



# Isotopy and energy of physical networks

Yanchen Liu<sup>1</sup>, Nima Dehmamy<sup>1,2</sup> and Albert-László Barabási<sup>3,4,5</sup>✉

**While the structural characteristics of a network are uniquely determined by its adjacency matrix<sup>1–3</sup>, in physical networks, such as the brain or the vascular system, the network's three-dimensional layout also affects the system's structure and function. We lack, however, the tools to distinguish physical networks with identical wiring but different geometrical layouts. To address this need, here we introduce the concept of network isotopy, representing different network layouts that can be transformed into one another without link crossings, and show that a single quantity, the graph linking number, captures the entangledness of a layout, defining distinct isotopy classes. We find that a network's elastic energy depends linearly on the graph linking number, indicating that each local tangle offers an independent contribution to the total energy. This finding allows us to formulate a statistical model for the formation of tangles in physical networks. We apply the developed framework to a diverse set of real physical networks, finding that the mouse connectome is more entangled than expected based on optimal wiring.**

Physical exclusion imposes severe limitations on most real systems, from granular media<sup>4</sup> to networks<sup>5,6</sup>. For example, if the links of a spatial network are physical objects and are unable to cross one another, in two dimensions only planar networks can exist. While in three dimensions any network can be embedded without link overlap<sup>7</sup>, for a given adjacency matrix a network can have an infinite number of layouts, differing in the positions of its nodes and the precise wiring of its links. Yet, many of these layouts are isotopic<sup>8</sup>, meaning that they can be transformed into one another through continuous topological deformations, without the need to cut a link (Fig. 1b). At the same time, a network can also have multiple non-isotopic embeddings, each defining a distinct isotopy class<sup>8</sup>.

To determine whether two network embeddings are non-isotopic, we start from the linking number, a knot invariant<sup>9,10</sup>, that measures the number of times two closed curves (cycles) wind around each other, capturing the number of tangles (Fig. 1c and Supplementary Fig. 1). We generalize it to define the graph linking number (GLN), which for a network with embedding  $\mathcal{E}$  represents the sum of the linking numbers of all pairs of cycles in the graph

$$\mathcal{G}(\mathcal{E}) = \sum_{c, c' \in \{C\}} |\mathcal{L}(\mathcal{E}; c, c')|, \quad (1)$$

where  $\{C\}$  is the set of cycles in the network, determined only by the adjacency matrix, and  $\mathcal{L}(\mathcal{E}; c, c')$  is the linking number between cycle  $c$  and  $c'$ , determined only by the embedding  $\mathcal{E}$ . Note that  $\mathcal{L}(\mathcal{E}; c, c')$  can be either positive or negative, depending on the orientation we select for  $c$  and  $c'$ , and  $\mathcal{L}(\mathcal{E}; c, c') = 0$  if cycles  $c$  and  $c'$  share nodes (Supplementary Information section SI.1 and

Supplementary Fig. 2). One can show that embeddings with different  $\mathcal{G}(\mathcal{E})$  belong to different isotopy classes (Fig. 1d,f). If, however, two embeddings have the same  $\mathcal{G}(\mathcal{E})$ , they can still be non-isotopic (Fig. 1d,e).

Usually, larger networks have more cycles<sup>11</sup>; hence, we expect more potential tangles between them, which would lead to higher  $\mathcal{G}$  values. To compare the layouts of networks with different sizes, we use the normalized graph linking number (nGLN)

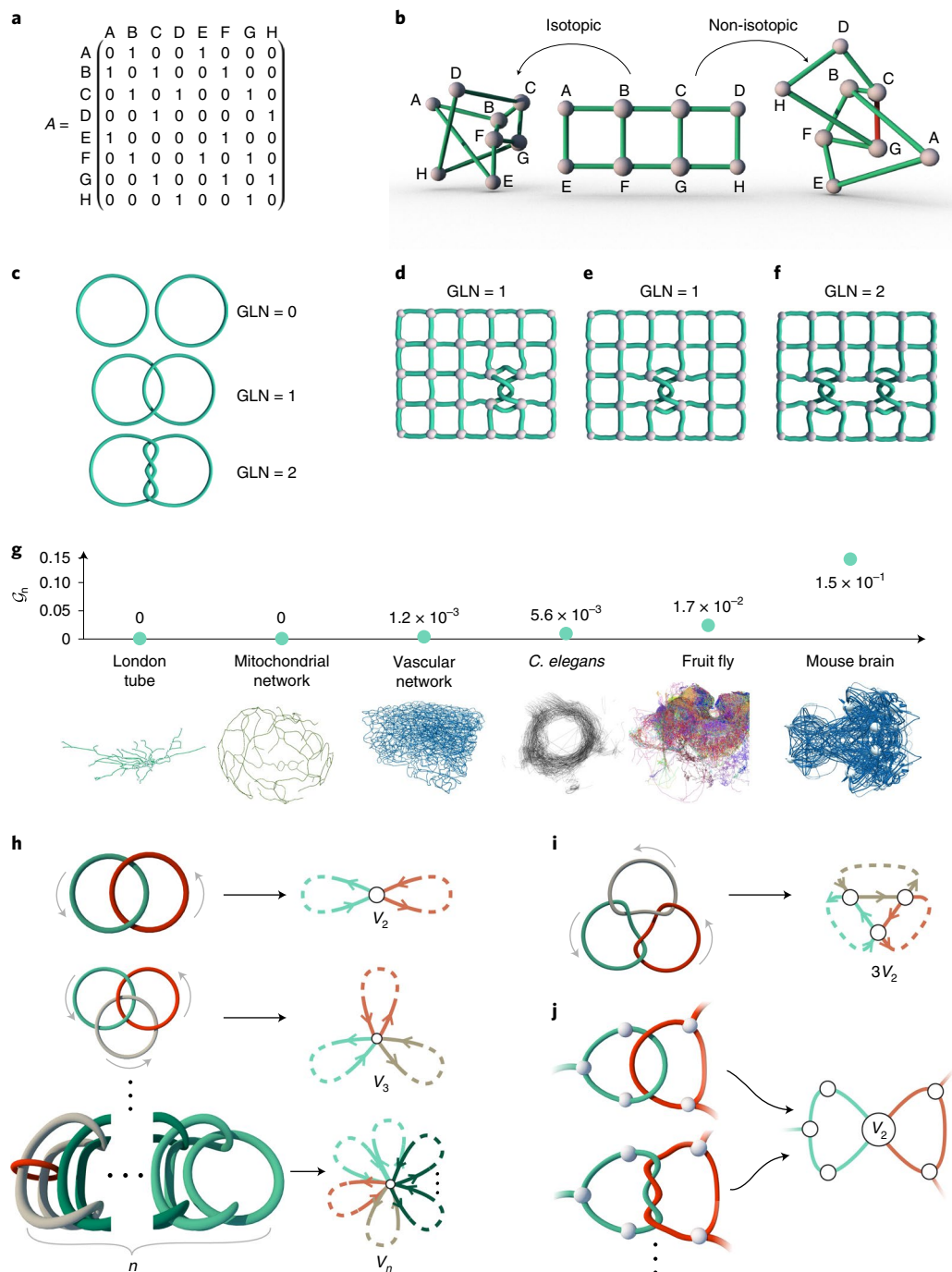
$$\mathcal{G}_n(\mathcal{E}) = \frac{\mathcal{G}(\mathcal{E})}{N_p}, \quad (2)$$

