thompson, Disrupted rich club network in behavioral variant frontotemporal dementia and early-onset Alzheimer's disease, Human Brain Mapping **37**, 868 (2015).
[29] M. Daianu, N. Jahanshad, T. M. Nir, C. R. Jack, M. W. Weiner, M. A. Bernstein, and P. M. Thompson, Rich club analysis in the Alzheimer's disease connectome reveals a relatively undisturbed structural core network, Human Brain Mapping **36**, 3087 (2015).

[30] L. L. Gollo, A. Zalesky, R. M. Hutchison, M. van den Heuvel, and M. Breakspear, Dwelling quietly in the rich club: Brain network determinants of slow cortical fluctuations, Philos. Trans. R. Soc. B: Biol. Sci. **370**, 20140165 (2015).

[31] R. Li, W. Liao, Y. Li, Y. Yu, Z. Zhang, G. Lu, and H. Chen, Disrupted structural and functional rich club organization of the brain connectome in patients with generalized tonic-clonic seizure, Human Brain Mapping **37**, 4487 (2016).

[32] X. Liu, C. He, D. Fan, Y. Zhu, F. Zang, Q. Wang, H. Zhang, Z. Zhang, H. Zhang, and C. Xie, Disrupted rich-club network organization and individualized identification of patients with major depressive disorder, Prog. Neuro-Psychopharmacol. Biol. Psychiatry **108**, 110074 (2021).

[33] P. Moretti and M.-T. Hütt, Link-usage asymmetry and collective patterns emerging from rich-club organization of complex networks, Proc. Natl. Acad. Sci. USA **117**, 18332 (2020).

[34] T. Opsahl, V. Colizza, P. Panzarasa, and J. J. Ramasco, Prominence and Control: The Weighted Rich-Club Effect, Phys. Rev. Lett. **101**, 168702 (2008).

[35] A. Ma, R. J. Mondragón, and V. Latora, Anatomy of funded research in science, Proc. Natl. Acad. Sci. USA **112**, 14760 (2015).

[36] M. Szell and R. Sinatra, Research funding goes to rich clubs, Proc. Natl. Acad. Sci. USA **112**, 14749 (2015).

[37] V. Colizza, A. Flammini, M. A. Serrano, and A. Vespignani, Detecting rich-club ordering in complex networks, Nat. Phys. **2**, 110 (2006).

[38] N. Pedreschi, D. Battaglia, and A. Barrat, The temporal rich club phenomenon, Nat. Phys. **18**, 931 (2022).

[39] M. Jusup, P. Holme, K. Kanazawa, M. Takayasu, I. Romić, Z. Wang, S. Geček, T. Lipić, B. Podobnik, L. Wang, W. Luo, T. Klanjšček, J. Fan, S. Boccaletti, and M. Perc, Social physics, Phys. Rep. **948**, 1 (2022).

[40] Z.-Q. Jiang and W.-X. Zhou, Statistical significance of the rich-club phenomenon in complex networks, New J. Phys. **10**, 043002 (2008).

[41] S. Zhou and R. J. Mondragón, Structural constraints in complex networks, New J. Phys. **9**, 173 (2007).

[42] M. Cinelli, G. Ferraro, and A. Iovanella, Rich-club ordering and the dyadic effect: Two interrelated phenomena, Phys. A (Amsterdam) **490**, 808 (2018).

[43] M. Á. Serrano, Rich-club vs rich-multipolarization phenomena in weighted networks, Phys. Rev. E **78**, 026101 (2008).

[44] M. Cinelli, Generalized rich-club ordering in networks, J. Complex Networks **7**, 702 (2019).

[45] V. Zlatic, G. Bianconi, A. Díaz-Guilera, D. Garlaschelli, F. Rao, and G. Caldarelli, On the rich-club effect in dense and weighted networks, Eur. Phys. J. B **67**, 271 (2009).

[46] J. J. McAuley, L. da Fontoura Costa, and T. S. Caetano, Rich-club phenomenon across complex network hierarchies, Appl. Phys. Lett. **91**, 084103 (2007).

[47] A. Avena-Koenigsberger, B. Misic, and O. Sporns, Communication dynamics in complex brain networks, Nat. Rev. Neurosci. **19**, 17 (2018).

