Supplementary information for "Network Bypasses Sustain Complexity"

Ernesto Estrada,* Lucas Lacasa
Institute for Cross-Disciplinary Physics and Complex Systems (IFISC)
CSIC-UIB, Palma de Mallorca, Spain

July 13, 2022

*Corresponding author: estrada@ifisc.uib-csic.es

S1 Network Communicability $G_{ij}(\beta)$

S1.1 Network communicability across the disciplines

The communicability function emerges in a variety of physical scenarios: it was shown in [1] that $G_{ij}(\beta)$ corresponds to the thermal Green's function of a network of coupled quantum harmonic oscillators where β represents the inverse temperature of the thermal bath in which the network is submerged. The self-communicability $G_{ii}(\beta)$ also appears in the definition of the probability of finding a network, represented by a tight-binding Hamiltonian, in a state with energy $E_j = -\lambda_j$ when the system has inverse temperature equal to β [2]. More recently, Lee et al. [3] found that the communicability function appears in the solution of a linearised upper bound to the susceptible-infected (SI) model when certain initial conditions are impossed to the epidemic dynamics. Similar results were extended by Bartesaghi and Estrada [4] for the SIS model. Finally, the communicability emerges as the solution of a linear autonomous system, which is a modification of the Kuramoto model, when the internal frequency of the oscillators is zero [5]. Nowadays communicability is used for the analysis of the structure and dynamics on brain networks, granular material, infrastructural systems, social networks, biological patterns, genomic and proteomic systems, among others (see references in the main article).

S1.2 Derivation of the communicability function in a concrete setting

For illustration, here we present the derivation of the communicability function and subsequent operators in case for the dynamics of a particle hopping on a network following the tight-binding Hamiltonian:

$$\mathcal{H} \coloneqq \sum_{v=1}^{n} \varepsilon_v \left|v\right\rangle \left\langle v\right| + \sum_{v \neq w}^{n} A_{vw} t_{mn} \left|w\right\rangle \left\langle v\right|, \tag{S1}$$

where $\mathbf{A} = \{A_{vw}\}_{v,w=1}^{n}$ is the network's adjacency matrix. In the initial value representation, the probability amplitude for a transition from state (node) v at time zero to state (node) w at time t is given by $G_{vw}(t) = \langle\psi_w| e^{-it\mathcal{H}/\hbar} |\psi_v\rangle$. Applying a Wick rotation $t \to i\beta\hbar$, the real-time propagator $e^{-it\mathcal{H}/\hbar}$ is mapped into the thermal propagator $e^{-\beta\mathcal{H}}$, where β is the inverse temperature of a thermal bath in which the network is submerged into. Setting $\varepsilon_v = 0$ for every node and $t_{mn} = -1$ for every pair of nodes, we identify $\mathcal{H} = -\mathbf{A}$ and the propagator reduces to the communicability function $G(\beta) = e^{\beta\mathbf{A}}$.

Then, for a particle located at the node v at time zero, the difference:

$$R_{vw}(\beta) \coloneqq \langle\psi_v| e^{\beta\mathbf{A}} |\psi_v\rangle - \langle\psi_w| e^{\beta\mathbf{A}} |\psi_v\rangle, \tag{S2}$$

accounts for the *resistance* offered by the network to the displacement of the particle to the node w at inverse temperature β. The smallest the values of $R_{vw}(\beta)$ the highest the probability that the particle does not get trapped at the origin and can propagate to node w. One can symmetrize this quantity to define the following operator:

$$\xi_{vw}^2(\beta) \coloneqq R_{vw}(\beta) + R_{wv}(\beta) = \langle\psi_v - \psi_w| e^{\beta\mathbf{A}} |\psi_v - \psi_w\rangle. \tag{S3}$$

Mathematically, one can prove that ξ_{vw} is an Euclidean distance between the corresponding nodes (see SI section S1.4). Conceptually, ξ_{vw} is indeed a measure of the network opposition to the flow between v and w. As a matter of fact, and as we will show below (SI section S1.3), ξ_{vw} is indeed formally equivalent to the so-called effective resistance found between a pair of nodes v and w in an electric circuit, which is expressed as $\langle \psi_v - \psi_w | L^* | \psi_v - \psi_w \rangle$, where L^* is the Moore-Penrose pseudoinverse of the circuit's Laplacian matrix. By analogy, we call ξ_{vw} the (thermal) resistance between nodes v and w in the network.

S1.3 Why we call ξ_{vw} a resistance

We start by considering a simple, undirected and connected network $\mathcal{G} = (V, E)$ as an electric circuit [8] in which every edge $(v, w) \in E$ is a resistor of $r_{vw} = 1$ Ohm. The resistance of the whole network when a voltage is connected across a pair of nodes is known as the effective resistance between these two nodes. If a voltage is connected between the nodes v and w and the net current out of source v and into the sink w is designated by I, tthe effective resistance between these nodes is given by

$$\Omega_{vw} = \frac{V_v - V_w}{I}, \tag{S4}$$

where V_i is the potential associated with node i. The following result is proved in the literature [9, 10].

Theorem 1. *The effective resistance Ω_{vw} between nodes v and w satisfies*

$$\Omega_{vw} = \left(e_v - e_w\right)^T L^+ \left(e_v - e_w\right), \tag{S5}$$

where e_i is a standard orthonormal basis, where

$$e_i\left(v\right) \coloneqq \begin{cases} 1 & \text{if } i = v, \\ 0 & \text{otherwise}, \end{cases}$$

and L^+ is the Moore-Penrose pseudoinverse of the graph Laplacian $L = K - A$, with K being a diagonal matrix of node degrees and A the adjacency matrix of $\mathcal{G}$.

Therefore, the effective resistance between a pair of nodes can be written as:

$$\varOmega_{vw} = L^{+}_{vv} + L^{+}_{ww} - 2L^{+}_{vw}. \tag{S6}$$

Notice at this point the formal analogy between $\sqrt{\Omega_{vw}}$ and $\xi_{vw} = \sqrt{G_{vv} + G_{ww} - 2G_{vw}}$ (see the main article).

Furthermore, the matrix of effective resistances of a graph $\mathcal{G}$ can be obtained as:

$$\Omega = diag\left(L^{+}\right)1^{T} + 1\left(diag\left(L^{+}\right)\right)^{T} - 2L^{+}. \tag{S7}$$

Now we use the following result from [11]:

Theorem 2. *Let D be a spherical Euclidean distance matrix (EDM). Then, there is a unique Laplacian matrix L_D such that*

$$D^{-1} = -\frac{1}{2}L_D + \frac{1}{\chi}ff^{T}, \tag{S8}$$

where $f = D^{-1}1$ and $\chi = 1^{T}D^{-1}1$.

Let $\varXi = [\xi_{vw}]_{n\times n}$ be the communicability distance matrix. This is indeed is a nonsingular spherical EDM [12]. Then, we can define the following unique Laplacian:

$$\mathcal{L} = 2\left(\frac{1}{\chi}ff^{T} - \varXi^{-1}\right), \tag{S9}$$

such that

$$\varXi = diag\left(\mathcal{L}^{+}\right)1^{T} + 1\left(diag\left(\mathcal{L}^{+}\right)\right)^{T} - 2\mathcal{L}^{+}, \tag{S10}$$

which clearly indicates that ξ_{vw} identifies with an effective resistance for a weighted graph whose Laplacian matrix is $\mathcal{L}$.