where  $N_p$  is the number of disjoint cycle-pairs in  $\{C\}$  (determined only by the adjacency matrix, Fig. 1a). Note, however, that identifying all loops in equations (1) and (2) is computationally expensive, and as many loops are parts of larger loops, equations (1) and (2) count the same tangles multiple times. Therefore, going forward, we limit equations (1) and (2) to the minimal loop set, representing all shortest loops from which any loop in the network can be constructed (Supplementary Information section SI.1). Finding all minimal loops is still a computationally expensive problem, which prompts us to develop a heuristic method that utilizes spanning trees to sample the minimal loops. We find that the resulting heuristic nGLN  $\mathcal{G}_n$  converges to a linear function of  $\mathcal{G}_n$ , a monotonicity that allows us to consistently compare the tangledness of networks with different sizes (Supplementary Information section SI.1 and Supplementary Figs. 4, 6, 7 and 10).

We collected data on the topology and the physical layout of several real-world physical networks, finding that their different  $\mathcal{G}_n(\mathcal{E})$  reflect differences in their function and structure (Fig. 1g). For example, the London tube network (number of nodes  $N=75$ , number of links  $M=118$ ) has a near-planar structure, ensuring the absence of entanglements and resulting in  $\mathcal{G}_n = 0$ . Interestingly, the mitochondrial network<sup>12</sup> ( $N=26$ ,  $M=27$ ), despite its full three-dimensional (3D) layout, also has  $\mathcal{G}_n = 0$ , due to the fact that it is embedded on a near-spherical surface, turning it into an effectively planar network. The vascular system network's<sup>13</sup> ( $N=1,544$ ,  $M=2,345$ ) need to transport oxygen and nutrients to tissues results in a space-filling network with low  $\mathcal{G}_n = 0.0012$ , reflecting its grid-like structure in three dimensions. The *Caenorhabditis elegans* anterior neuronal network<sup>14</sup> ( $N=31,346$ ,  $M=36,216$ ) and the BL(L) region of the fruit fly brain network<sup>15</sup> ( $N=7,841$ ,  $M=9,497$ ) are connectomes mapped at the single-neuron resolution, whose elongated neuron branches result in tangled layouts with higher  $\mathcal{G}_n$  equal to 0.0056 and 0.017, respectively. The mouse brain network<sup>16</sup> (J. A. Brum et al., manuscript in preparation) ( $N=26$ ,  $M=472$ ) connects brain regions with axon bundles that represent thousands to millions of neurons, resulting in a highly entangled wiring with  $\mathcal{G}_n = 0.15$ . The three orders of magnitude spanned by  $\mathcal{G}_n$  reflect the diversity

<sup>1</sup>Network Science Institute, Northeastern University, Boston, MA, USA. <sup>2</sup>Kellogg School of Management, Northwestern University, Evanston, IL, USA.

<sup>3</sup>Network Science Institute and Department of Physics, Northeastern University, Boston, MA, USA. <sup>4</sup>Division of Network Medicine and Department of Medicine, Brigham and Womens Hospital, Harvard Medical School, Boston, MA, USA. <sup>5</sup>Department of Network and Data Science, Central European University, Budapest, Hungary. ✉e-mail: [a.barabasi@northeastern.edu](mailto:a.barabasi@northeastern.edu)



**Fig. 1 | Graph linking number.** **a, b**, A network with the adjacency matrix  $A$  (**a**) can have multiple physical layouts (**b**). **b**, The first two layouts are isotopic, meaning that we can transform one into the other without cutting links. The third layout is non-isotopic to the first two, as to transform into the previous ones, we need to cut and re-attach the red link. **c**, Linking number in knot theory<sup>9</sup> measures the number of tangles between two loops. The figures show three pairs of loops with GLN = 0, 1 and 2. **d–f**, Non-isotopic layouts of the same 2D lattice with  $G=1$  (**d**), 1 (**e**) and 2 (**f**). Layout (**d**) and layout (**e**) both have one loop pair that has linking number 1, therefore they both have  $G=1$ . However, since the location of the tangle in (**d**) and (**e**) are different, they are non-isotopic. **g**, Examples of several real-world physical networks we analysed, including the London tube network, the mitochondrial network<sup>12</sup>, the mouse brain vascular network<sup>13</sup>, the *C. elegans* anterior neuron network<sup>14</sup>, the fruit fly brain BL(L) network<sup>15</sup> and the mouse brain network<sup>16</sup> (J. A. Brum et al., manuscript in preparation). Note that for the fruit fly, the ‘missing’ part on the right is not included in the dataset for BL(L), possibly because it belongs to another brain region. The plot shows the  $G_n$  value for each network. **h**, Feynman diagram representation of loop tangles. Each loop is like the trajectory of a particle that enters from far away, interacts with other loops and then leaves again. The interactions of two loops are like four-point interactions. Analogously, interactions of three loops are like six-point interactions, and  $n$  loops, like  $2n$ -point interactions. **i**, Three pairwise interactions between three loops. **j**, Pairwise interaction between two loops can include multiple windings.

of processes that influence the entangledness of physical networks, challenging us to develop a theoretical framework that can explain the observed differences.

