Why Are There Six Degrees of Separation in a Social Network?

I. Samoylenko[1,2,*] D. Aleja[3,4,5,*] E. Primo[3,5] K. Alfaro-Bittner[3,5] E. Vasilyeva[1,6] K. Kovalenko[7] D. Musatov[1,8,9] A. M. Raigorodskii[1,9,10,11] R. Criado[3,5,12] M. Romance[3,5,12] D. Papo[13,14] M. Perc[15,16,17,18,19] B. Barzel[20,21,22] and S. Boccaletti[1,3,5,23,24]

[1]*Moscow Institute of Physics and Technology, 9 Institutskiy Pereulok, Dolgoprudny, Moscow Region 141701, Russia*
[2]*National Research University Higher School of Economics, 6 Ulitsa Usacheva, Moscow 119048, Russia*
[3]*Universidad Rey Juan Carlos, Calle Tulipán s/n, 28933 Móstoles, Madrid, Spain*
[4]*Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan 48109, USA*
[5]*Laboratory of Mathematical Computation on Complex Networks and their Applications, Universidad Rey Juan Carlos, Calle Tulipán s/n, Móstoles, 28933 Madrid, Spain*
[6]*P.N. Lebedev Physical Institute of the Russian Academy of Sciences, 53 Leninsky Prospekt, Moscow 119991, Russia*
[7]*Scuola Superiore Meridionale, Largo S. Marcellino, 10, 80138 Napoli NA, Italy*
[8]*Russian Academy of National Economy and Public Administration, 84 Prospekt Vernadskogo, Moscow 119606, Russia*
[9]*Caucasus Mathematical Center at Adyghe State University, 208 Pervomayskaya Ulitsa, Maykop, Adygea 385000, Russia*
[10]*Moscow State University, 1 Leninskie Gory, Moscow 119991, Russia*
[11]*Institute of Mathematics and Computer Science, Buryat State University, 5 Ulitsa Ranzhurova, Ulan-Ude, Buryatia 670000, Russia*
[12]*Data, Complex Networks and Cybersecurity Research Institute, Universidad Rey Juan Carlos, Plaza Manuel Becerra 14, 28028 Madrid, Spain*
[13]*Department of Neuroscience and Rehabilitation, University of Ferrara, Ferrara, Italy*
[14]*Center for Translational Neurophysiology of Speech and Communication, Fondazione Istituto Italiano di Tecnologia, Ferrara, Italy*
[15]*Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia*
[16]*Department of Medical Research, China Medical University Hospital, China Medical University, Taichung 404332, Taiwan*
[17]*Complexity Science Hub Vienna, Josefstädterstraße 39, 1080 Vienna, Austria*
[18]*Alma Mater Europaea, Slovenska ulica 17, 2000 Maribor, Slovenia*
[19]*Department of Physics, Kyung Hee University, 26 Kyungheedae-ro, Dongdaemun-gu, Seoul, Republic of Korea*
[20]*Department of Mathematics, Bar-Ilan University, Ramat-Gan, 52900, Israel*
[21]*The Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan, 52900, Israel*
[22]*Network Science Institute, Northeastern University, Boston, Massachusetts, 02115, USA*
[23]*CNR—Institute of Complex Systems, Via Madonna del Piano 10, I-50019 Sesto Fiorentino, Italy*
[24]*Complex Systems Lab, Department of Physics, Indian Institute of Technology, Indore–Simrol, Indore 453552, India*

A wealth of evidence shows that real-world networks are endowed with the small-world property, i.e., that the maximal distance between any two of their nodes scales logarithmically rather than linearly with their size. In addition, most social networks are organized so that no individual is more than six connections apart from any other, an empirical regularity known as the six degrees of separation. Why social networks have this ultrasmall-world organization, whereby the graph's diameter is independent of the network size over several orders of magnitude, is still unknown. We show that the "six degrees of separation" is the property featured by the equilibrium state of any network where individuals weigh between their aspiration to improve their centrality and the costs incurred in forming and maintaining connections. We show,

*These authors contributed equally to this work.

moreover, that the emergence of such a regularity is compatible with all other features, such as clustering and scale-freeness, that normally characterize the structure of social networks. Thus, our results show how simple evolutionary rules of the kind traditionally associated with human cooperation and altruism can also account for the emergence of one of the most intriguing attributes of social networks.

I. INTRODUCTION

In the short story *Chains* (1929), the Hungarian writer Frigyes Karinthy described a game where a group of people were discussing how the members of the human society were closer together than ever before. To prove this point, one participant proposes that any person out of the entire Earth population (around 1.8 billion at that time) could be reached using nothing except each personal network of acquaintances, betting that the resulting chain would be of no more than five individuals. The story coined the expression "six degrees of separation" to reflect the idea that all people of the world are six or fewer social connections apart from each other. The concept was later generalized to that of "small-world" networks, where the maximal social distance (the diameter of the network) scales logarithmically, rather than linearly, with the size of the population [1].

After early studies on the structure of social networks by Gurevitch [2] and de Sola Pool and Kochen [3], Milgram performed his 1967 famous set of experiments on social distancing [4] (see also Ref. [5]) where, with a limited sample of 1000 individuals, it was shown that people in the U.S. are indeed connected by a small number of acquaintances. Later on, Dodds *et al.* recreated Milgram's experiments with Internet email users [6] by tracking 24 163 chains aimed at 18 targets from 13 countries and confirmed that the average number of steps in the chains was around six. Furthermore, many experiments conducted at a planetary scale on various social networks verified the ubiquitous character of this feature: (i) a 2007 study by Leskovec and Horvitz (with a dataset of 30 billion conversations among 240 million Microsoft Messenger users) revealed the average path length to be six [7] (see also Ref. [8]), (ii) the average degree of separation between two randomly selected Twitter users was found to be 3.435 [9], and (iii) Facebook's network in 2011 (721 million users with 69 billion friendship links) displayed an average distance between nodes of 4.74 [10].

Such abundant and consistent evidence points to the fact that the structure of these networks radically differs from either that of regular networks (where the diameter scales linearly with the size) and that of classical small-world networks (where, instead, the scaling law is logarithmic) [1]. A clear explanation of the mechanisms through which social networks organize into *ultrasmall-world* states (where the diameter does not depend on the system size over several orders of magnitude) is, however, still missing. Why does such a collective property emerge? What are its fundamental mechanisms? Why is the common shortest path length between units of a social network six, rather than five or seven or any other number, implying an average distance which is also not far from six?