S2 ξ_{vw} is an Euclidean distance

Theorem 3. *Let $\xi^2_{vw} := \left(e^{A}\right)_{vv} + \left(e^{A}\right)_{ww} - 2\left(e^{A}\right)_{vw}$, where A is the adjacency matrix of a graph. Then, ξ_{vw} is an Euclidean distance between the nodes v and w.*

Proof. Let $A = U\Lambda U^T$ where the jth column of U is ψ_j and $\Lambda = diag\left(\lambda_j\right)$. Then,

$$\begin{aligned} \xi^2_{vw} &= \sum_{j=1}^{n} e^{\lambda_j}\psi^2_{jv} + \sum_{j=1}^{n} e^{\lambda_j}\psi^2_{jw} - 2\sum_{j=1}^{n} e^{\lambda_j}\psi_{jv}\psi_{jw} \\ &= \sum_{j=1}^{n} e^{\lambda_j}\left(\psi_{jv} - \psi_{jw}\right)^2 \end{aligned} \tag{S11}$$

Let us define φ_v to be the vth column of U^T, such that we can write

$$\xi^2_{vw} = \left(\varphi_v - \varphi_w\right)^T e^{\Lambda}\left(\varphi_v - \varphi_w\right), \tag{S12}$$

then, because e^{Λ} is positive definite we can write

$$\xi^2_{vw} = \left(e^{\Lambda/2}\varphi_v - e^{\Lambda/2}\varphi_w\right)^T\left(e^{\Lambda/2}\varphi_v - e^{\Lambda/2}\varphi_w\right). \tag{S13}$$

Let us designate $x_v := e^{\Lambda/2}\varphi_v$, such that we have

$$\begin{aligned} \xi^2_{vw} &= \left(x_v - x_w\right)^T\left(x_v - x_w\right) \\ &= \left\|x_v - x_w\right\|^2, \end{aligned} \tag{S14}$$

which means that $\xi_{vw} = \sqrt{\left\|x_v - x_w\right\|^2}$ is a Euclidean distance between v and w. Therefore, x_v and x_w are the coordinates of the nodes v and w in such Euclidean space. $\square$

S3 How communicability and effective distance change with network rewiring: a microscopic analysis

Let us consider the effective length of a path P between the nodes i and j as

$$\mathbb{L}_{P(i,j)} = \sum_{e_k \in P}\xi_{e_k} = \left(\sum_{e_k \in P}\xi^2_{e_k} + 2\sum_{e_k \neq e_l}\xi_{e_k}\xi_{e_l}\right)^{1/2}, \tag{S15}$$

where e_k is an edge in the path P, where the second equality is used for convenience and simply uses the fact that $a + b = \sqrt{a^2 + b^2 + 2ab}$ (which can always be made since ξ is a distance and thus non-negative).

The first term inside the square root is given by

$$\sum_{e_k \in P} \xi_{e_k}^2 = G_{ii} + 2\sum_{s \in P} G_{ss} + G_{jj} - 2 \sum_{(r,t) \in e_k \in P} G_{rt}, \tag{S16}$$

where i and j are the endpoints of the path P, s are intermediate nodes in the path and (r, t) are every adjacent pairs of nodes belonging to the corresponding path.

Then, in order to drop the effective length of the path P relative to another path connecting the same pair of nodes i and j –for instance, relative to the SP– we can act in any of the following two cases:

1. Dropping $\sum_{s \in P} G_{ss}$ to any node in P relative to the similar sum in the reference path. This can be done by dropping the number of small subgraphs in which nodes in the path P participates. For instance, we can delete edges or triangles attached to any node in the path P.

2. Increasing $\sum_{(r,t) \in e_k \in P} G_{(rt)}$ to any pair of nodes forming an edge in P relative to the similar sum in the reference path. This can be done by increasing the communication capacity of pairs of nodes in the path P. For instance, by dropping the length of the cycles in which the two nodes are involved or by increasing the number of paths between pairs of nodes in the path P.

Using only this contribution we can see how the two growing mechanisms studied in this paper may affect these two cases. The rewiring process of the WS mechanism can influence the two cases analyzed in the following way. The rewiring may remove edges attached to nodes which are in the path under consideration (Case 1), at the same time such rewires can create shortcuts that increases the communicability between alternative paths (Case 2). Additionally, the rewiring may attach edges to pairs of nodes in the SP and also can drop the communicability between others which are in the SP. In the case of the BA model, in must be remarked that the addition of a node connected to the nodes of the network will not drop the self-communicaility of any node (Case 1 is not possible). Therefore, such mechanism can drop the effective length of a path by increasing the communicability between pairs of nodes in the path or by increasing the self-communicability of nodes in the SP.

The second term in Eq.S15 represents the interaction between pair of edges in the graph, where each individual term can be written as:

$$\xi_{e_k}\xi_{e_l} = \left(\sum_{s\in e_k, r\in e_l} G_{ss}G_{rr} - 2 \sum_{s\in e_k,(r,t)\in e_l} G_{ss}G_{rt} + 2 \sum_{(p,q)\in e_k,(r,t)\in e_l} G_{pq}G_{rt} \right)^{1/2}, \qquad \text{(S17)}$$

where the first term is the sum of the product of self-communicabilities of nodes s and r, where the two nodes are always in different edges, the second term is the product of the self-communicability of a node in one edge with the communicability between a pair of nodes in another edge, and the last term is the product of the communicability between two nodes in one edge and the one between the two nodes in another. This second term in the expression of $\mathbb{L}_{P(i,j)}$ indeed introduces nonlinearities in the influence of the self-communicability and internode communicability on the change of the length of a path after a given transformation. For, instance, we have seen that dropping $\sum_{s\in P} G_{ss}$ will make the path less resistant, according to the first term of the expression. This dropping also contributes to drop $\sum_{s\in e_k, r\in e_l} G_{ss}G_{rr}$, which makes the path less resistant, but it contributes to increase the effective length of the path P via the term $2\sum_{s\in e_k,(r,t)\in e_l} G_{ss}G_{rt}$. A similar analysis is possible with the communicability G_{rt}, whose increment drops the term $2\sum_{s\in e_k,(r,t)\in e_l} G_{ss}G_{rt}$, but increases $2\sum_{(p,q)\in e_k,(r,t)\in e_l} G_{pq}G_{rt}$. As a consequence of the different types of contributions emerging in both terms, there will be a trade off between the self-communibility and the internode communicability, such that changing them in the appropriate direction will drop the effective length of the path P relative to the SP only up to certain point after which the effective length will increase.