In physical networks, the links (for example, axons) do not have arbitrary lengths, as they are costly to build and to maintain. We therefore measured the total elastic energy of the

layout, representing the sum of the elastic energies of all links  $l$ , defined as

$$V_{\text{el}}[\{\gamma_l\}] = \sum_l \int_{\gamma_l} \left| \frac{d\mathbf{x}_l}{ds} \right|^2 ds, \quad (3)$$

where the integral is over the path  $\gamma_l$  for link  $l$ . Here  $s \in [0, 1]$  parameterizes the length of a link  $l$  and  $\mathbf{x}_l(s)$  is the location of the segment of link  $l$  from length parameter value  $s$  to  $s + ds$ . To avoid the crossing of links and nodes, we add a short-range node–node repulsion  $V_{\text{NN}} \approx A_N \sum_{ij} \exp[-(|\mathbf{X}_i - \mathbf{X}_j|/2r_N)^p]$  and link–link repulsion  $V_{\text{LL}} \approx A_L \sum_{lm} \int \int ds_l ds_m \exp[-(|\mathbf{x}_l - \mathbf{x}_m|/2r_L)^p]$  with  $p \geq 2$  (ref. <sup>17</sup>), where  $A_N, A_L$  are amplitudes for the potentials,  $\mathbf{X}_i, \mathbf{X}_j$  are location for node  $i, j$ ,  $\mathbf{x}_l, \mathbf{x}_m$  are directed segments on link  $l, m$ ,  $r_N, r_L$  are parameters for node and link interaction ranges, and the exponent  $p$  determines how hard or soft the potentials are. As the total elastic energy increases monotonically with the total link length<sup>17,18</sup>, the energy minima of the energy landscape corresponds approximately to the shortest wiring length embedding<sup>17</sup>. Given the numerous local minima of the energy landscape<sup>19,20</sup>, finding the global energy minimum is a problem of complexity similar to protein folding, an NP-complete (non-deterministic polynomial-time complete) problem<sup>21,22</sup>. Yet, as we show next, a direct relationship between isotopy and energy offers us unique insights about the energy landscape. Indeed, all isotopic layouts can be continuously transformed into one another without link crossings, implying that they belong to the same energy well. As the energy cost of a link crossing is infinite, the energy landscape consists of distinct wells, one for each isotopy class, separated by infinite energy barriers (Fig. 2a).

To unveil the relation between  $\mathcal{G}_n$  and the elastic energy  $V_{\text{el}}$ , we model a statistical ensemble of layouts, assuming a stochastic component characterized by temperature  $T$  in a canonical ensemble. The partition function for this ensemble sums over all possible paths for all links with a fixed set of node locations  $\{X_i\}$ , integrated over all possible node locations

$$Z = \sum_{\{\text{Link paths } \gamma\}} \sum_{\{\text{Node pos.}\}} \exp[-\beta V_{\text{el}}] \\ = \int \prod_{i=1}^N d^3 X_i \prod_{l=1}^{N_L} [d\gamma_l] \exp[-\beta V_{\text{el}}] \quad (4)$$

$$= \int \prod_{i=1}^N d^3 X_i Z_L[\{X_i\}], \quad (5)$$

where  $\beta = 1/T$ ;  $d\gamma_l$  are variations in the path of link  $l$ ;  $N$  and  $N_L$  are the number of nodes and links, respectively; and  $Z_L[\{X_i\}]$  denotes the partition function for all link configurations for fixed node positions  $\{X_i\}$ .

We can rewrite equation (5) in terms of loop interactions within a network. When the link thickness  $r_L$  is much smaller than the node–node repulsion range  $r_N$ , loop interactions can be formally represented as Feynman diagrams (Fig. 1h–j and Supplementary Information section SI.3). Tangles between a pair of loops are captured by  $V_2$  (interaction energy between a pair of loops, Fig. 1i,j), but irreducible  $V_n$  (interaction energy between  $n$  loops) interactions with  $n > 2$  also exist (Fig. 1i). For example, a three-loop interaction (that is, a Borromean ring) contributes to the partition function with a Boltzmann factor  $\exp[-\beta(\|\mathbf{l}_1\|^2 + \|\mathbf{l}_2\|^2 + \|\mathbf{l}_3\|^2 + V_3)]$ , that is,  $\exp[-\beta(\|\mathbf{l}_3\|^2 + V_3 - V_2)]$  smaller than a pairwise interaction (Supplementary Information section SI.3), indicating that higher interactions offer exponentially diminished contributions to the layout's expected energy  $\langle E \rangle$ , compared with pairwise interactions (Supplementary Information section SI.3).

We further simplify the partition function using a mean-field approach, where  $\varepsilon_n \equiv \text{mean}(V_n + \sum_i^n \delta \|\mathbf{l}_i\|^2)$  captures the energy contribution from an  $n$ -loop irreducible interaction  $V_n$  and the extra elastic energies  $\delta \|\mathbf{l}_i\|^2$  needed to stretch loop  $i$

relative to its ground state. Note that there is an infinite number of energy states for  $n$ -loop interactions: pairwise loop interactions include the untangled state (linking number 0, elastic energy  $\varepsilon_0$ ), simple tangle (linking number 1, energy  $\varepsilon_0 + \varepsilon_t$ ,  $\varepsilon_t$  is the energy contribution of a tangle), tangled with  $L$  windings around each other (linking number  $L$ , energy  $\varepsilon_0 + \varepsilon_t + (L-1)\varepsilon_w$ ,  $\varepsilon_w$  is the energy contribution of extra windings), as well as more complex knotted states. Similarly, irreducible higher-order interactions  $V_n$  include infinite numbers of interactions  $V_n^\mu$ , where  $\mu$  represents the distinct topological ways  $n$  loops can get tangled irreducibly, with average energy  $\varepsilon_n^\mu$ . Therefore, for each set of  $n$  loops, the partition function  $z_n = \sum_\mu \exp[-\beta \varepsilon_n^\mu]$  gives expected energy  $\langle \varepsilon_n \rangle = -\frac{1}{z_n} \frac{\partial z_n}{\partial \beta}$ . The partition function for all irreducible  $V_n$  interactions is the product of  $z_n$  over all  $C_n \approx \binom{\text{number of loops}}{n}$  possible choices of  $n$  loops out of all loops in the layout ( $Z_n = \prod_{C_n} z_n = (z_n)^{C_n}$ ), and the full partition function is the product of  $Z_n$

$$Z = e^{-\beta E_0} \prod_n Z_n = e^{-\beta E_0} \prod_n \left( \sum_\mu \exp[-\beta \varepsilon_n^\mu] \right)^{C_n}. \quad (6)$$