[48] A. Avena-Koenigsberger, X. Yan, A. Kolchinsky, M. P. van den Heuvel, P. Hagmann, and O. Sporns, A spectrum of routing strategies for brain networks, PLoS Comput. Biol. **15**, e1006833 (2019).

[49] R. R. Coifman and S. Lafon, Diffusion maps, Appl. Comput. Harmonic Anal. **21**, 5 (2006).

[50] M. Boguñá, I. Bonamassa, M. De Domenico, S. Havlin, D. Krioukov, and M. Á. Serrano, Network geometry, Nat. Rev. Phys. **3**, 114 (2021).

[51] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, Random Struct. Algorithms **6**, 161 (1995).

[52] L. Bohlin, D. Edler, A. Lancichinetti, and M. Rosvall, Community detection and visualization of networks with the map equation framework, in *Measuring Scholarly Impact: Methods and Practice*, edited by Y. Ding, R. Rousseau, and D. Wolfram (Springer International, Cham, Switzerland, 2014), pp. 3–34.

[53] D. Edler and M. Rosvall, The MapEquation software package, 2013, https://www.mapequation.org/downloads/netscicoauthor2010.net.

[54] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.4.033185 for a completely labeled neighborhood subnetwork of the 2010 Network Scientists Network, functional rich-club coefficients for other synthetic network (along with their parameters), details regarding the “rings and crosses” pattern in the network model, and a description of other types of interactomes and connectomes. The Supplemental Material includes Refs. [62–67].

[55] S.-y. Takemura, A. Bharioke, Z. Lu, A. Nern, S. Vitaladevuni, P. K. Rivlin, W. T. Katz, D. J. Olbris, S. M. Plaza, P. Winston, T. Zhao, J. A. Horne, R. D. Fetter, S. Takemura, K. Blazek, L.-A. Chang, O. Ogundeyi, M. A. Saunders, V. Shapiro, C. Sigmund *et al.*, A visual motion detection circuit suggested by *Drosophila* connectomics, Nature (London) **500**, 175 (2013).

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

[57] P. Erdős and A. Rényi, On random graphs, Publ. Math. Debrecen **6**, 290 (1959).

[58] A. Lancichinetti, S. Fortunato, and F. Radicchi, Benchmark graphs for testing community detection algorithms, Phys. Rev. E **78**, 046110 (2008).

[59] A. Chatr-aryamontri, R. Oughtred, L. Boucher, J. Rust, C. Chang, N. K. Kolas, L. O’Donnell, S. Oster, C. Theesfeld, A. Sellam, C. Stark, B.-J. Breitkreutz, K. Dolinski, and M. Tyers, The BioGRID interaction database: 2017 update, Nucleic Acids Res. **45**, D369 (2017).

[60] C. Stark, BioGRID: A general repository for interaction datasets, Nucleic Acids Res. **34**, D535 (2006).

[61] J. T. Vogelstein, E. Perlman, B. Falk, A. Baden, W. Gray Roncal, V. Chandrashekhar, F. Collman, S. Seshamani, J. L. Patsolic, K. Lillaney, M. Kazhdan, R. Hider, D. Pryor, J. Matelsky, T. Gion, P. Manavalan, B. Wester, M. Chevillet, E. T. Trautman, K. Khairy *et al.*, A community-developed open-source computational ecosystem for big neuro data, Nat. Methods **15**, 846 (2018).

[62] C.-T. Shih, O. Sporns, S.-L. Yuan, T.-S. Su, Y.-J. Lin, C.-C. Chuang, T.-Y. Wang, C.-C. Lo, R. J. Greenspan, and A.-S. Chiang, Connectomics-based analysis of information flow in the *Drosophila* brain, Curr. Biol. **25**, 1249 (2015).

[63] M. D. Domenico, V. Nicosia, A. Arenas, and V. Latora, Structural reducibility of multilayer networks, Nat. Commun. **6**, 6864 (2015).

[64] D. J. Watts and S. H. Strogatz, Collective dynamics of small-world networks, Nature (London) **393**, 440 (1998).

[65] M. Penrose, *Random Geometric Graphs* (Oxford University Press, Oxford, 2003).

[66] P. W. Holland, K. B. Laskey, and S. Leinhardt, Stochastic blockmodels: First steps, Social Networks **5**, 109 (1983).

[67] K. Faust and S. Wasserman, Blockmodels: Interpretation and evaluation, Social Networks **14**, 5 (1992).