Networks: The Visual Language of Complexity

Blai Vidiella, Salva Duran-Nebreda and Sergi Valverde

Abstract Understanding the origins of complexity is a fundamental challenge with implications for biological and technological systems. Network theory emerges as a powerful tool to model complex systems. Networks are an intuitive framework to represent inter-dependencies among many system components, facilitating the study of both local and global properties. However, it is unclear whether we can define a universal theoretical framework for evolving networks. While basic growth mechanisms, like preferential attachment, recapitulate common properties such as the power-law degree distribution, they fall short in capturing other system-specific properties. Tinkering, on the other hand, has shown to be very successful in generating modular or nested structures 'for-free', highlighting the role of internal, non-adaptive mechanisms in the evolution of complexity. Different network extensions, like hypergraphs, have been recently developed to integrate exogenous factors in evolutionary models, as pairwise interactions are insufficient to capture environmentally-mediated species associations. As we confront global societal and climatic challenges, the study of network and hypergraphs provides valuable insights, emphasizing the importance of scientific exploration in understanding and managing complexity.

Key words: *Networks; Evolution; Hypergraphs; Complex Systems; Tinkering*

Blai Vidiella (iD) : Institute of Evolutionary Biology, CSIC-UPF, Pg. Barceloneta 37, Barcelona 08003, Spain. Theoretical and Experimental Ecology Station, CNRS, Moulis, France. e-mail: `blai.vidiella-rocamora@sete.cnrs.fr`

Salva Duran-Nebreda (iD) : Institute of Evolutionary Biology, CSIC-UPF, Pg. Barceloneta 37, Barcelona 08003, Spain. e-mail: `salva.duran@ibe.upf-csic.es`

Sergi Valverde (iD) : Institute of Evolutionary Biology, CSIC-UPF, Pg. Barceloneta 37, Barcelona 08003, Spain. European Centre for Living Technology (ECLT), Ca' Bottacin, Dorsoduro 3911, 30123 - Venice, Italy. e-mail: `s.valverde@csic.es`

1 Introduction

Understanding complexity is a main scientific challenge [16]. This quest has been a complex process in and of itself, motivated by both theoretical and practical reasons. Consider two examples of using networks to comprehend complexity, one from biology and one from technology. Bertrand Russell proposed in the 1930s that notations play an important role in distilling abstract ideas into clearly intelligible forms, hence aiding communication [124]. Almost 90 years later, the Synthetic Biology Open Language (SBOL) [63] introduced a formal notation in modern biology, aspiring to be an industrial standard for those engineering or simply conveying biological designs (particularly genetic constructs and pathways) (see Figure 1a).

Prior to SBOL, computer scientists were seeking for a language that would allow them to control the complexity of software. Alan Kay, a former math and biology student and Xerox PARC computer scientist, recognized that the level of complexity in software was only comparable to that of biology. Kay played a significant role in the development of Smalltalk [55], a programming language that conceptualized software as a network of code units, drawing inspiration from biological analogies and the interactions between cells: "I thought of objects being like biological cells and/or individual computers on a network, only able to communicate with messages."

Smalltalk is a successful programming language, but neither it nor its contemporary descendants have been the ultimate solution to software development [106]. One of us participated in a video game project in 2002, where we used software engineering techniques, including graphical notation (see Figure 1b). The large-scale structure of our video game proved to be extremely complex and difficult to understand, despite the talent of the software development team and ongoing attempts to optimize, improve, and modularize the software architecture [107, 113] (see Figure 1c). Unwanted complexity is certainly not surprising to many software developers, but neither is there a convincing explanation for it (or a strategy to prevent it). For these reasons, we decided to launch an ongoing research project to investigate the origins of software complexity from a scientific (not engineering) perspective [110]. This example also suggests that SBOL's comparable initiatives may have a limited impact in managing biological complexity (especially when dealing with large non-orthogonal biological designs [22, 23]). But why are notation schemes insufficient? What is missing?