S4 Derivation of the bypass consolidation criterion and measures of impact of bypass consolidation in network navigability

S4.1 Bypass consolidation criterion

Take two arbitrary nodes s and t, and denote $\mathcal{S}_{(s,t)}$ the set of all shortest paths connecting s and t (this set has at least one element, but it might have many more, and that is typically the case when the network is large and the nodes s,t are distant, topologically speaking). Let us denote $\mathcal{P}_{(s,t)}$ the set of potential bypasses. This consists of all paths connecting s and t except the set of shortest paths

$$\mathcal{P}_{(s,t)} = \{\mathfrak{p}(s \to t)\} \setminus \mathcal{S}_{(s,t)}.$$

We say that any given path $\mathfrak{p}(s \to t)$ is a consolidated bypass if

$$\mathbb{L}_{\mathfrak{p}(s\to t)} < \min_{\text{SP}(s,t)\in\mathcal{S}_{(s,t)}} [\mathbb{L}_{\text{SP}(s,t)}]. \tag{S18}$$

It is clear that, for a given pair of nodes (s,t), applying Eq.S18 for all elements in $\mathcal{P}_{(s,t)}$ can yield no solutions, one solution, or multiple solutions. To make the quantification simpler, instead of looking for the total number of consolidated bypasses of (s,t), we only consider a specific path, i.e. the one which has the smallest effective length, i.e. the SRP. In this way, we say that the shortest path(s) are bypassed if

$$\mathbb{L}_{SRP(s,t)} < \min_{\text{SP}(s,t)\in\mathcal{S}_{(s,t)}} [\mathbb{L}_{\text{SP}(s,t)}]. \tag{S19}$$

Observe that, by definition of the Shortest Resistive Path, for any shortest path we always have

$$\mathbb{L}_{SRP(s,t)} \leq \mathbb{L}_{SP(s,t)},$$

and at the same time, by definition of the Shortest Path(s), we always have

$$\ell_{SP(s,t)} \leq \ell_{SRP(s,t)}. \tag{S20}$$

These two properties imply that if Eq.S19 is satisfied, then necessarily we have as a consequence that

$$\ell_{SRT(s,t)} > \ell_{SP(s,t)}. \tag{S21}$$

In fact, if $\ell_{SRT(s,t)} > \ell_{SP(s,t)}$ then it follows that $\mathbb{L}_{SRP(s,t)} < \min_{\mathrm{SP}(s,t)\in\mathcal{S}_{(s,t)}}[\mathbb{L}_{\mathrm{SP}(s,t)}]$ (the other alternative would imply that the SRT is indeed a shortest path and thus not a bypass, resulting in a contradiction). Accordingly, SRT(s,t) is a consolidated bypass if and only if Eq.S21 holds. In fact the same criterion holds to assess if any $\mathfrak{p}(s \to t)$ is a consolidated bypass, but as previously mentioned we won't be analysing the abundance of bypasses, only whether for a given s, t, one consolidated bypass exists. It is trivial to see that checking the SRT is a sufficient condition, so we focus on the SRT.

S4.2 Quantification of bypass effect

Once we have a principled way of assessing if the SRT is a consolidated bypass, we wish to characterize its effect on the navigability of the network. There are two independent aspects one can look at.

First, one can assess how larger is the topological distance of the SRT with respect to the SP, i.e. how small the ratio $\frac{\ell_{SP(s,t)}}{\ell_{SRP(s,t)}} < 1$ is. From this we can define the *topological length excess* $\epsilon_{(s,t)}$

$$\epsilon_{(s,t)} = \left(1 - \frac{\ell_{\mathrm{SP}(s,t)}}{\ell_{\mathrm{SRP}(t,t)}}\right) \cdot 100, \tag{S22}$$

which indicates that, for a particle traveling between two arbitrary nodes i and j, the choice of consolidated bypass SRP over the SP, while beneficial, leads to an *apparent* excess of $\epsilon_{(s,t)}\%$ from the topological distance travelled via the shortest path. Since Eq.S22 is a ratio, then it can be properly averaged across all pairs to construct an average topological length excess

$$\epsilon = \frac{2}{N(N-1)} \sum_{s<t}^{N} \epsilon_{(s,t)}, \tag{S23}$$

that quantifies the average effect of bypasses across all pairs of nodes in terms of the resulting increase in topological length.

Second, we can ask ourselves why would a walker prefer to take a path of, say, topological length $\ell = 9$, than another of, say, topological length $\ell = 3$? Of course this is the case if the first path has smaller effective distance (the walker will much more likely be jammed in the second route, or can easily get lost using that other route, etc), but this information does not seem to be easily interpreted from Eqs.S22 and S23. So to some extent knowing that a walker still prefers a path of topological length three times larger gives also rough idea of the underlying (hidden) resistive geometry. Let us explain this: For the sake or argument, let us suppose that we have detected a consolidated bypass SRP with a certain effective length $\mathbb{L}_{SRP} < \mathbb{L}_{SP}$, and consider two hypothetical situations: (1) in the first one we would have $\ell_{SRP(s,t)} \gg \ell_{SP(s,t)}$, and (2) in the second, $\ell_{SRP(s,t)}$ is only marginally larger than $\ell_{SP(s,t)}$. Clearly, in the first situation the walker navigating from s to t choosing SRP would need to traverse a proportionally much larger amount of edges than by choosing the SP than in the second situation. Since, we need to recall, this is utterly beneficial (because SRP is at the end of the day a consolidated bypass of the SP), this necessarily means that not only the effective length of the SRP is *smaller* than the effective length of the SP (true by definition of the consolidated bypass), but furthermore, it means that the ratio between the effective length *per edge* in the SRP and in the SP is notably smaller in the first situation than in the second. In formulas, the average effective length per edge in the SRP and SP are

$$\langle \mathbb{L}_{SRP(s,t)} \rangle = \frac{\mathbb{L}_{SRP(s,t)}}{\ell_{SRP(s,t)}}; \;\; \langle \mathbb{L}_{SP(s,t)} \rangle = \frac{\mathbb{L}_{SP(s,t)}}{\ell_{SP(s,t)}}$$

and then the ratio

$$\rho = \frac{\langle \mathbb{L}_{SRP(s,t)} \rangle}{\langle \mathbb{L}_{SP(s,t)} \rangle} \tag{S24}$$

quantifies the net benefit (reduction) in effective distance per edge in choosing SRP over SP. ρ is thus the correct measure to compare (and also to average) such benefit for different bypasses. Unfortunately, evaluating Eq.S24 is only computationally efficient for small networks, and in general is not scalable (see however next section for a concrete illustration). Interestingly, by using the properties of SRP and SP we can show that

$$\rho = \frac{\mathbb{L}_{SRP(s,t)}}{\mathbb{L}_{SP(s,t)}} \frac{\ell_{SP(s,t)}}{\ell_{RSP(s,t)}} \leq \frac{\ell_{SP(s,t)}}{\ell_{RSP(s,t)}} \leq 1. \tag{S25}$$

In words, ρ that quantifies how much more beneficial it is –in terms of effective distance travelled per edge– to travels through the SRP than through the SP, and this gain is bounded by $\ell_{\text{SP}(s,t)}/\ell_{\text{TRSP}(s,t)}$. This upper bound is computationally efficient to work out for large networks and in fact, we have already computed it in the bypass consolidation criterion and in Eqs.S22 and S23 ! Using this upper bound as a conservative approximation of ρ, it turns out that the *topological length excess* $\epsilon_{(s,t)}$ can now be interpreted as a lower bound of the *effective distance gain per edge* of using RSP over SP, and likewise ϵ is the network's *average effective distance saving* in the navigability that bypasses have globally provided.