Taken together, the average of the energy of the ensemble can be expressed as (Supplementary Information section SI.3.5)

$$\langle E \rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = E_0 + \sum_n C_n \langle \varepsilon_n \rangle, \quad (7)$$

which we can expand to the  $n=2$  term with the expected number of type- $\mu$  pairwise interactions  $\langle N_2^\mu \rangle$

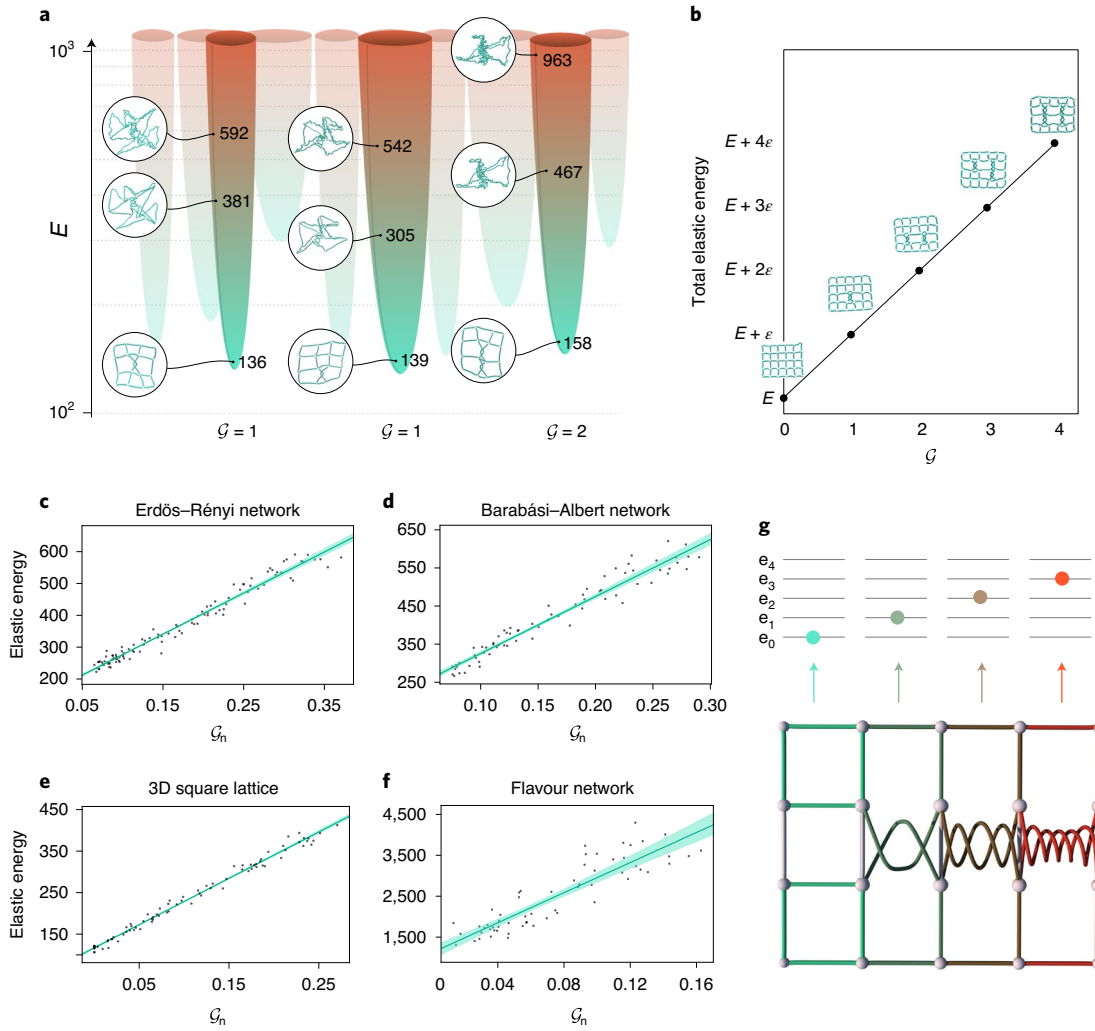
$$\langle E \rangle = E_0 + \langle \mathcal{G} \rangle \varepsilon_w + \langle N_2^t \rangle (\varepsilon_t - \varepsilon_w) + \sum_{\mu \neq t, \{t,w\}} \langle N_2^\mu \rangle \varepsilon_2^\mu + \sum_{n \geq 3} C_n \langle \varepsilon_n \rangle. \quad (8)$$

This indicates that the average graph linking number  $\langle \mathcal{G} \rangle$  determines the lowest-level contribution to the energy. Note, however, that the higher-order contributions decreases greatly the correlation between energy and  $\mathcal{G}$ , and together with the difficulty in detecting irreducible high-order interactions, it is impossible to estimate the minimal energy for an isotopy class of layouts. To move forward, we simulated various network topologies, starting from random layouts, and performing gradient descent to settle into a local minimum, measuring the total energy of the obtained layouts. Surprisingly, as Fig. 2b–f shows, we find that there is a statistically significant approximately linear correlation between the minimum elastic energy and  $\mathcal{G}_n$ , indicating that the higher-order interactions offer diminishing contributions to the total energy ( $\langle \varepsilon_2 \rangle \gg \langle \varepsilon_n \rangle$ ), and for pairwise interactions  $\varepsilon_t \approx \varepsilon_w \approx \varepsilon$ , and  $\langle N_2^\mu \rangle \ll \langle \mathcal{G} \rangle$  for  $\mu \neq t, \{t,w\}$ . Therefore, the total elastic energy of a network layout with  $\mathcal{G}$  is well approximated by

$$E(\mathcal{G}) \approx E_0 + \mathcal{G} \varepsilon, \quad (9)$$

a finding allowing us to continue the analytical treatment by mapping the network layout into a system of non-interacting particles, where the effective degrees of freedom are loop pairs with infinitely many evenly spaced, discrete energy states. A physical network behaves like a canonical ensemble, whose energy fluctuations are determined by the temperature  $T$  of the 'heat bath'. For  $T \rightarrow 0$ , the system converges to the ground-state layout, with close-to-straight links and minimal tangles, whereas at higher  $T$  the links are curved, which leads to additional tangles and higher elastic energy. The likelihood that a network has elastic energy  $E$  follows the Boltzmann distribution

$$P(E(\mathcal{G})) = \frac{n(\mathcal{G}) e^{-\beta E(\mathcal{G})}}{Z}, \quad (10)$$