We here answer these questions in exact terms, by adopting a game theoretical approach for describing the network evolution, a line of studies which started almost five decades ago by Myerson [11] analyzing cooperation structures in a wide class of games. A couple of decades later, games on adaptive networks were introduced, for instance, by Jackson and Wolinsky [12], with the purpose of studying the stability and efficiency of social and economic networks when self-interested individuals could form or sever links with others. Further on, the influential work by Nowak and May [13] showed how spatial structure could provide an evolutionary escape hatch for cooperation in social dilemmas. Coevolutionary networks have then been considered in a series of works where players could improve their topological position, for example, by cutting links to defectors or rewiring their links to gain larger payoffs [14–19]. Related research also covered game theoretical models as the basis for cooperation on networks [20], for network formation and growth [21–24], as well as for agents to achieve a position of high centrality while minimizing the number of contacts they have to maintain [25].

So far, the few available studies on ultrasmall-world states have focused on finding the relationship between the scaling properties of distances in a graph and those of the node's degree distribution. It was indeed proved that scale-free networks with degree distribution $p(k) \sim k^{-\gamma}$ and $2 < \gamma < 3$ (as it is observed in all real-world networks) display a scaling of the diameter as $D \sim \ln \ln N$ [26], which departs from the classic logarithmic scaling of small-world networks and yet maintains an explicit dependence on the network size N. On the other hand, scale-free networks featuring an asymptotically invariant shortest path (called *Mandala networks* [27]) may be synthesized, which however have an associated value of γ strictly equal to 2 and therefore do not match any case observed in the real world.

Rather than being dependent on global (i.e., degree distribution) scaling properties, in this article we show that the mechanism behind such observed regularity can be found, instead, in a dynamic evolution of the network. Precisely, we rigourously show that, when a simple compensation rule between the cost incurred by nodes in maintaining connections and the benefit accrued by the chosen links is governing the evolution of a network, the asymptotic equilibrium state (a Nash equilibrium where no further actions would produce more benefit than cost [28])
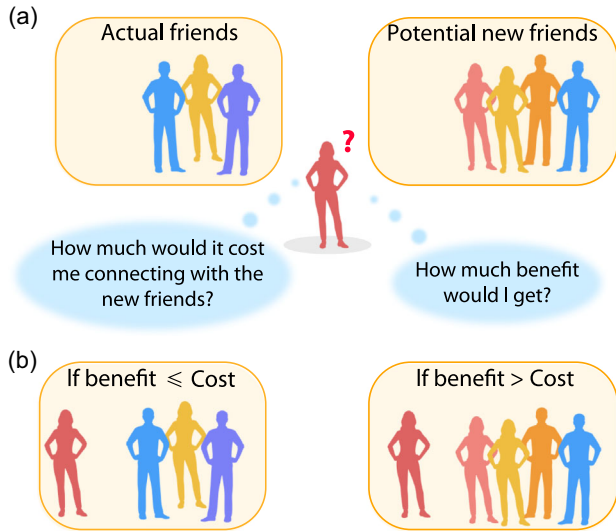
FIG. 1. The game theoretical framework. The structure of a social network evolves following simple rules of a game. (a) At each step of the game, the individuals forming part of the network (like the red woman in the picture) have to decide whether to stay with the neighborhood formed by their actual friends or to change to another neighborhood formed by potential new friends. The current and new neighborhoods may overlap (in our picture, the blue man and the yellow woman are members of both sets). The decision is based on a careful evaluation of the cost incurred and of the benefit gained with the change. (b) The decision is merely utilitarian. If the benefit is not overcoming the cost, then individuals maintain their current neighborhood (left-hand picture). If, on the contrary, the payoff exceeds the cost, then individuals relinquish their current neighborhood and move to the new one (right-hand picture). The structure of the network then evolves until converging to its Nash equilibrium (if it exists), i.e., to the configuration where no changes of neighborhood are allowed, as no individual has anything to gain in abandoning acquaintances.

features a diameter which does not depend on the system's size, and is equal to six. In other words, we theorematically prove that any network where nodes strive to increase their centrality by forming connections if and only if their cost is smaller than the payoff tends to evolve into an ultrasmall-world state endowed with the six degree of separation property, irrespective of its initial structure. Our study points out, therefore, that evolutionary rules of the kind traditionally associated with human cooperation and altruism [29–33] can in fact account also for the emergence of this attribute of social networks. Furthermore, we show that such a global network feature can emerge even from situations where individuals have access to only partial information on the overall structure of connections, which is indeed the case in almost all social networks.

II. RESULTS

A. Game theoretical model for network evolution

Consider the general case which is schematically depicted in Fig. 1, where the N nodes of a network V

are rational agents of a game. At each step m of the game, each agent $v \in V$ selects (independently of the choices made by the other agents at the same step) a potential neighborhood $\mathcal{N}_v(m)$ made of $k_v(m)$ other nodes of V. The agent then decides whether it is more profitable to form connections with the nodes in $\mathcal{N}_v(m)$ or to remain connected with the nodes in $\mathcal{N}_v(m-1)$. The decision is based on a balance between the payoff and the cost functions associated with the change of neighborhood.

As for the cost function, we assume that node v pays a unitary cost $c > 0$ to maintain a connection with each node u belonging to its neighborhood (and that node u cannot refuse the connection paid for by v). Moreover, to be as generic as possible, we assume the unitary cost either to be a constant or to depend on the network size as $c = c(N)$.

As for the benefit function, if agents are rational, it is logical to assume that their goal is to increase their importance within the network. This can be naturally framed in terms of betweenness centrality [34], which indeed provides a measure of the influence exerted by a node on the information flow within a network. This is defined as follows. First of all, if v and s are two nodes of a connected network, the distance $l(v, s)$ is taken to be the number of edges forming the shortest path between them. Then, the betweenness centrality (or degree of mediation) B_v is taken to be $\sum_{s \neq v \neq t} [\sigma_{st}(v)/\sigma_{st}]$, where $s, t \in V$ are all possible pairs of different vertices that do not match with v, $\sigma_{st}(v)$ is the number of shortest paths between the vertices s and t passing through the vertex v, and σ_{st} is the total number of shortest paths between the vertices s and t.

B_v quantifies how relevant the intermediary role played by v in the graph is. However, one immediately realizes that the contribution in B_v of the shortest paths in which v is *the unique intermediary* between s and t is equal to that of paths in which v is just one of a long chain of intermediaries. To account for such a difference, one may adopt a generic weighted version of the betweenness centrality $WBC(v)$ which is defined as

$$WBC(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \cdot f[l(s, t)], \tag{1}$$

where f is a strictly decreasing function of its argument (as longer paths must contribute less). One can think of Eq. (1) as follows: each pair s, t of vertices creates some utility, which is then distributed equally among all shortest paths from s to t, and then each intermediary vertex in each path obtains a fraction equal to $f[l(s, t)]/\sigma_{st}$.