An important requirement is the ability to capture the structure-dynamics relationship as well as the evolutionary component inherent in complex systems. The crucial role of dynamics in complex systems is exemplified by social insect colonies,

whose properties cannot be reduced to the agents building them [123]. Nest architecture emerges from the coordinated actions of many insects, which are, in turn, channeled through the emerging 3-D network of galleries [75, 108]. Feed-backs are a universal feature of the evolution of complexity, which can be a driver of adaptations and innovations emerging at multiple scales [30]. Examples include not only social insects [123, 88], but multicellularity [78, 19, 31, 89], human language [26] and many others [104]. Evolutionary transitions frequently occur at critical points [99, 87] that often involve (or can be understood by) some kind of structural property. Network theory is a tool for not only graphically displaying the structure of a complex system, but also for understanding the logic and evolutionary principles that underpin these systems.
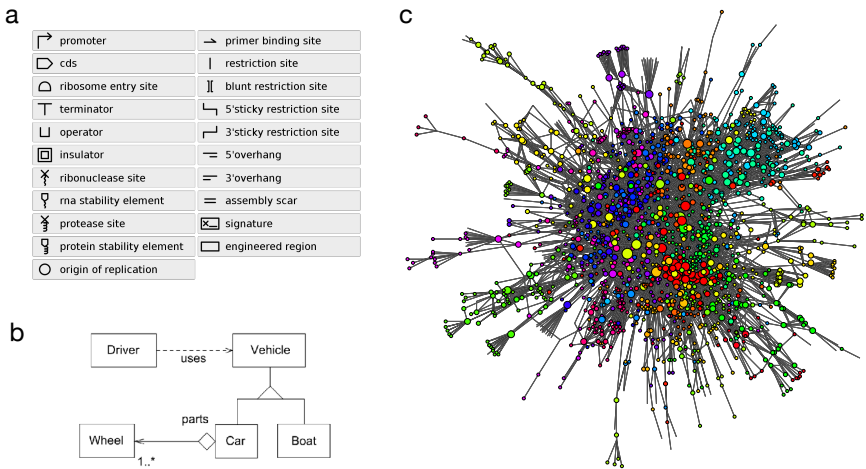

Fig. 1 Visual notations in biology and software. (a) The Synthetic Biology Open Language (SBOL) defines standardised visual notations for representing genetic constructions and biological designs such as DNA sequences, regulatory elements, and functional parts, as well as their connections and interactions within a biological system. Visual symbols represent promoters, coding sequences, terminators, and other genetic components, making it easier to communicate biological ideas across communities. (b) In software engineering, class diagrams are a form of structural diagram implemented in the Unified Modeling Language (UML). They serve the purpose of conceptualizing and representing the distinct attributes of software components, as well as their interrelationships, within a given system or software application. The illustration depicts a section of the class diagram used to create a racing game. (c) The network representation of the class diagram in the video game Prorally 2002 (Ubisoft, 2002). Colour denotes the subsystem the software component belongs to (i.e., physics, rendering, audio, etc.), and node size represents the in-degree.

Networks provide a means to investigate the relationship between system structure and dynamics, offering an intuitive framework to represent intricate interdependencies among system components. Their versatility lies in their ability to capture both local and global properties, enabling the analysis of small-scale interactions and large-scale emergent phenomena. Whether it's understanding the spread of diseases, analyzing information flow in social media, or optimizing transportation networks, graphs offer a robust mathematical language for studying complex systems and their dynamics. The effectiveness of networks as a mathematical representation of complexity has led researchers to embrace them in various scientific disciplines. This multidisciplinary approach facilitates the integration of theoretical and empirical approaches, allowing us to better comprehend and predict abrupt human-induced changes in ecosystems. Many examples demonstrate how networks can deepen our understanding of the impact of human actions on the biosphere. By leveraging the power of networks, we can develop strategies to mitigate the effects of human actions, a crucial challenge in the era of climate change.

2 Measures

A first step in analyzing the structure of complex systems is often to make a picture of it. The human eye is extremely powerful at picking out visual patterns and allow us to put visualization to work on our studies. A network $G = (V, E)$ is a graphical representation that consist of just two fundamental elements: nodes (or vertices) $v_i \in V$ and pairs of nodes $(v_i, v_j) \in E$ so-called links or edges. Despite this apparent simplicity, networks can express complex patterns by connecting nodes and links in numerous ways. This approach may be beneficial even in systems that do not inherently suggest a network [60] (see Figure 2).