S5 A concrete example

Let us put the formalism in action in a concrete example, a network of $n = 9$ nodes illustrated in Fig.S1. In particular, P_1 is the shortest path connecting the nodes p and q (highlighted in red), and has topological length $\ell_{P_1} = 3$. This path however includes two 'hubs', proned to be jammed, and there exists an alternative route (P_2, highlighted in yellow). P_2 is a potential bypass with larger topological distance $\ell_{P_2} = 4$. P_2 is indeed a consolidated bypass since $\mathbb{L}_{P_1} \approx 5.09$ and $\mathbb{L}_{P_2} \approx 4.80$, so $\mathbb{L}_{P_2} < \mathbb{L}_{P_1}$: while P_1 is topologically shorter, in units of the hidden resistive

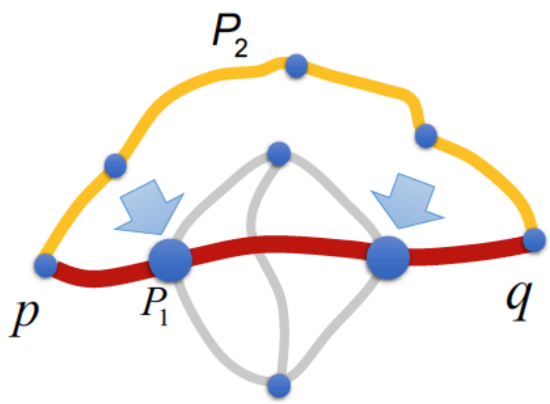

Figure S1: **Explicit example** Cartoon network where P_2 (highlighted in yellow) is a consolidated bypass of the shortest path P_1 (highlighted in red) since P_1 has smaller effective length than P_2, despite of having a longer topological length. This happens because the shortest path involves some nodes with higher degree, which could be proned to jamming and are sources where a walker could diffuse out.

geometry, if is effectively longer! The topological length excess is $\epsilon_{(p,q)} = 100(1-3/4) = 25\%$, i.e. the walker will appear to be choosing a path with 25% more links than the shortest path. Now, the effective distance per edge ratio

$$\rho = \frac{\mathbb{L}_{P_2}}{\ell_{P2}} \cdot \frac{\ell_{P1}}{\mathbb{L}_{P_1}} \approx 0.7073 < 0.75 = \frac{\ell_{P1}}{\ell_{P2}},$$

In words, by actually choosing P_2, the walker will only travel in each link a distance which is about 70% of the actual (effective) distance travelled via P_1. This is less than 75%, the bound offered by the topological ratio. So the net saving of taking P_2 instead of the shortest path P_1 is indeed $100(1 - 0.7073) \approx 30\%$, but our lower bound offers a conservative $\epsilon = 25\%$.

Putting this example in a general context, this helps to illustrate the fact that by choosing the SRP over the SP, the walkers are reducing the effective length of their walk, therefore enhancing navigability. Our estimation of much this is enhanced (ϵ) is only a conservative lower bound of the actual number.

S6 Super-hubs and the Umbrella effect

As we have seen the preferential attachment method produces significantly more bypasses than a random interconnection of nodes. The question that emerges is if there is a simple mechanism which produces even more bypasses than those introduced by the skew degree distribution of the BA method. Let us consider a sparse connected random graph G (see Fig. S2). The chances that there are some bypasses in G are scarce, so we can consider that such number is nill. As the graph G is connected there is a SP between every pair of nodes. Now, let us create G' by adding a new node connected to every node of G. This new node v is a super-hub, which is connected to $n' - 1$ nodes, where n' is the number of nodes in G'. Therefore, a particle trying to go from one node $i \neq v$ to another $j \neq v$ will avoid going through the super-hub as this will be a very resistive trajectory. However, the particle can travel by using the SP connecting them in G. For instance, it can go through the path: $1-2-5$, $1-2-5-6$, $1-2-5-4$, $1-2-5-4-3$ instead of through the paths $1-2-j$. Consequently, all those paths which are of length larger than 2 become potential bypasses by the effect of the newly added super-hub. Due to the shape of the graph illustrated in Fig. S2 we call this the "umbrella effect". In closing, this mechanism potentially transforms all SP of length larger than 2 in a sparse graph into bypasses.

The umbrella effect is expected to occurs in the BA mechanism when the mean degree k_{mean} of the network is large. That is, suppose that at certain "time" t of the BA evolutionary mechanism, the number of existing nodes n_t is smaller than $\langle k \rangle$. This frequently happens at the initial stages of the evolution of BA networks when $\langle k \rangle$ is relatively large. In these cases, the new nodes attached to the network need to be connected to every existing node in the graph, producing an umbrella effect during these periods of the process. The resulting effect could be an explosion of bypasses in the final BA network when $\langle k \rangle$ is relatively large [7].

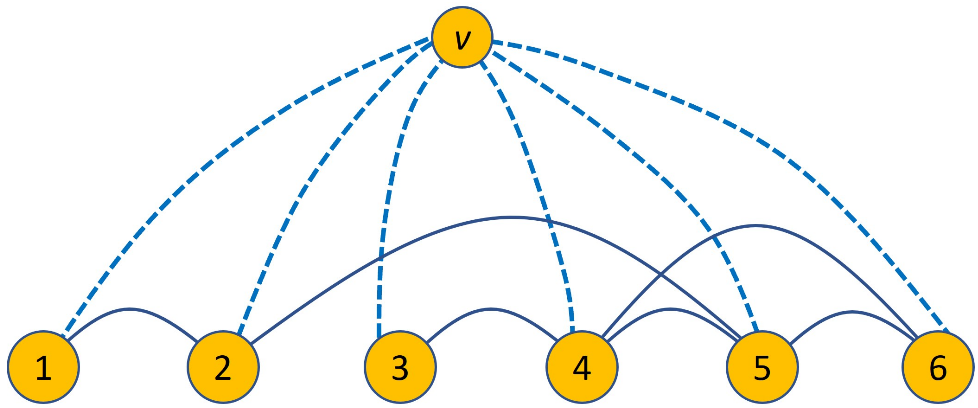

Figure S2: Illustration of the umbrella effect. Initially we have a simple sparse connected network G with 6 nodes and 6 edges. A new graph G' is created by adding a node connected to every existing node of G. This new super-hub makes all previous paths in the original network to be bypasses of the shortest paths, that now involve the super-hub. Adding this super-hub makes ϵ to experience a sudden increase, what we call the umbrella effect.

S7 Effect on Dynamics

S7.1 Synchronization

The synchronizability of a network of coupled oscillators is governed by the Laplacian eigenratio μ_{max}/μ_2, where μ_{max} and μ_2 are the largest and smallest (non-null) eigenvalues of the network's Laplacian matrix, such that the smaller such eigenratio, the more synchronizable the network is. In panel (a) of Fig. S3 we illustrate the change of such eigenratio with the rewiring probability p in the Watts-Strogatz model. Interestingly, we find a qualitative change of behavior at $p^* \approx p_{\text{GNP}}$, where for $p < p^*$ the eigenratio scales as $\mu_{\text{max}}/\mu_2 \sim p^{-\gamma}$, with $\gamma \approx 1.153$, whereas for $p > p^*$ the eigenratio's behavior smoothly crossovers to a flat dependence. To derive the crossover value, we compute the sample standard deviation

$$s = \sqrt{\sum_{i=1}^{n(p^*)} (y_i - f(x_i))^2/(n(p^*) - 1)} \tag{S26}$$

between the actual eigenratio and the scaling function above in the range $p < p^*$ and find the value p^* which minimises such error (see inset of panel (a) in Fig. S3). We find that the error in