**Fig. 2 | Energy and GLN.** **a**, Each isotopy class of a network corresponds to a different energy well, as illustrated for a 2D lattice network. The walls of the energy well go to infinity, as due to the prohibition of link crossings, a network cannot be transformed from one isotopy class to another. Next to the wells, we show three isotopic layouts of the same network for each well, together with their linking number (shown as numbers) and total layout energy (left axis). **b**, The relation between the elastic energy and  $\mathcal{G}$  for a 2D lattice. Each tangle stretches the links, increasing the elastic energy by  $\epsilon$  relative to that of their shortest-length states. We find that each loop crossing induces a comparable link length increase, implying a proportionality between energy and  $\mathcal{G}_n$ . **c–f**, The elastic energy versus  $\mathcal{G}_n$  for networks generated by the Erdős–Rényi<sup>30</sup> (**c**) and the Barabási–Albert<sup>31</sup> (**d**) models, for  $N=100$  and average degree  $\langle k \rangle = 6$ , a 3D square lattice ( $N=100$ ) (**e**) and the flavour network ( $N=184$ ,  $\langle k \rangle = 7.8$ ) (**f**)<sup>32</sup>. For each network, the energy grows linearly with  $\mathcal{G}_n$ . The shaded areas indicate the 95% confidence region of the linear fit. **g**, Four pairs of loops highlighted in a 2D lattice, with linking number 0, 1, 2 and 3, respectively. Each loop pair can be mapped into a particle with a different discrete energy. The first loop pair has no tangle, corresponding to a particle at the ground state  $e_0$ ; the second has one tangle, corresponding to a particle at the first excited energy state  $e_1 = e_0 + \epsilon$ ; while the last two correspond to particles at the second and third excited energy states,  $e_2 = e_0 + 2\epsilon$  and  $e_3 = e_0 + 3\epsilon$ .

where  $Z = e^{-\beta E_0} (1 - e^{-\beta \epsilon})^{N_p}$ ,  $\beta = \frac{1}{T}$  and  $n(\mathcal{G}) = \binom{N_p + \mathcal{G} - 1}{\mathcal{G}}$  is the degeneracy of energy state  $E(\mathcal{G})$  (Supplementary Information section SI.3.6), allowing us to derive the expected energy of the layout

$$\langle E \rangle = \frac{N_p \epsilon}{e^{\beta \epsilon} - 1} + E_0, \quad (11)$$

the expected nGLN

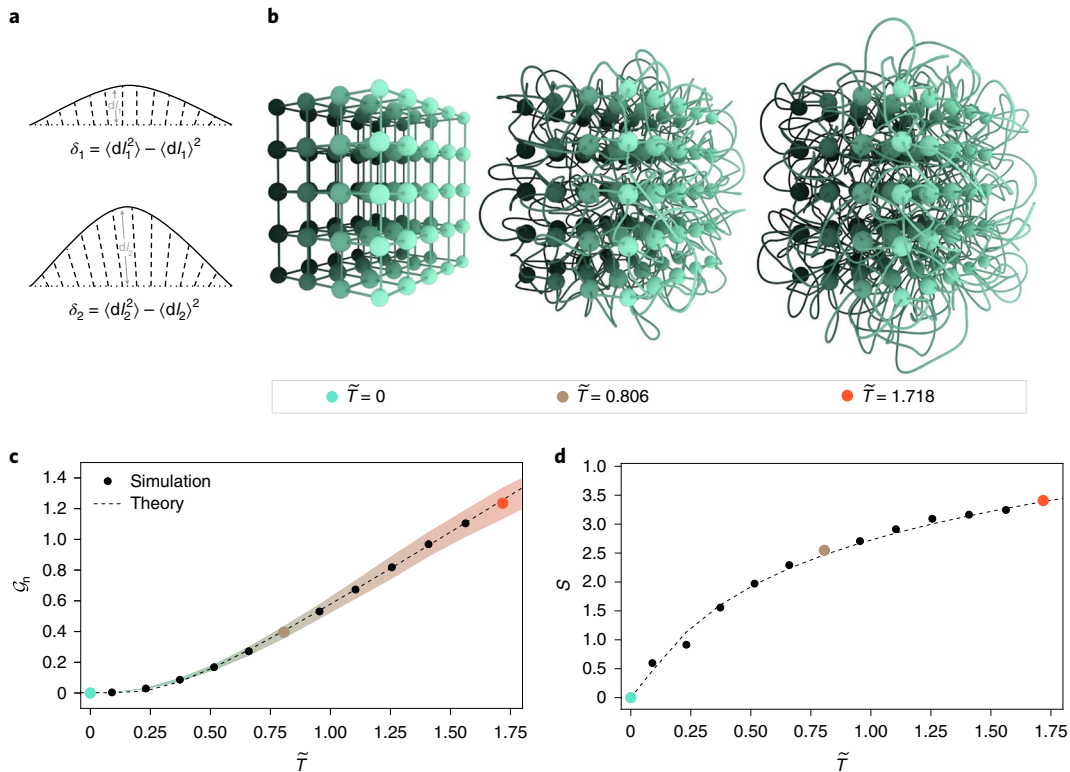
$$\langle \mathcal{G}_n \rangle = \frac{1}{e^{\beta \epsilon} - 1}, \quad (12)$$

and the expected entropy

$$S = -N_p \log(1 - e^{-\beta \epsilon}) + \frac{N_p \beta \epsilon}{e^{\beta \epsilon} - 1}, \quad (13)$$

the latter capturing the variability in the layout. Note that the energy cost of a tangle  $\epsilon$  is a function of the network structure, link thickness and elastic constants, resulting in different tangledness  $\langle \mathcal{G}_n \rangle$  at the same  $T$  for different kinds of network. To ensure that the temperature is independent of such system-specific parameters, we use the rescaled temperature  $\tilde{T} = \frac{T}{\epsilon}$  ( $\beta$  to  $\tilde{\beta} = \beta \epsilon$ ), which guarantees that networks at the same  $\tilde{T}$  have the same degree of entangledness  $\langle \mathcal{G}_n \rangle = 1 / (e^{\tilde{\beta}} - 1)$ . Variations in the layout energy are typically caused by geometric deviations ( $\delta l_i(s)$ ) from the paths of the zero temperature layout ( $\int dl_{0i} = l_{0i}$ ), resulting in new link paths parameterized as  $d\mathbf{l}_i = d\mathbf{l}_{0i} + \delta \mathbf{l}_i(s)$ . We can explicitly relate the layout temperature  $\tilde{T}$  to the deviation  $\delta$  of a link trajectory from straight lines (Fig. 3a and Supplementary Information section SI.4),