With these simple rules in mind, the N agents play the game. When the game converges to a Nash equilibrium (a configuration where no agent has anything to gain by changing its own neighborhood, as all of them have already attained their optimal adjacency), we can demonstrate rigorously that the obtained structure is endowed with the six degrees of separation attribute.
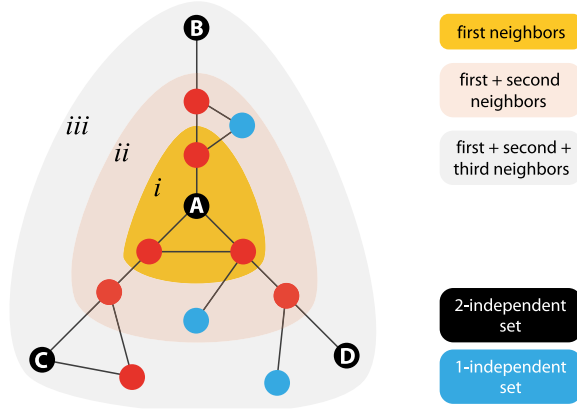
FIG. 2. *l*-independence of nodes. Sketch of a generic graph, with node A at the center. The first, second, and third neighbors of node A are, respectively, located within the yellow, pink, and gray region. The *l*-independent set of a graph is the set of nodes such that the distance between any two of them is larger than *l*. The black nodes (A, B, C, and D) form the 2-independent set of the graph, as all of them are at a distance larger than 2 from each other. The black nodes together with the ones depicted in light blue form, instead, the 1-independent set. Note that the light blue nodes do not participate in the 2-independent set. Finally, the red nodes belong neither to the 1-independent set nor to the 2-independent set.

B. 2-independent sets and the emergence of ultrasmall-world states

Before we demonstrate our main results, we first need to introduce the concept of 2-independence of network's nodes. In traditional graph theory, a 1-independent set (or internally stable set, or anticlique) S is a set of vertices such that any pair of them is not connected by a graph's edge. This is to say that each edge in the graph has at most one end point in S. As a consequence, any two vertices of S are at a distance which is strictly larger than one.

One can now generalize the latter definition and designate as an *l*-independent set S_l the set of the network's nodes such that the distance between any pair of its members is larger than *l* [35]. It follows that nodes belonging to S_l do not necessarily belong to S_{l+1} (see Fig. 2 for an illustrative sketch of the comparison between a 2-independent set and a classical 1-independent set).

Why are 2-independent sets important in our framework? This can be understood by looking at Fig. 3. In Fig. 3(a), the three vertices 1,2,7 are originally part of a 1-independent set. Now, if vertex 7 forms the yellow edges (7,1) and (7,2), it is removed from the set but it does not change the distance between nodes 1 and 2. It only contributes to the multiplicity of shortest paths between nodes 1 and 2. As the number of alternative shortest paths may be very large in large sized networks, the minimum possible benefit obtained from gluing a 1-independent set (as node 7 would do by forming edges with nodes 1 and 2) may be very small with the growth of the network's size. From the latter point it follows that the presence of
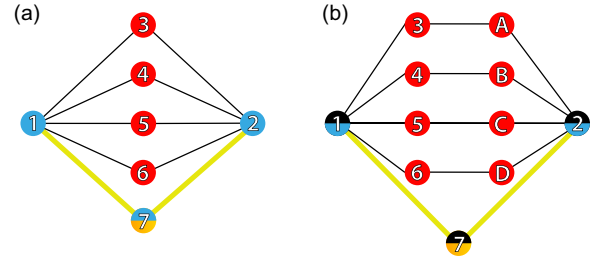

FIG. 3. Independence of nodes and Nash equilibrium. (a) When only black links are considered, vertices 1,2,7 form a 1-independent set. For consistency with Fig. 2, nodes 1 and 2 are colored in light blue. As vertex 7 forms the yellow edges (7, 1) and (7, 2) it is removed from the 1-independent set (this change is depicted by coloring the lowest part of the node in yellow), but the two new connections do not remove nodes 1 and 2 from the 1-independent set, since they only contribute to the multiplicity of the shortest paths between 1 and 2. (b) When only black links are considered, vertices 1,2,7 form a 2-independent set. As the yellow connections are formed, vertex 7 reduces the distance between nodes 1 and 2 from at least 3 down to 2. As a consequence, nodes 1 and 2 can only be part of a 1-independent set. For this reason, the upper half of nodes 1 and 2 is depicted in black, indicating that these nodes initially belonged to the 2-independent set, and the lower half in light blue, indicating that by receiving the connections from node 7, they become members of a 1-independent set. Node 7 is half colored in black (as it initially belonged to the 2-independent set) and half in yellow (as the two new connections remove it from all independent sets).

independent sets of large size may be compatible with the Nash equilibrium.

A totally different situation occurs when we consider 2-independent sets, as in Fig. 3(b). Indeed, when forming the yellow connections with nodes 1 and 2, vertex 7 is actually reducing their distance from at least 3 down to 2. Therefore, regardless of which other edge exists in the network involving vertices 1 and 2, vertex 7 receives a minimum benefit equal to $f(2)$. This is equally valid for any other vertex of the 2-independent set which would form edges with all other members of the set: it would receive at least the same benefit from each pair of nodes in the set. Therefore, the minimal benefit obtained from gluing a 2-independence set of size x is $\binom{x-1}{2}f(2)$, which may be rather substantial. For this reason, sizable 2-independent sets cannot exist in the Nash equilibrium.

The process of gluing large size 2-independence sets is precisely what regulates the spontaneous emergence of the six degrees of separation. Namely, it can be proved theorematically that such a process determines that
(i) at the Nash equilibrium the graph necessarily contains at least a vertex v whose degree k scales as the cube root of the system's size i.e., $k \sim \sqrt[3]{N}$, and
(ii) node v is at the center of the network and displays the remarkable property of being at a distance of no more than 3 from any other node of the graph.

The latter implies that the shortest path between any pair of nodes i, j in the graph will be smaller than or equal to 6, as there will be maximum three edges forming the shortest path from i to v and maximum three edges also to form the shortest path from v to j. Therefore, the diameter D of the network will be exactly 6.

The proofs of the theorems and lemmas involved are available in the Supplemental Material (SM) [36].

C. Illustrative case

For the sake of a better illustration, let us now focus on the case described as follows.

  (i) The agents start the game when they are already connected by means of a pristine graph where, in addition, there exists at least one node with sufficiently high degree.