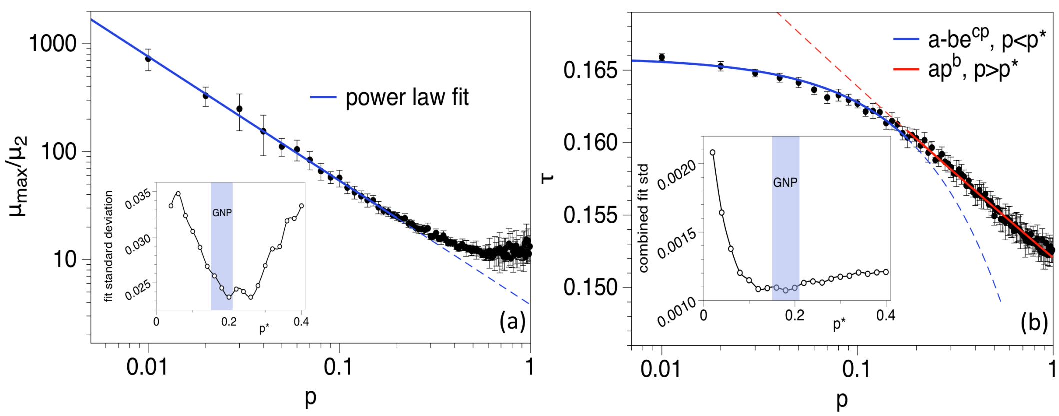

Figure S3: **Dynamics.** (a) Plot of the eigenratio $\mu_{\max}/\mu_2$ of the Laplacian matrix versus the rewiring probability p in a WS model with $N = 500$ nodes and average degree $k = 6$. The straight line corresponds to the power law decay of the eigenratio $\sim p^{-1.15}$. Note that the power law decay is a very good fit only in the range $0 < p < p^*$, thereafter smoothly crossovering to a flat dependence. (Inset) Sample standard deviation of the power law fit as a function of the upper bound p^*. This fitting error is minimized close to $p^* \approx p_{\text{GNP}}$, i.e. the crossover between a power law decay of the eigenratio and a flat dependence on p takes place at the network's Good Navigational Point, where bypasses have a maximal effect. (b) Plot of the epidemic threshold $\tau = 1/\lambda_1$ as a function of the rewiring probability p in the WS model with $N = 500$ nodes and average degree $k = 6$ (dots and error bars denote average and standard deviation over 100 realizations). The blue and red lines correspond to two qualitatively different fits, for $p < p^*$ (blue) and $p > p^*$ (red). (Inset) To find the value of p^* that marks the crossover between scaling behaviors, we compute the aggregate fit deviation of the two-function model, finding that such deviation reaches a minimum for $p^* \approx p_{\text{GNP}}$, i.e. the change of behavior onsets at the regime where bypasses take a predominant role.

the fit is minimised for $p^* \approx 0.2$, indeed close to p_{GNP}.

S7.2 Epidemic spreading

The second process is a SIR-type epidemic spreading over the network, where α is the birth rate of the epidemic and δ is its death rate. It is well known that when $\alpha/\delta < \tau$ infection dies out and when $\alpha/\delta > \tau$ infection survives and becomes an epidemic, where τ is the epidemic threshold (the smallest τ, the easiest that an infection outbreak becomes epidemic). For undirected networks $\tau = 1/\lambda_1$, where λ_1 is the spectral radius of the adjacency matrix. As can be seen in panel (b) of Fig. S3 the epidemic threshold of a Watts-Strogatz network also shows a qualitative change: there is a crossover from a slow decrease for $p < p^*$ to a faster, power-law decrease for $p > p^*$. To find p^*, we compute the sample standard deviation of each branch (eq.S26). The inset of panel (b) shows the combined sample standard deviations, that find a minimum for $p^* \approx p_{\mathrm{GNP}}$. That is, up to the point in which bypasses become predominant, an infection is relatively easier to control than after that point.

Altogether, these results suggest that dynamical processes running on the network perceive such bypasses in a nontrivial way, and these therefore play a role in the network's function.

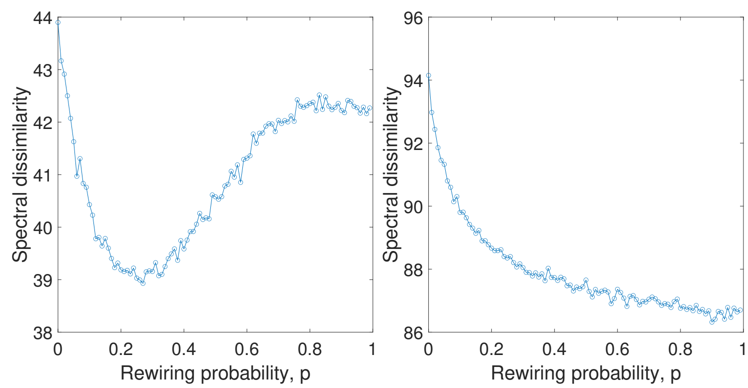

Figure S4: Typical dissimilarity plots obtained in the analysis of empirical networks: we either we find a non-monotonic behavior that highlights an optimal value at intermediate values of p (left panel) or a monotonically decreasing function, for which $p^* = 1$ (right panel). The former suggests a clear SW structure while the latter suggests that the network is closer to a random graph.

S8 Empirical networks

S8.1 Additional details on empirical networks

Below we provide details on all 37 empirical networks, grouped by topic (see also table S1 for a summary of characteristics).

Brain networks

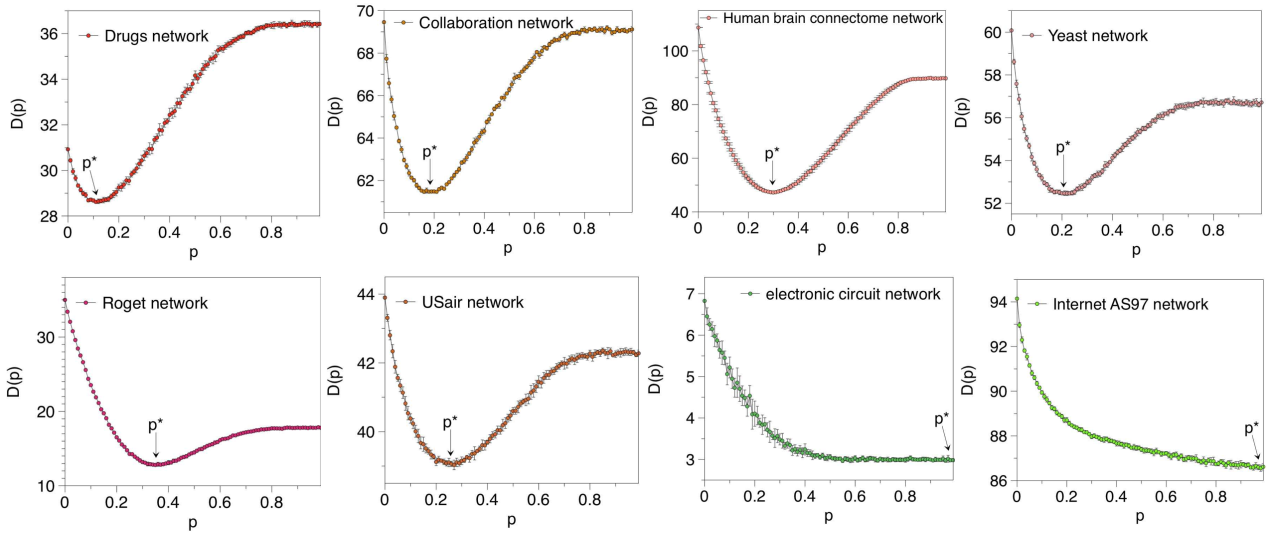

Figure S5: Dissimilarity curves $\mathcal{D}(p)$ as a function of the rewiring probability p, for different empirical networks. The minimum of these curves characterizes the type of SW network which is more similar to the empirical network.