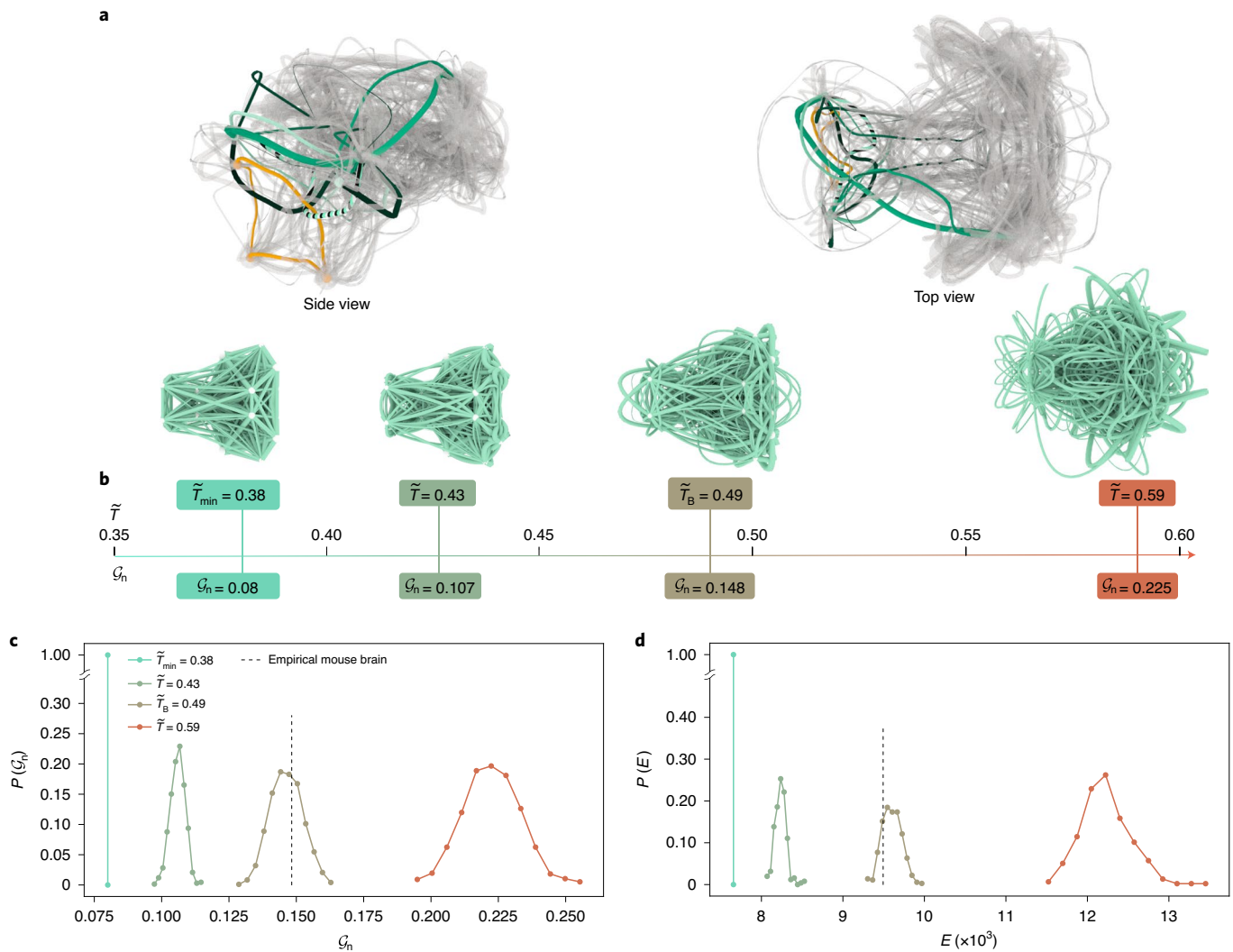
**Fig. 3 | Statistical physics of physical networks.** **a**, The deviation  $\delta$  from the straight link trajectories used in equation (14) is defined as the square of the standard deviation of  $d_l$ , representing the distance of points along the link from their locations on the straight path (dotted line). In the figure, the second link has a larger deviation than the first link ( $\delta_2 > \delta_1$ ), which contributes to the higher temperature of the layout (according to equation (14)). **b**, Three-dimensional lattice layouts with different temperatures, while keeping the node positions fixed, illustrating that higher  $T$  corresponds to more curving of the links. In turn, the more curved links lead to more tangles, resulting in a higher nGLN. **c**, Comparison between the analytic and numerical results for  $\mathcal{G}_n$  for 3D lattice layouts at different temperatures (different canonical ensembles). The dots are the mean  $\langle \mathcal{G}_n(T) \rangle$  at different temperatures (the coloured dots correspond to the lattices in **b**), and the shaded area represents two standard deviations above and below the mean. The dashed line is obtained by fitting the analytical formula (12) to the data. **d**, Analytical and numerical results for the entropy for the canonical ensemble of 3D lattice layouts at different temperatures. The dots are the numerical results, and the dashed line is the analytical result (equation (13)).