  (ii) Each agent v adopts as benefit function

$$WBC(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \cdot \frac{1}{l(s,t)^{\alpha}}, \qquad (2)$$

with α being a strictly positive parameter. Comparing with Eq. (1), this means that the weighting factor is $f[l(s,t)] = 1/l(s,t)^{\alpha}$, and that Eq. (2) coincides, for $\alpha = 1$, with the classical weighted betweenness centrality [34].

  (iii) Agents sequentially add new connections to their neighborhood if and only if there is a positive balance between the extra utility brought by the new connections and the extra cost.

In practice, at each step m of the game, the potential neighborhood $\mathcal{N}_v(m)$ of each agent $v \in V$ is equal to $\mathcal{N}_v(m-1)$ plus p other nodes. The p new edges are then added only if $\Delta WBC(v) \geq pc$, i.e., only if the extra weighted betweenness centrality is larger than or equal to the extra cost pc.

When no agent is able to incorporate any further edge, the network is said to have reached its asymptotic equilibrium. It should be remarked that such a final state cannot formally be associated to a Nash equilibrium, because the option of removing existing links is not contemplated in the game, and therefore there is no certainty that agents, in their asymptotic states, are in their optimal adjacency configuration. In this respect, it is worth highlighting that another mechanism (beyond that of link addition and deletion) that one can consider at the basis of the emergence of the six degrees of separation is that of link rewiring, which would actually imply the invariance of the network density during its evolution toward the asymptotic equilibrium. We plan to report on the effects of this latter mechanism elsewhere.

The following theorem can be proved.

  (i) If v is a node of the pristine graph with k original connections, and

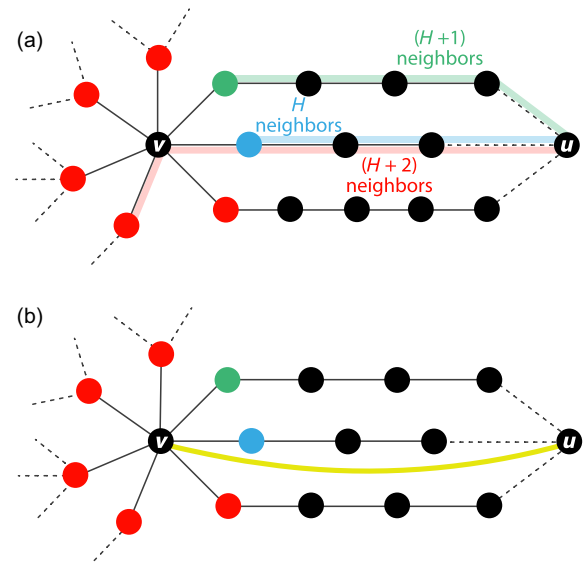 (ii) if $H \in \{3, 4, 5, \ldots\}$ is some integer number strictly larger than 2, and

FIG. 4. The emergence of the ultrasmall-world state. (a) Sketch of a hypothetical network where nodes v and u are separated by a distance $H + 1$. The neighbors of v are then at either H (the light blue node), or $H + 1$ (the green node s), or $H + 2$ (the red nodes) edges from u. For a better visualization, paths of different lengths are marked with the corresponding colors. (b) A direct (yellow) link is added between v and u. Our study demonstrates rigorously (see Theorem 3 of the SM [36]) that the network configuration of (a) is incompatible with an equilibrium state. Note that, at variance with the case depicted in Fig. 2, here distances between nodes depend on the value of the parameter H.

  (iii) if, for the considered values of c and α, the inequality

$$\left(\frac{1}{2^{\alpha}} - \frac{1}{(H+2)^{\alpha}} \right) k \geq c \qquad (3)$$

is satisfied, then, in the equilibrium state of the network, the node v is linked to all other nodes of the graph by no more than H links, implying that the diameter of the equilibrium network does not exceed $2H$.

In practice, the theorem guarantees that the asymptotic state of a network evolving from an initial condition that satisfies condition (3) is an ultrasmall-world state (and, for $H = 3$, also the emergence of the six degrees of separation property).

The proof of the theorem (see SM for details [36]) is given by contradiction i.e., by supposing that there is a node u in the final state of the network whose distance from v is at least $H + 1$, i.e., $l(u,v) \geq H + 1$. To better illustrate the situation, we depict in Fig. 4(a) the case in which nodes v and u are separated by a distance $H + 1$. In that circumstance, the nodes directly connected to v (the neighbors of v) may be found at either H (the light blue node), or $H + 1$ (the green node), or $H + 2$ (the red nodes) edges from u. Looking at the figure, it is easy to understand that

all network's shortest paths which end in u and start in either the green or the light blue node *cannot* pass through v. Therefore, the only contribution to the benefit function of v from shortest paths ending in u is coming from those paths which start in the red nodes, the neighbors of v that are at distance $H + 2$ from u.

When one, instead, includes a direct link between v and u [the yellow link in Fig. 4(b)], then the shortest path between any neighbor of v (denoted by w) and u becomes $w - v - u$, since $H \geq 3$. Calculating then the value of $\Delta WBC(v)$ corresponding to the addition of such a link, and recalling that the equilibrium requires $\Delta WBC(v)$ to be smaller than the cost c, one easily gets to an expression which is in explicit contradiction with condition (3) (see the SM for full details [36]).

D. Realistic case

We remark that our approach is valid independently of the specific degree distribution properties of the pristine graph. However, the maximum degree of a node in a scale-free network generated by the preferential attachment method [37] is known to scale as $\sqrt{N}$ [38,39] and this implies that, for these networks, condition (3) is (from a given size on) always verified for any value of fixed cost c and any value of α, thus making them very good candidates for initializing the formation of ultrasmall-world structures.

Therefore, to illustrate power and generality of the above theorem, we perform a massive numerical trial by initializing our game on networks of N nodes generated with the Barabási-Albert (BA) algorithm [37], for $\alpha = 1$ (i.e., adopting as benefit the weighted betweenness centrality), $H = 3$, and $c = 0.15\sqrt{N}$ (to ensure a coherent scaling of the cost with that of the maximum degree in the network). With these stipulations, condition (3) becomes $0.3k \geq 0.15\sqrt{N}$. As $k \approx 2\sqrt{N}$ [38,39], this means that condition (3) is verified at each value of N, and one then expects that the diameter at equilibrium would not exceed 6.

It is important to remark here that estimating the benefit function (2) requires the retrieval of the global structure of the network's pathways at each step of the game. However, such information is in general not available to the agents of real social networks. Indeed, computing Eq. (2) becomes prohibitively costly as the size of the network increases, requiring (with the fastest existing algorithms) $\mathcal{O}(NL)$ operations (L being the total number of links in the network) [40,41].