- **Neurons:** Neuronal synaptic network of the nematode *C. elegans*. Included all data except muscle cells and using all synaptic connections [13]
- **Human brain**: task-related activation of human brain regions obtained through a meta-analysis of over $1,600$ studies published between 1985–2010 [14].

Ecological networks

- **Benguela**: Marine ecosystem of Benguela off the southwest coast of South Africa [15]
- **Bridge Brook**: Pelagic species from the largest of a set of 50 New York Adirondack lake food webs [16];
- **Canton Creek**: Primarily invertebrates and algae in a tributary, surrounded by pasture, of the Taieri River in the South Island of New Zealand [17];
- **Chesapeake Bay**: The pelagic portion of an eastern U.S. estuary, with an emphasis on larger fishes [18];

- **Coachella**: Wide range of highly aggregated taxa from the Coachella Valley desert in southern California [19];
- **El Verde**: Insects, spiders, birds, reptiles and amphibians in a rainforest in Puerto Rico [20];
- **Grassland**: all vascular plants and all insects and trophic interactions found inside stems of plants collected from 24 sites distributed within England and Wales [21];
- **Little Rock**: Pelagic and benthic species, particularly fishes, zooplankton, macroinvertebrates, and algae of the Little Rock Lake, Wisconsin, U.S. [22];
- **Reef Small**: Caribbean coral reef ecosystem from the Puerto Rico-Virgin Island shelf complex [23];
- **Scotch Broom**: Trophic interactions between the herbivores, parasitoids, predators and pathogens associated with broom, Cytisus scoparius, collected in Silwood Park, Berkshire, England, UK [24];
- **Shelf**: Marine ecosystem on the northeast US shelf [25];
- **Skipwith**: Invertebrates in an English pond [26];
- **St. Marks**: Mostly macroinvertebrates, fishes, and birds associated with an estuarine seagrass community, Halodule wrightii, at St. Marks Refuge in Florida [27];
- **St. Martin**: Birds and predators and arthropod prey of Anolis lizards on the island of St. Martin, which is located in the northern Lesser Antilles [28];
- **Stony Stream**: Primarily invertebrates and algae in a tributary, surrounded by pasture, of the Taieri River in the South Island of New Zealand in native tussock habitat [29];

- **Ythan1**: Mostly birds, fishes, invertebrates, and metazoan parasites in a Scottish Estuary [30] ;
- **Ythan2**: Reduced version of Ythan1 with no parasites [31].
- **Termite**: The networks of three-dimensional galleries in termite nests [32];

Informational networks

- **Roget**: Vocabulary network of words related by their definitions in Roget's Thesaurus of English. Two words are connected if one is used in the definition of the other [33];

Biological networks

- **Protein-protein interaction networks** in: *S. cerevisiae* (yeast) [34, 35];
- **TransE.coli, Transyeast**: Direct transcriptional regulation between genes in *Saccaromyces cerevisae*. [13, 36].

Social networks

- **Geom**: Collaboration network of scientists in the field of Computational Geometry [38];
- **QcGr** Collaboration network of scientists in the field of Quantum Gravity [39];
- **Drugs**: Social network of injecting drug users (IDUs) that have shared a needle in the last six months [40];

Technological and infrastructural networks

- **Electronic**: Three electronic sequential logic circuits parsed from the ISCAS89 benchmark set, where nodes represent logic gates and flip-flop [13];
- **USAir97**: Airport transportation network between airports in US in 1997 [38];

- **Internet**: The internet at the Autonomous System (AS) level as of September 1997 [41];
- **Power Grid**: The power grid network of the Western USA [42].

Software networks

- Collaboration networks associated with six different **open-source software** systems, which include collaboration graphs for three Object Oriented systems written in C++, and call graphs for three procedural systems written in C. The class collaboration graphs are from version 4.0 of the **VTK** visualization library; the CVS snapshot dated 4/3/2002 of Digital Material (DM), a library for atomistic simulation of materials; and version 1.0.2 of the AbiWord word processing program. The call graphs are from version 3.23.32 of the **MySQL** relational database system, and version 1.2.7 of the **XMMS** multimedia system. Details of the construction and/or origin of these networks are provided in Myers [43].

S8.2 Additional details on calculation of p^*

As can be seen in table 1 of the main article, we find two classes of networks, one class is formed by networks for which a minimum dissimilarity exists between the real-world network and a WS model operating close to p_{GNP} (see Fig. S4, left panel). The other class is formed by networks which are more similar to the random graph, i.e., WS with $p = 1$ (see Fig. S4 right panel). The dissimilarity curve is plotted for some empirical networks in Fig. S5.

S8.3 ϵ and ϵ/ϵ_{BA}: further analysis

Here we give a closer inspection into the values of table 1 of the main article. Overall, we can see that for a groups of networks (not necessarily domain specific), the bypass navigability gain ϵ is similar to the one expected for a BA model with the same number of nodes N and mean degree $\langle k \rangle$, suggesting that the type of heterogeneity is similar and that the contribution

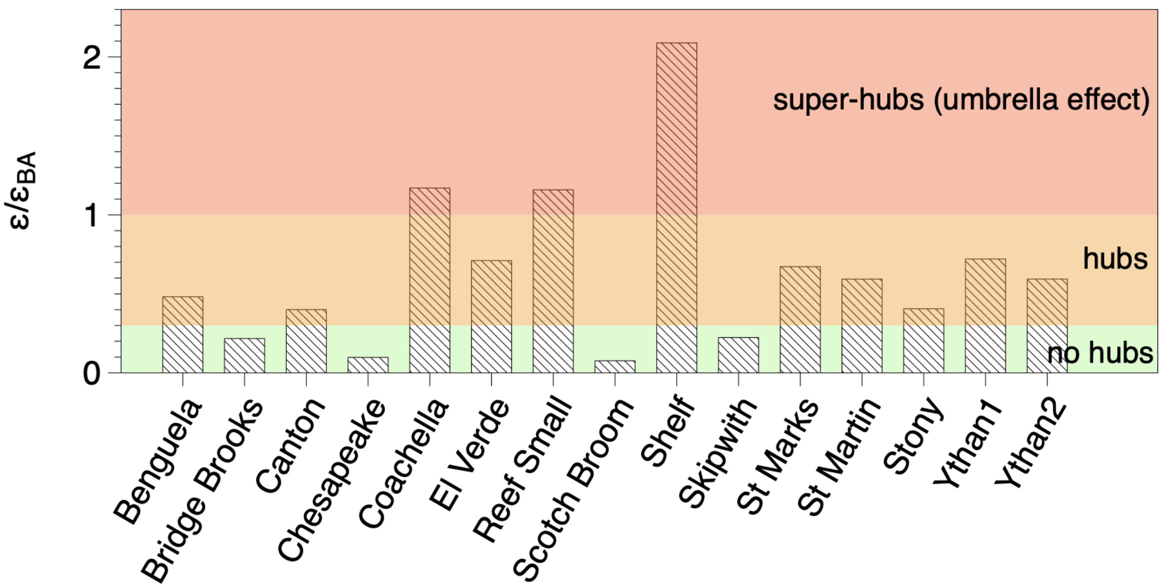

Figure S6: ϵ/ϵ_{BA} for all 15 food webs. We have highlighted three regimes: $\epsilon/\epsilon_{BA} < 0.5$, where the network is essentially not bypassing hubs, $0.5 < \epsilon/\epsilon_{BA} < 1$, where the real network and its analogous BA network have fairly similar associated navigability gain and thus the real network is bypassing hubs, and $\epsilon/\epsilon_{BA} > 1$ where the real network is bypassing hubs stronger than in the BA model. This last situation takes place when super-hubs emerge, i.e. nodes that connect to a large portion of the rest of nodes in the network.