$$\tilde{T}(\delta) - \tilde{T}_{\min} = \frac{1}{2 \operatorname{arcsinh}\left(\sqrt{\frac{N_p \epsilon^2}{8kE_0} \frac{1}{\delta}}\right)} = \frac{\delta}{\epsilon \sqrt{\frac{N_p}{8kE_0}}} + O\left(\frac{1}{\delta}\right), \quad (14)$$

connecting the statistical formalism to a physical network's measurable geometric characteristics. To test the validity of the statistical formalism, we generated layouts of 3D lattices with varying temperatures, by varying the deviation  $\delta$  used in simulations (Fig. 3a,b), and extracted  $\tilde{T}$  from the network geometry using equation (14). The higher the temperature, the larger is the deviation  $\delta$  from the straight link trajectories, and the more tangles are observed. Figure 3c,d shows the analytic results for  $\langle \mathcal{G}_n \rangle$  (equation (12)) and entropy (equation (13)), and the numerical values extracted from the simulations (Fig. 3b). The excellent agreement between the analytic and the simulation results confirms that the statistical mechanical model of entanglement (equations (9)–(13)) offers an accurate description of the geometrical layouts of physical networks. Next, we show that this statistical model can offer insights on the layout of a real physical network, which we illustrate in the case of the mouse brain.

Given the high metabolic cost of each neuron, brain wiring is often hypothesized to be minimizing the wiring cost<sup>23</sup>. To quantify the degree of optimality of the mouse connectome, we extracted the 3D map of the axon bundles (links) connecting different brain regions from the Allen Mouse Brain Atlas<sup>16</sup> (J. A. Brum et al., manuscript in preparation). At the highest level of hierarchy, we study a network of  $N = 24$  brain regions<sup>16</sup> (J. A. Brum et al., manuscript in

preparation) connected by  $L = 472$  axon bundles. This natural layout is highly entangled, with  $\mathcal{G}_n = 0.148$  (Fig. 4a), but the number itself does not tell us whether it represents an optimal wiring. To unveil the optimal layout, we fixed the anatomically predetermined node locations (brain regions) and generated the most economical layout with no link crossings, corresponding to the lowest possible temperature  $\tilde{T} = \tilde{T}_{\min}$  in the statistical formalism. This optimal layout has  $\mathcal{G}_n = 0.08$ , roughly half of the real brain's  $\mathcal{G}_n$  (Fig. 4b), suggesting that the layout (red) of the brain substantially deviates from the optimal one. Using the relation (12) between  $\mathcal{G}_n$  and  $\tilde{T}$ , we calculated the true temperature of the brain layout, obtaining  $\tilde{T}_B = 0.49$  (we also estimated the maximum temperature of the mouse brain, limiting the total link length by the limits imposed by the brain size, obtaining  $\tilde{T}_{\max} \approx 0.66$ ). To see whether the temperature  $\tilde{T}_B$  alone determines the observed structural characteristics of the real brain, we generated artificial mouse brain layouts at different temperatures, measuring for each layout nGLN and the energy (Fig. 4c,d). We find that nGLN and energy, as well as the width of their distributions, increase with temperature, consistent with the predictions of the canonical ensemble (Supplementary Information section SI.4). Fixing the temperature at  $\tilde{T}_B$  by setting the link deviation  $\delta$  using equation (14), we created 1,000 alternative layouts of the mouse brain with the same adjacency matrix and node locations but different link geometries. We find that the observed nGLN and energy of the synthetic brain ensemble at  $\tilde{T} = \tilde{T}_B$  are consistent with the values observed for the mouse brain (Fig. 4c,d). The fact



**Fig. 4 | Modelling brain layout.** **a**, The mouse brain is a tangled network, with  $\mathcal{G}_n = 0.148$ . To illustrate the roots of its tangledness, we highlighted the tangles between one specific loop (orange) and four other loops (green to black). **b**,  $\mathcal{G}_n$  and  $\tilde{T}$  for several different layouts of the brain network. All layouts have the same node positions as the real layout shown in **a**, but different link curvatures. The first corresponds to straight links with temperature  $\tilde{T} = 0.38$  and  $\mathcal{G}_n = 0.08$ . The second layout is at  $\tilde{T} = 0.43$ , obtained by adding link deviations from straight paths, and resulting in  $\mathcal{G}_n = 0.107$ . The third layout is obtained by raising the temperature until it matches the brain temperature  $\tilde{T}_B = 0.49$  as the real mouse brain, and has  $\mathcal{G}_n = 0.148$ . The fifth layout is at  $\tilde{T} = 0.59$  with  $\mathcal{G}_n = 0.225$ . **c**, Probability distributions for  $\mathcal{G}_n$  of the brain network at different temperatures, with the  $\mathcal{G}_n$  of the real mouse brain indicated with a dashed line. **d**, Probability distributions for the elastic energy of the mouse brain network at different temperatures, with the energy of the real mouse brain indicated with a dashed line.

that the expected GLN characterizing the ensemble at  $\tilde{T} = \tilde{T}_B$  is largely indistinguishable from the actual brain's nGLN suggests that solely controlling the level of fluctuations results in layouts similar to the brain. These results indicate that the real mouse brain is more tangled than expected for the most economical layout. Most importantly, a single parameter,  $\tilde{T}$ , extracted from the wiring through the developed statistical framework, captures the inherent layout variability of a real brain.

In summary, the introduced GLN allows us to systematically compare different network layouts, and determines the elastic energy of a physical network. While the optimization landscape of the layout problem is similar to glassy polymer-like systems, the existence of independent energy wells for each isotopy class allowed us to build a self-consistent statistical mechanical model of the system, which helped us derive the layout characteristics from a single parameter, the layout temperature. Although our algorithm measuring  $\mathcal{G}_n$  for a given network layout can be computationally taxing,

we find that our algorithm generally scales similar to a polynomial with the number of nodes ( $N$ ) (Supplementary Information section SI.6), and at least quadratically with the number of links, which limits our ability to explore very large networks.

Our results also show that normal brains do display a predictable degree of layout variability, which could be experimentally confirmed once data on multiple brains become available. Diseases such as schizophrenia have been associated with abnormalities in the axon bundles<sup>24,25</sup>, suggesting that quantifying layout changes in specific brain regions may offer avenues to identify structural differences between healthy and diseased brain tissues, helping diagnosis<sup>26</sup>, and allowing us to quantify the underlying resilience of a physical network<sup>27</sup>.

Lastly, our results indicate that certain difficult network optimization problems simplify if we map them to dual problems in topology and vice versa. Force-directed network layouts using gradient descent are analogous to perturbed diffusion processes on the

network, helping us explore the relationship between optimization processes on networks and further topological invariants, or characterize dynamical and control processes on networks<sup>28,29</sup>.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41567-020-1029-z>.