For this reason, it is much more realistic and much less computationally demanding to assume that agents use only local information. We then consider a scenario wherein at each step m of the game, a (large degree) node v is chosen. v incorporates an edge with another node u if
(a) $0.3k \geq c$, where k is the degree of v,
(b) the distance between u and v is larger than 3.

In this way, it is only required to check that the subgraph formed by v and its first and second neighbors has zero
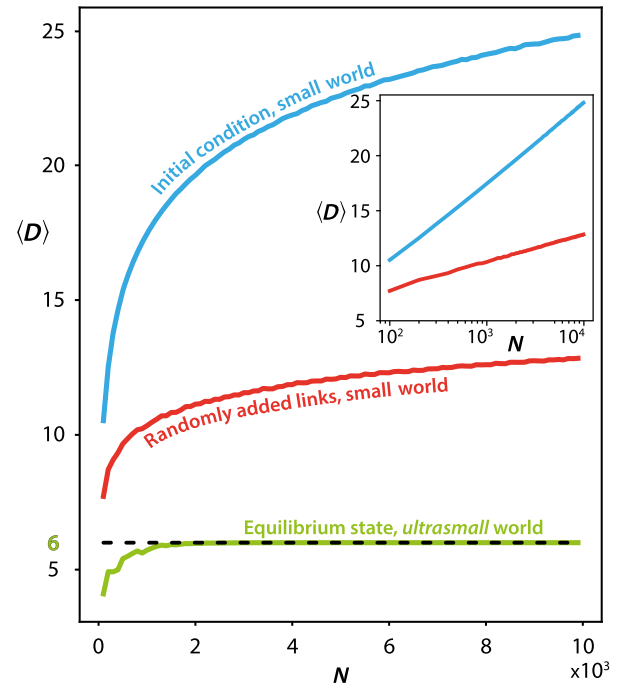

FIG. 5. The emergence of the six degrees of separation. Ensemble average $\langle D \rangle$ versus N for different sets of networks. Light blue line: BA scale-free networks that are used as initial conditions for the evolution of the game. Green line: networks generated at the equilibrium state of the game. Red line: networks constructed by randomly adding to the initial condition of each game the same number of links needed to reach the game equilibrium. A horizontal black dashed line is positioned at $\langle D \rangle = 6$ to indicate that the network's structure obtained at the equilibrium features the ultrasmall-world property, with the concurrent emergence of the six degrees of separation. Inset: log-lin plot of $\langle D \rangle$ versus N. The logarithmic scaling of the light blue and red lines is clearly visible.

overlap with the subgraph formed by u and its first neighbors, and the method is not hurting for the knowledge of the overall shortest paths' structure. At the same time, the adoption of local information makes our study's main claims even stronger, because it proves that a global network property (the network diameter) may emerge as a result of a game in which agents share only local information, which is what actually happens in almost all real circumstances.

Note that, if an edge connecting u and v is added, the above conditions imply that $\Delta WBC(v) \geq c$. Indeed, if the node u satisfies condition (b), it can easily be shown (using the same arguments as in the SM for the proof of the theorem [36]) that the maximum contribution to v of the shortest paths between u and a neighbor of v is $1/5$. Adding the new edge, such a contribution raises to $1/2$, and this means that

$$\Delta WBC(v) \geq 0.3k,$$

where k is the number of connections of v. Therefore, if condition (a) holds, then condition $\Delta WBC(v) \geq c$ is also

satisfied. This implies that our local method is actually more restrictive when incorporating edges, and yet sufficient to give evidence of the predictions of the theorem for nodes satisfying condition (3).

The results of our simulation trial are presented in Fig. 5. At each value of the network size N, 10 000 different realizations of a BA scale-free network are generated. The ensemble average $\langle D \rangle$ of the value of the network diameter is plotted as a light blue line in the Fig. 5, showing a small-world behavior (a logarithmic scaling with N, well visible in the log-lin plot of the inset).

Each of the generated networks is then taken as initial condition for the evolution of our game, following the conditions (a) and (b) described above, until reaching the final, equilibrium state. $\langle D \rangle$ for the reached equilibria is reported as a green line in the Fig. 5, and it is clearly seen that an ultrasmall-world state emerges with $\langle D \rangle = 6$ (a value marked by a horizontal dashed line).

A legitimate objection is that adding links to a graph (and therefore increasing the graph's density) always results in decreasing the network's diameter, and therefore a proper comparison has to be offered to assess the relevance of the obtained results. For this purpose, in all trials we take diligent note of the total number of links added before reaching the equilibrium. Then, we take back the initial condition of the specific trial, and add exactly the same number of links, but this time in a fully random way, i.e., without caring about the fulfillment of the game conditions (a) and (b). The obtained values of $\langle D \rangle$ are reported as a red line in Fig. 5. As expected, the red line is always located below the light blue line, but the remarkable result is that the new network ensemble maintains exactly the same logarithmic scaling with N (once again well visible in the inset), which is destined to depart more and more from the constant value characterizing ultrasmall-world states and emerging at the equilibrium of our game.

Finally, we move to show that the mechanism proposed by us and leading to the emergence of the six degrees of separation is, in fact, perfectly compatible with all major structural properties that are observed in real social networks, and in particular with scale-freeness in the degree distribution and with the presence of prominent and hierarchical clustering features. The former attribute indicates that the distribution of the nodes' degrees scales as $p(k) \sim k^{-\gamma}$ (with $2 < \gamma \leq 3$ in real social networks); the latter implies that the clustering coefficient $c(k)$ of a connectivity class k (the average clustering coefficient of all nodes with a given degree k) does depend on k as $c(k) \sim k^{-\omega}$ [42].