of bypasses is substantial due to the presence of hubs. This is the case, for instance, in the Human brain, collaboration networks, C. Elegans neurons, or USA airports: for all these cases $\epsilon/\epsilon_{BA} > 0.75$.

There is a second group of networks where such ratio is lower, suggesting that either the role of hubs is, at least, less prominent than what we would expect in a BA model with the same characteristics, or that even if hubs. exist, the network is designed to not bypasses. In this group we find a whole spectrum of values ϵ: networks where the role of bypasses is still substantial but considerably less than in a comparable BA model (e.g. protein-protein interaction, with $\epsilon \approx 25$, with $\epsilon/\epsilon_{BA} \approx 0.55$), and networks where the bypasses have only a marginal effect in the network's navigability (e.g. termite mounds, with $\epsilon \approx 3$, with $\epsilon/\epsilon_{BA} \approx 0.12$, or the power

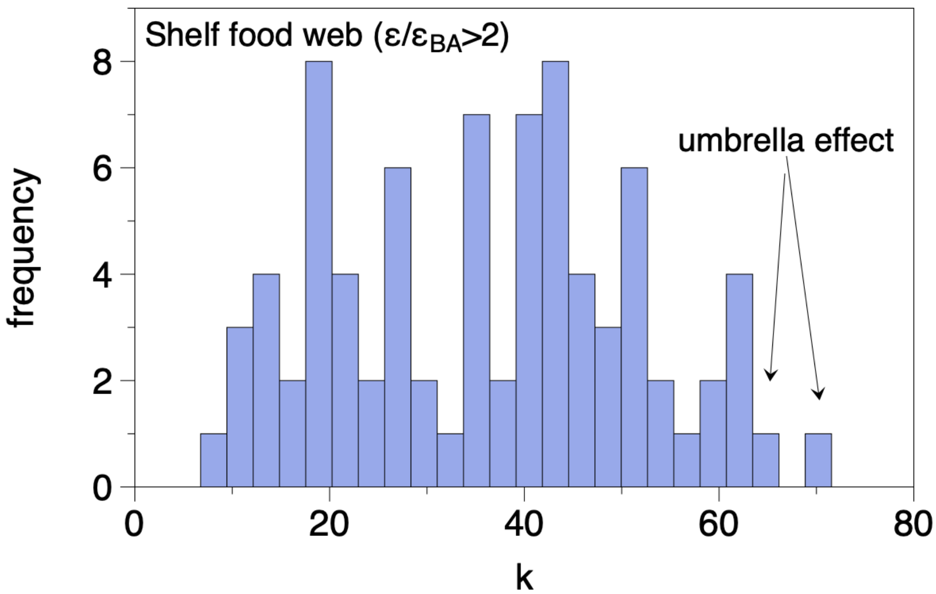

Figure S7: Degree frequency histogram of the Shelf food web (the one with $\epsilon/\epsilon_{BA} > 2$). This network has $N = 81$ nodes and a large number of them have very high degree. Some of these nodes will convert into super-hubs and then be likely to evidence the umbrella effect (see SI section S6) which make the navigability gain ϵ larger than the expected for a BA network.

grid, with $\epsilon \approx 2.6$, with $\epsilon/\epsilon_{BA} \approx 0.05$). The result on the power grid is interesting: originally seen as an example of a scale-free network [44] (albeit with an exponent 4, i.e. much more homogeneous than a BA network), its scale-freeness has been a matter of debate [45]. Nevertheless, it is clear that this network do have hubs. So our measures –that highlights that this network has bypasses have only about 5% of the impact than a respective BA network would have– determine that this network is not engineered to have routes bypassing hubs. This is in agreement to previous evidence on blackouts and cascading failures in such type of networks: these are examples of networks with fat-tailed degree distributions that do not follow an evolutionary process, which would certify the presence of bypasses.

Next, we disaggregate the calculation of ϵ/ϵ_{BA} performed on the 15 food webs. In Fig. S6 we plot the estimate of ϵ/ϵ_{BA} in each of them, where ϵ_{BA} is computed by generating BA models with the same number of nodes and mean degree that the empirical network, and averaging the resulting value of the navigability gain over 100 realizations of the BA model. We find food-webs across three regimes: the regime of low ϵ/ϵ_{BA}, which indicates that the network does not have hubs or that the network has not been designed to bypass these, the regime where $\epsilon \approx \epsilon_{BA}$, which indicates that the network is compatible with a BA model and the navigability gain is therefore associated to hubs being bypassed, and a (new) third regime ϵ/ϵ_{BA}, found for very dense networks, where there are some nodes in the empirical network that likely evidence the umbrella effect. Figure S7 illustrates this fact for the case of the Shelf food web ($\epsilon/\epsilon_{BA} > 2$). In this figure we plot the degree frequency of this food web, showing that this is a notably dense network where we find some nodes with extreme degree (note that this network has $N = 81$ nodes) which are candidates for the umbrella effect.

References

[1] Estrada, E., Hatano, N., & Benzi, M. (2012). The physics of communicability in complex networks. Physics reports, 514(3), 89-119.

[2] Estrada, E., & Hatano, N. (2007). Statistical-mechanical approach to subgraph centrality in complex networks. Chemical Physics Letters, 439(1-3), 247-251.

[3] Lee, C. H., Tenneti, S., & Eun, D. Y. (2019, July). Transient dynamics of epidemic spreading and its mitigation on large networks. In Proceedings of the twentieth ACM international symposium on mobile ad hoc networking and computing (pp. 191-200).

[4] Bartesaghi, P., & Estrada, E. (2021). Where to cut to delay a pandemic with minimum disruption? Mathematical analysis based on the SIS model. Mathematical Models and Methods in Applied Sciences, 31(12), 2571-2596.

[5] Arola Fernandez, L. Synchronization in Complex Networks Under Uncertainty (Doctoral dissertation, Universitat Rovira i Virgili). 2022. Chapter 6.

[6] Bartesaghi, P., Clemente, G. P., & Grassi, R. (2020). Community structure in the World Trade Network based on communicability distances. Journal of Economic Interaction and Coordination, 1-37.

[7] L. Lacasa, E. Estrada, The Umbrella effect: a new structural transition in networks, in preparation

[8] Doyle, P. G., & Snell, J. L. (1984). Random walks and electric networks (Vol. 22). American Mathematical Soc.