Received: 19 November 2019; Accepted: 6 August 2020;

Published online: 19 October 2020

### References

- Cohen, R. & Havlin, S. *Complex Networks: Structure, Robustness and Function* (Cambridge Univ. Press, 2010).
- Caldarelli, G. *Scale-free Networks: Complex Webs in Nature and Technology* (Oxford Univ. Press, 2007).
- Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. Critical phenomena in complex networks. *Rev. Mod. Phys.* **80**, 1275–1335 (2008).
- Song, C., Wang, P. & Makse, H. A. A phase diagram for jammed matter. *Nature* **453**, 629–632 (2008).
- Barthélemy, M. Spatial networks. *Phys. Rep.* **499**, 1–101 (2011).
- Barrat, A., Barthélemy, M. & Vespignani, A. *Dynamical Processes on Complex Networks* (Cambridge Univ. Press, 2008).
- Whitney, H. *Collected Papers of Hassler Whitney* (Nelson Thornes, 1992).
- Armstrong, M. A. *Basic Topology* (Springer Science & Business Media, 2013).
- Gauss, C. F. in *Zur Mathematischen Theorie der Electrodynamische Wirkungen* Collected Works, Vol. 5, 605 (Königlichen Gesellschaft des Wissenschaften, 1833).
- Kauffman, L. *Knots and Physics, Series on Knots and Everything 1* (World Scientific, 1991).
- Bianconi, G. & Capocci, A. Number of loops of size  $h$  in growing scale-free networks. *Phys. Rev. Lett.* **90**, 078701 (2003).
- Viana, M. P. et al. Mitochondrial fission and fusion dynamics generate efficient, robust, and evenly distributed network topologies in budding yeast cells. *Cell Syst.* **10**, 287–297 (2020).
- Gagnon, L. et al. Quantifying the microvascular origin of BOLD-fMRI from first principles with two-photon microscopy and an oxygen-sensitive nanoprobe. *J. Neurosci.* **35**, 3663–3675 (2015).
- Cook, S. J. et al. Whole-animal connectomes of both *Caenorhabditis elegans* sexes. *Nature* **571**, 63–71 (2019).
- Xu, C. S. et al. A connectome of the adult *Drosophila* central brain. Preprint at <https://doi.org/10.1101/2020.01.21.911859> (2020).
- Lein, E. S. et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007).
- Dehmamy, N., Milanlouei, S. & Barabási, A.-L. A structural transition in physical networks. *Nature* **563**, 676–680 (2018).
- Dubrovin, B., Fomenko, A. & Novikov, S. *Modern Geometry Methods and Applications. Part 1: The Geometry and Topology of Manifolds* (trans. Burns, R. G.) (Graduate Texts in Mathematics 104, Springer-Verlag, 1984).
- Mézard, M. & Parisi, G. Replica field theory for random manifolds. *J. Phys. I* **1**, 809–836 (1991).
- Parisi, G. The physical meaning of replica symmetry breaking. Preprint at <https://arxiv.org/abs/cond-mat/0205387> (2002).
- Fraenkel, A. S. Complexity of protein folding. *Bull. Math. Biol.* **55**, 1199–1210 (1993).
- Ngo, J. T., Marks, J. & Karplus, M. in *The Protein Folding Problem and Tertiary Structure Prediction* (eds Merz K. M. & Le Grand S. M.) 433–506 (Springer, 1994).
- Bullmore, E. & Sporns, O. The economy of brain network organization. *Nat. Rev. Neurosci.* **13**, 336–349 (2012).
- Kubicki, M., McCarley, R. W. & Shenton, M. E. Evidence for white matter abnormalities in schizophrenia. *Curr. Opin. Psychiatry* **18**, 121 (2005).
- Zalesky, A. et al. Disrupted axonal fiber connectivity in schizophrenia. *Biol. Psychiatry* **69**, 80–89 (2011).
- Del Ferraro, G. et al. Finding influential nodes for integration in brain networks using optimal percolation theory. *Nat. Commun.* **9**, 2274 (2018).
- Smaldino, P. E., D'Souza, R. M. & Maoz, Z. Resilience by structural entrenchment: dynamics of single-layer and multiplex networks following sudden changes to tie costs. *Netw. Sci.* **6**, 157–175 (2018).
- Cornelius, S. P., Kath, W. L. & Motter, A. E. Realistic control of network dynamics. *Nat. Commun.* **4**, 1942 (2013).
- Hens, C., Harush, U., Haber, S., Cohen, R. & Barzel, B. Spatiotemporal signal propagation in complex networks. *Nat. Phys.* **15**, 403–412 (2019).
- Erdős, P. & Rényi, A. *On Random Graphs I* (Publicationes Mathematicae, 1959).
- Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
- Ahn, Y.-Y., Ahnert, S. E., Bagrow, J. P. & Barabási, A.-L. Flavor network and the principles of food pairing. *Sci. Rep.* **1**, 196 (2011).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2020

### Data availability

Source data are provided with this paper. All other data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

### Code availability

Code is available for this paper at <https://github.com/YanchenLiu1/GLN>. All other code that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request.

### Acknowledgements

We thank J. A. Brum and E. Towilson for useful discussions and for providing the processed Allen Institute mouse brain data, S. Cook for providing the *C. elegans* data, M. Viana for providing the mitochondrial network data and A. Grishchenko for 3D and data visualizations. We were supported by grants from the NSF (grant nos. 1735505 and 1734821), ERC (grant no. 810115 - DYNASNET) and John Templeton Foundation (grant no. 61006). N.D. was also supported by the Office of Naval Research (grant no. 00014-18-9-001).

### Author contributions

Y.L. performed the mathematical modelling, developed the algorithm, ran and analysed the simulations, generated the figures, and contributed to writing the manuscript. N.D. contributed to the mathematical modelling, running the simulations and writing the manuscript. A.-L.B. contributed to the conceptual design of the study and was the lead writer of the manuscript.

### Competing interests

A.-L.B. is the founder of Scipher, Nomix and Foodome that bring network tools to health science.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41567-020-1029-z>.

**Correspondence and requests for materials** should be addressed to A.-L.B.

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