To that purpose, we repeat the same extensive simulations which lead us to obtain the results reported in Fig. 5, but this time we adopt as initial conditions for each trial networks that are originated by means of the procedure described in Ref. [43], which indeed provides graphs endowed with degree distributions $p(k) \sim k^{-3}$, with a very
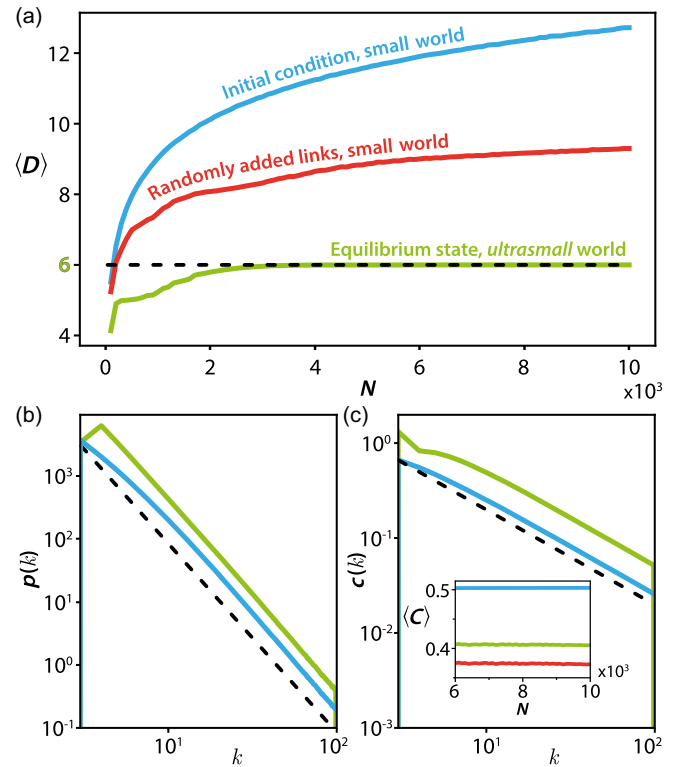

FIG. 6.   Scale-free distribution and hierarchical clustering. (a) Ensemble average $\langle D \rangle$ versus N for the three considered ensembles of networks. Light blue line: networks generated by the procedure of Ref. [43] and that are used as initial conditions for the evolution of the game. Green line: the equilibrium network states of the game. Red line: networks constructed by randomly adding to the initial condition of each game the same number of links needed to reach the game equilibrium. A horizontal black dashed line is positioned at $\langle D \rangle = 6$. (b) The degree distribution $p(k)$ versus k for the set of initial conditions (light blue line) and the set of reached equilibria (green line). $N = 10\,000$. For visibility, the green line plotting $p(k)$ at the equilibria has been slightly vertically shifted. The black dashed line reports the scaling $p(k) \sim k^{-3}$. (c) The hierarchical clustering coefficient $c(k)$ (see text for definition) versus k for the set of initial conditions (light blue line) and the set of reached equilibria (green line). $N = 10\,000$. For visibility, the green line plotting $c(k)$ at the equilibria has been slightly vertically shifted. The black dashed line reports the scaling $c(k) \sim k^{-1}$. The small inset reports the global clustering coefficients $\langle C \rangle$ versus N for the three considered ensembles.

high clustering value ($c \sim 0.5$ for an average degree of $\langle k \rangle = 6$), and with a hierarchical structure of the clustering described by $c(k) \sim k^{-1}$.

Once again, for each value of N, an ensemble of 10 000 different networks are synthesized by the technique of Ref. [43], and each of the generated networks is taken as initial condition for the evolution of the game, until reaching the equilibrium state. In each trial, moreover, note is taken of the total number of links added before reaching the equilibrium, and a network is constructed, for

comparison, by randomly adding exactly the same number of links to the used initial condition.

The results are reported in Fig. 6. Precisely, Fig. 6(a) clearly shows that the scenario obtained is identical to that of Fig. 5: the values of $\langle D \rangle$ averaged over the ensemble of the initial conditions (light blue line) and over the constructed set of networks with randomly added links (red line) are both scaling logarithmically with N, while the reached equilibria (green line) are ultrasmall-world states with $\langle D \rangle = 6$ (marked by a horizontal dashed line).

Figures 6(b) and 6(c) compare, instead, the structures of the initial conditions and that of the reached equilibria, for $N = 10\,000$, and one immediately sees a very remarkable fact: *all other* structural properties imprinted in the initial conditions are conserved in the final state. Precisely, Fig. 6(b) [Fig. 6(c)] reports the degree distribution $p(k)$ [the clustering coefficient $c(k)$] for the light blue case corresponding to the used initial conditions and for the green case corresponding to the reached equilibria, and one immediately sees that the scaling $p(k) \sim k^{-3}$ [$c(k) \sim k^{-1}$], highlighted by a black dashed line, is fully preserved within the range 10^0–10^2 of the degree, i.e., across 2 orders of magnitude, and with minimal differences occurring only at larger degrees due to the addition of the new links that create a few new hubs at the equilibrium. For visibility, $p(k)$ [$c(k)$] of the equilibria has even been multiplied by 2, in order to shift the line in the Fig. 6 panels (b) and (c), because otherwise there would be an almost complete overlap between the light blue and the green curves.

In the small inset in Fig. 6(b), the values of the global clustering coefficients are reported versus N for the three ensembles. One sees that the addition of links in the process of relaxation to equilibrium leads to a slight decrease of $\langle C \rangle$ (from $\langle C \rangle \sim 0.5$ to $\langle C \rangle \sim 0.42$) which, however is maintained to a level pointing to the presence of very prominent and important clustering features. However, the most remarkable trait here is that the value of $\langle C \rangle$ at equilibria is *larger* than that pertinent to the ensemble of networks constructed by randomly adding to the initial conditions the same number of links needed to reach the equilibria.

III. DISCUSSION AND OUTLOOK

The compensation of costs and benefits is certainly a natural interaction strategy through which rational agents determine their connections [44–49], and therefore our study contributes to the understanding of why the six degrees of separation is such a ubiquitous property across vastly different social networks. It is, moreover, reasonable to assume that a similar evolutionary principle may also apply to the design of manmade or technological networks [50]: take, for instance air or sea transportation networks [51–53], in which airports or ports may increase their volume of trades and/or tourism industry by "being in between" the main routes of interchanges of passengers and goods, and in doing so they are keen to incur the relative costs of maintaining (or even enlarging) the number of local connections.

On the other hand, the units of biological networks are in general not rational agents, and it is not straightforward to argue that benefits in terms of betweenness centrality shaped, for instance, the structure of metabolic, genetic, or brain networks along their million-year-long evolutionary path [54–57]. However, one cannot rule out that other compensation mechanisms could have played a pivotal role in this case too, with different benefit functions (e.g., resilience to random perturbations or failures [58,59], or local or global efficiency [60]) recouping for the cost to form or maintain a specific adjacency structure. In neural structures, for instance, it is well known that the functional gains associated with link formation must offset the associated structural costs [61–63]. Note that, for neural structures, while this principle holds in general at evolutionary and developmental timescales, it may also take place at much shorter scales, comparable to those of social network dynamics.

Finally, our study also sheds light on the so-called strength of weak ties phenomenon. This concept was introduced by Granovetter who showed that the most common way of finding a new job is through personal contacts with distant acquaintances, and not via close friends, as one would instead have expected [64,65]. Distant acquaintances represent links connecting different groups of people, and therefore provide each individual with a unique way to receive useful information about distant groups.