[9] Ellens, W., Spieksma, F. M., Van Mieghem, P., Jamakovic, A., & Kooij, R. E. (2011). Effective graph resistance. Linear algebra and its applications, 435(10), 2491-2506.

[10] Klein, D. J., & Randić, M. (1993). Resistance distance. Journal of mathematical chemistry, 12(1), 81-95.

[11] Balaji, R. R. B. B., & Bapat, R. B. (2007). On Euclidean distance matrices. Linear algebra and its applications, 424(1), 108-117.

[12] Estrada, E., Sanchez-Lirola, M. G., & De La Peña, J. A. (2014). Hyperspherical embedding of graphs and networks in communicability spaces. Discrete Applied Mathematics, 176, 53-77.

[13] R. Milo, S. Shen–Orr, S. Itzkovitz, et al., “Network motifs: simple building blocks of complex networks,” Science **298**, 824–827 (2002).

[14] Crossley, N. A., Mechelli, A., Vertes, et al (2013). Cognitive relevance of the community structure of the human brain functional coactivation network. Proceedings of the National Academy of Sciences, 110(28), 11583-11588.

[15] P. Yodzis, "Local trophodynamics and the interaction of marine mammals and fisheries in the Benguela ecosystem," J. Anim. Ecol. **67**, 635-658 (1998).

[16] G. A. Polis, "Complex trophic interactions in deserts: an empirical critique of food-web theory," Am. Nat. **138**, 123-155 (1991).

[17] C. Townsend, R. M. Thompson, and A. R. McIntosh, et al. "Disturbance, resource supply, and food-web architecture in streams," Ecol. Lett. **1**, 200 (1998).

[18] R. R. Christian, and J. J. Luczkovich, "Organizing and understanding a winter's seagrass foodweb network through effective trophic levels," Ecol. Model. **117**, 99-124 (1999).

[19] P. H. Warren,"Spatial and temporal variation in the structure of a fresh-water food web," Oikos **55**, 299-311 (1989).

[20] R. B. Waide, and W. B. Reagan, (Eds.) *The Food Web of a Tropical Rainforest*. (University Chicago Press, Chicago, 1996).

[21] N. D. Martinez, B. A. Hawkins, H. A. Dawah, and B. P. Feifarek, "Effects of sampling efforts on characterization of food web structure," Ecology **80**, 1044–1055 (1999).

[22] K. Havens, "Scale and structure in natural food webs," Science **257**, 1107-1109 (1992).

[23] S. Opitz, "Trophic Interactions in Caribbean coral reefs," ICLARM Tech. Rep. 43, Manila, Philippines, (1996).

[24] J. Memmott, N. D. Martinez, and J. E. Cohen, “Predators, parasites and pathogens: species richness, trophic generality, and body sizes in a natural food web,” J. Animal Ecol. **69**, 1-15 (2000).

[25] J. Link, “Does food web theory work for marine ecosystems?,” Mar. Ecol. Prog. Ser. 230, 1-9 (2002).

[26] P. Yodzis, “Diffuse effects in food webs,” Ecology **81**, 261-266 (2000).

[27] L. Goldwasser, and J. A. Roughgarden, “Construction and analysis of a large Caribbean food web,” Ecology **74**, 1216-1233 (1993).

[28] N. D. Martinez, “Artifacts or attributes? Effects of resolution on the Little Rock Lake food web,” Ecol. Monogr. **61**, 367-392 (1991).

[29] D. Baird, and R. E. Ulanowicz, “The seasonal dynamics of the. Chesapeake Bay ecosystem,” Ecol. Mon. **59**, 329-364 (1989).

[30] M. Huxman, S. Beany, and D. Raffaelli, “Do parasites reduce the chances of triangulation in a real food web?,” Oikos **76**, 284-300 (1996).

[31] S. J. Hall, and D. Rafaelli, “Food-web patterns - lessons from a species-rich web,” J. Anim. Ecol. **60,** 823–842 (1991).

[32] A. Perna, S. Valverde, and J. Gautrais, et al. “Topological efficiency in the three-dimensional gallery networks of termite nests,” Physica A **387**, 6235-6244 (2008).

[33] Roget’s Thesaurus of English Words and Phrases, Project Gutenberg. `http://gutenberg.net/etext/22` (2002).

[34] D. Bu, Y. Zhao, and L. Cai, et al. "Topological structure analysis of the protein-protein interaction network in budding yeast," Nucleic Acids Res. **31**, 2443-2450 (2003).

[35] C. von Mering, R. Krause, and B. Snel, et al. "Comparative assessment of large-scale data sets of protein-protein interactions," Nature **417**, 399-403 (2002).

[36] R. Milo, S. Itzkovitz, and N. Kashtan, et al. "Superfamilies of evolved and designed networks," Science **303**, 1538-1542 (2004).

[37] S. Shen-Orr, R. Milo, S. Mangan, and U. Alon, "Network motifs in the transcriptional regulation network of Escherichia coli," Nature Gen. **31**, 64-68 (2002).

[38] V. Batagelj, and A. Mrvar, Pajek datasets. Available at: http://vlado.fmf.uni-lj.si/pub/networks/data/ (2006).

[39] J. Leskovec, J. Kleinberg and C. Faloutsos. Graph Evolution: Densification and Shrinking Diameters. ACM Transactions on Knowledge Discovery from Data (ACM TKDD), 1(1), 2007.

[40] J. Moody, Data for this project was provided in part by NIH grants DA12831 and HD41877, those interested in obtaining a copy of these data should contact James Moody (moody.77@sociology.osu.edu), (2001).

[41] M. Faloutsos, P. Faloutsos, and C. Faloutsos, "On power-law relationships of the internet topology" Comp. Comm. Rev. **29**, 251-262 (1999).

[42] D. J. Watts, and S. H. Strogatz, "Collective dynamics of small-world networks" Nature **393**, 440-442 (1998).

[43] C. R. Myers, "Software systems as complex networks: Structure, function, and evolvability of software collaboration graphs," Phys. Rev. E **68**, 046116 (2003)

[44] Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science* 286, 5439 (1999), 509-512.

[45] Hartmann, B., & Sugar, V. (2021). Searching for small-world and scale-free behaviour in long-term historical data of a real-world power grid. *Scientific Reports*, 11, 1 (2020): 1-10.

Name	N	m
Benguela	29	191
BridgeBrook	75	542
Canton	108	707
Chesapeake	33	71
Coachella	30	241
Drugs	616	2012
Electronic1	122	189
Electronic2	252	399
Electronic3	512	819
ElVerde	156	1439
Collaboration CoGe	3621	9461
Collaboration QcGr	4158	13428
Internet1997	3015	5156
LittleRockA	181	2318
Neurons	280	1973
PIN_Yeast	2224	6608
Power_grid	4941	6594
ReefSmall	50	503
Roget	994	3640
ScotchBroom	154	366
Shelf	81	1451
Skipwith	35	353
Software_Abi	1035	1719
Software_Digital	150	198
Software_Mysql	1480	4140
Software_VTK	771	1357
Software_XMMS	971	1802
StMarks	48	218
StMartin	44	218
Stony	112	830
Termite_1	507	676
Termite_2	260	280
Termite_3	268	437
Transc_yeast	662	1062
USAir97	332	2126
Ythan1	134	593
Ythan2	92	416

Table S1: Dataset of real-world networks studied in this paper, their size N (number of nodes), and number of edges m.