Formally speaking, weak ties are links connecting nodes that were originally located at rather large distances and they are therefore called *bridges* or *local bridges* (see the discussion and references in Chap. 3 of Ref. [66]). Their importance for social interaction and communication is strongly supported by a wide range of studies [67,68].

The formation of links connecting nodes from 2-independent sets as the key to the emergence of the six degrees of separation describes exactly the case of a local bridge formation, i.e., a *weak tie* in Granovetter's sense. Therefore, our model can also be viewed as the game theoretical foundation for the strength of weak ties phenomenon.

---

[1] D. J. Watts and S. H. Strogatz, *Collective Dynamics of 'Small-World' Networks*, Nature (London) **393,** 440 (1998).

[2] M. Gurevitch, *The social structure of acquaintanceship networks*, Ph.D. thesis, Massachusetts Institute of Technology, 1961.

[3] I. de Sola Pool and M. Kochen, *Contacts and Influence*, Soc. Networks **1,** 5 (1978).

[4] S. Milgram, *The Small World Problem*, Psychol. Today **2,** 60 (1967), http://snap.stanford.edu/class/cs224w-readings/milgram67smallworld.pdf.

[5] J. Travers and S. Milgram, *An Experimental Study of the Small World Problem in Social Networks* (Elsevier, New York, 1977), pp. 179–197.

[6] P. S. Dodds, R. Muhamad, and D. J. Watts, *An Experimental Study of Search in Global Social Networks*, Science **301,** 827 (2003).

[7] J. Leskovec and E. Horvitz, *Planetary-Scale Views on a Large Instant-Messaging Network*, in *Proceedings of the 17th international conference on World Wide Web* (2008), pp. 915–924.

[8] K. Sanderson, *Six degrees of messaging*, Nature (London) (2008), 10.1038/news.2008.670.

[9] R. Bakhshandeh, M. Samadi, Z. Azimifar, and J. Schaeffer, *Degrees of Separation in Social Networks*, in *Proceedings of the International Symposium on Combinatorial Search* (AAAI Press, 2011), Vol. 2, 10.1609/socs.v2i1.18200.

[10] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna, *Four Degrees of Separation*, in *Proceedings of the 4th Annual ACM Web Science Conference* (Association for Computing Machinery, New York, NY, 2012), pp. 33–42.

[11] R. B. Myerson, *Graphs and Cooperation in Games*, Math. Oper. Res. **2,** 225 (1977).

[12] M. O. Jackson and A. Wolinsky, *A Strategic Model of Social and Economic Networks*, J. Econ. Theory **71,** 44 (1996).

[13] M. A. Nowak and R. M. May, *Evolutionary Games and Spatial Chaos*, Nature (London) **359,** 826 (1992).

[14] H. Ebel and S. Bornholdt, *Coevolutionary Games on Networks*, Phys. Rev. E **66,** 056118 (2002).

[15] M. G. Zimmermann, V. Eguíluz, and M. San Miguel, *Coevolution of Dynamical States and Interactions in Dynamic Networks*, Phys. Rev. E **69,** 065102(R) (2004).

[16] M. G. Zimmermann and V. Eguíluz, *Cooperation, Social Networks and the Emergence of Leadership in a Prisoner's Dilemma with Local Interactions*, Phys. Rev. E **72,** 056118 (2005).

[17] J. M. Pacheco, A. Traulsen, and M. A. Nowak, *Coevolution of Strategy and Structure in Complex Networks with Dynamical Linking*, Phys. Rev. Lett. **97,** 258103 (2006).

[18] J. Gómez-Gardeñes, M. Campillo, Y. Moreno, and L. M. Floría, *Dynamical Organization of Cooperation in Complex Networks*, Phys. Rev. Lett. **98,** 108103 (2007).

[19] T. Gross and B. Blasius, *Adaptive Coevolutionary Networks: A Review*, J. R. Soc. Interface **5,** 259 (2008).

[20] H, Ohtsuki, C. Hauert, E. Lieberman, and M. A. Nowak, *A Simple Rule for the Evolution of Cooperation on Graphs and Social Networks*, Nature (London) **441,** 502 (2006).

[21] S. Goyal and F. Vega-Redondo, *Network Formation and Social Coordination*, Games Econ. Behavior **50,** 178 (2005).

[22] F. Bloch and M. O. Jackson, *The Formation of Networks with Transfers among Players*, J. Econ. Theory **133,** 83 (2007).

[23] J. Poncela, J. Gómez-Gardeñes, A. Traulsen, and Y. Moreno, *Evolutionary Game Dynamics in a Growing Structured Population*, New J. Phys. **11,** 083031 (2009).

[24] X. Bei, W. Chen, S.-H. Teng, J. Zhang, and J. Zhu, *Bounded Budget Betweenness Centrality Game for Strategic Network Formations*, Theor. Comput. Sci. **412,** 7147 (2011).

[25] P. Holme and G. Ghostal, *Dynamics of Networking Agents Competing for High Centrality and Low Degree*, Phys. Rev. Lett. **96,** 098701 (2006).

[26] R. Cohen and S. Havlin, *Scale-Free Networks Are Ultrasmall*, Phys. Rev. Lett. **90,** 058701 (2003).

[27] C. I. N. Sampaio Filho, A. A. Moreira, R. F. S. Andrade, H. J. Herrmann, and J. S. Andrade, *Mandala Networks: Ultra-Small-World and Highly Sparse Graphs*, Sci. Rep. **5,** 1 (2015).

[28] J. F. Nash, *Equilibrium Points in n-Person Games*, Proc. Natl. Acad. Sci. U.S.A. **36,** 48 (1950).

[29] M. A. Nowak, *Five Rules for the Evolution of Cooperation*, Science **314,** 1560 (2006).

[30] D. G. Rand, S. Arbesman, and N. A. Christakis, *Dynamic Social Networks Promote Cooperation in Experiments with Humans*, Proc. Natl. Acad. Sci,. U.S.A. **108,** 19193 (2011).

[31] D. G. Rand and M. A. Nowak, *Human Cooperation*, Trends Cognit. Sci. **17,** 413 (2013).

[32] F. Fu, C. E. Tarnita, N. A. Christakis, L. Wang, D. G. Rand, and M. A. Nowak, *Evolution of In-Group Favoritism*, Nat. Sci. Rep. Ochanomizu Univ. **2,** 460 (2012).

[33] F. Fu, M. A. Nowak, N. A. Christakis, and J. H. Fowler, *The Evolution of Homophily*, Nat. Sci. Rep. Ochanomizu Univ. **2,** 845 (2012).

[34] L. C. Freeman, *A Set of Measures of Centrality Based on Betweenness*, Sociometry **40,** 35 (1977).

[35] J. F. Fink and M. S. Jacobson, *On n-Domination, n-Dependence and Forbidden Subgraphs*, Graph Theory with Applications to Algorithms and Computer Science (John Wiley & Sons, Inc. USA, 1985), pp. 301–311, 10.5555/21936.25447.

[36] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevX.13.021032 for all details of the theorematic proofs which are referred to in the main text.

[37] A. L. Barabási and R. Albert, *Emergence of Scaling in Random Networks*, Science **286,** 509 (1999).

[38] B. Bollobás and O. M. Riordan, *Mathematical Results on Scale-Free Random Graphs*, Handbook of Graphs and Networks: From the Genome to the Internet (John Wiley & Sons, Ltd., 2003), pp. 1–34, 10.1002/3527602755.ch1.

[39] A. Flaxman, A. Frieze, and T Fenner, *High Degree Vertices and Eigenvalues in the Preferential Attachment Graph*, Internet Math. **2**, 1 (2005).

[40] M. E. Newman, *Scientific Collaboration Networks. II. Shortest Paths, Weighted Networks, and Centrality*, Phys. Rev. E **64**, 016132 (2001).

[41] U. Brandes, *A Faster Algorithm for Betweenness Centrality*, J. Math. Sociol. **25**, 163 (2001).

[42] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A. L. Barabási, *Hierarchical Organization of Modularity in Metabolic Networks*, Science **297**, 1551 (2002).

[43] S. Boccaletti, D.-U. Hwang, and V. Latora, *Growing Hierarchical Scale-Free Networks by Means of Nonhierarchical Processes*, Int. J. Bifurcation Chaos Appl. Sci. Eng. **17**, 2447 (2007).

[44] V. M. Eguíluz, M. G. Zimmermann, C. J. Cela-Conde, and M. S. Miguel, *Cooperation and the Emergence of Role Differentiation in the Dynamics of Social Networks*, Am. J. Sociol. **110**, 977 (2005).

[45] M. Perc and A. Szolnoki, *Coevolutionary Games—A Mini Review*, BioSystems **99**, 109 (2010).

[46] M. O. Jackson, *Social and Economic Networks* (Princeton University Press, Princeton, NJ, 2010).

[47] N. A. Christakis and J. H. Fowler, *Social Contagion Theory: Examining Dynamic Social Networks and Human Behavior*, Stat. Med. **32**, 556 (2013).

[48] A. Nishi, H, Shirado, D. G. Rand, and N. A. Christakis, *Inequality and Visibility of Wealth in Experimental Social Networks*, Nature (London) **526**, 426 (2015).

[49] U. Alvarez-Rodriguez, F. Battiston, G. F. de Arruda, Y. Moreno, M. Perc, and V. Latora, *Evolutionary Dynamics of Higher-Order Interactions in Social Networks*, Nat. Hum. Behav. **5**, 586 (2021).

[50] P. Holme and J. Saramäki, *Temporal Networks*, Phys. Rep. **519**, 97 (2012).

[51] R. Guimera, S. Mossa, A. Turtschi, and L. N. Amaral, *The Worldwide Air Transportation Network: Anomalous Centrality, Community Structure, and Cities' Global Roles*, Proc. Natl. Acad. Sci. U.S.A. **102**, 7794 (2005).

[52] H. Barbosa, M. Barthelemy, G. Ghoshal, C. R. James, M. Lenormand, T. Louail, R. Menezes, J. J. Ramasco, F. Simini, and M. Tomasini, *Human Mobility: Models and Applications*, Phys. Rep. **734**, 1 (2018).

[53] W. Lei, L. G. Alves, and L. A. N. Amaral, *Forecasting the Evolution of Fast-Changing Transportation Networks Using Machine Learning*, Nat. Commun. **13**, 1 (2022).

[54] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A. L. Barabási, *The Large-Scale Organization of Metabolic Networks*, Nature (London) **407**, 651 (2000).

[55] D. S. Bassett and O. Sporns, *Network Neuroscience*, Nat. Neurosci. **20**, 353 (2017).

[56] I. Zwir, C. Del-Val, M. Hintsanen, K. M. Cloninger, R. Romero-Zaliz, A. Mesa, J. Arnedo, R. Salas, G. F. Poblete, E. Raitoharju, O. Raitakari, L. Keltikangas-Järvinen, G. A. de Erausquin, I. Tattersall, T. Lehtimäki, and C. R. Cloninger, *Evolution of Genetic Networks for Human Creativity*, Mol. Psychiatry **27**, 354 (2022).

[57] E. A. Tripp, F. Fu, and S. D. Pauls, *Evolutionary Kuramoto Dynamics*, Proc. R. Soc. B **289**, 20220999 (2022).

[58] R. Albert, H. Jeong, and A. L. Barabási, *Error and Attack Tolerance of Complex Networks*, Nature (London) **406**, 378 (2000).

[59] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, *Resilience of the Internet to Random Breakdowns*, Phys. Rev. Lett. **85**, 4626 (2000).

[60] F. Morone and H. A. Makse, *Influence Maximization in Complex Networks through Percolation*, Nature (London) **524**, 65 (2015).

[61] S. Ramón y Cajal, *Histologie du Systéme Nerveux de l'Homme et des Vertébrés* (1909); English translation as *Histology of the Nervous System of Man and Vertebrates*, translated by N. Swanson and L. Swanson (Oxford University Press, New York, 1995).

[62] E. Bullmore and O. Sporns, *The Economy of Brain Network Organization*, Nat. Rev. Neurosci. **13**, 336 (2012).

[63] P. Sterling and S. Laughlin, *Principles of Neural Design* (MIT Press, Cambridge, MA, 2015).

[64] M. S. Granovetter, *The Strength of Weak Ties*, Am. J. Sociol. **78**, 1360 (1973).

[65] M. S. Granovetter, *Getting a Job: A Study of Contacts and Careers* (University of Chicago Press, Chicago, 2018).

[66] D. Easley and J. Kleinberg, *Networks, Crowds, and Markets: Reasoning about a Highly Connected World* (Cambridge University Press, Cambridge, England, 2010).

[67] R. S. Burt, *Structural Holes: The Social Structure of Competition* (Harvard University Press, Cambridge, MA, 1992).

[68] R. S. Burt, *Structural Holes and Good Ideas*, Am. J. Sociol. **110**, 349 (2004).