Network visualization is, however, only really useful for (a) small networks (up to a few hundred or thousands of vertices), and for (b) sparse networks, that is, when the density of connections is low as not to obscure the main structural features. If there are too many vertices or links, then renderings will be complicated and difficult to understand (such network diagrams are popularly known as 'hairballs'). Many of the networks that interest scientists today have hundreds of thousands or even millions of nodes, which means that visualization is not of much help in their analysis and we need to employ other techniques to determine their structural properties. Here, network theory has developed a large toolkit of measures and metrics that can be useful to understand what data are telling us, even in cases where visualization is not
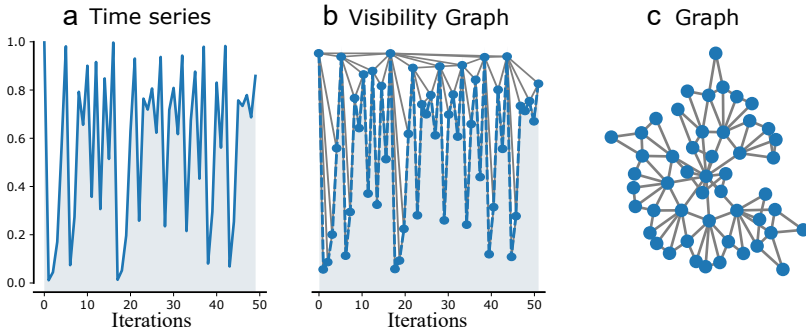
possible.

Fig. 2 Network representation of time series. (a) A Horizontal Visibility Graph (HVG) is a graph representation constructed from time series or one-dimensional data. The HVG is formed by converting time series data into a graph using an algorithm based on the idea of horizontal visibility. (b) Each time series point is treated as a node in the HVG. When a data point at one time step "sees" another data point at a different time step and no other point between them has a higher value, a link is established. (c) The HVG enables the use of graph theory to understand the underlying structure or features of time series data in a variety of domains such as signal processing, economics, and natural sciences.

2.1 Degree

A convenient notation for mathematical purposes is the adjacency matrix. The adjacency matrix A of a simple network $G = (V, E)$ is the matrix with elements $A_{ij} = 1$ iff $(i, j) \in E$ and $A_{ij} = 0$ otherwise. The structure of the adjacency matrix is informative. For an adjacency matrix with no self-edges, the diagonal elements are all zero, while a symmetric matrix shows that the network edges are all bi-directional (so-called 'undirected' edges). In some cases, we would like to represent edges as having a strength, or a weight, or a value attached to them. For example, Internet edges might have weights representing the amount of data flowing along them. In a social network, edges might have weights representing the frequency of interactions. In neural networks, learning processes reinforce specific connections between neurons, and weights might represent the likelihood of signal propagation from one neuron to the next. These weighted networks can be represented by giving the elements of the adjacency matrix values equal to the weights of the corresponding connections,

i.e., the volume of traffic carried by a communication link, the number of social interactions, or the strength of a neuronal connection.

The most basic structural property of any node i is its degree $k(i)$ or the number of links connected to it. The degree can be written in terms of the adjacency matrix as follows:

$$k_i = \sum_{j=1}^{N} A_{ij} \qquad (1)$$

where $N = |V|$ is the number of nodes. Every edge of an undirected network has two endpoints and if there are m edges in total there are $2m$ ends of edges:

$$2m = \sum_{i=1}^{N} k(i) \qquad (2)$$

or

$$m = \frac{1}{2} \sum_{i=1}^{N} k(i) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} A_{ij} \qquad (3)$$

The above result is often used in the study of networks. Notice that the maximum possible number of edges in an undirected network is

$$\binom{N}{2} = \frac{1}{2} N(N-1) \qquad (4)$$

By dividing the number of edges m by this maximum possible, we obtain the connectance (ρ) or density of a network, i.e., the fraction of edges that are actually present [76]:

$$\rho = \frac{m}{\binom{N}{2}} = \frac{2m}{N(N-1)} = \frac{\langle k \rangle}{N-1} \qquad (5)$$

where $\langle k \rangle$ is the average degree of a vertex in an undirected network. A dense network is one in which the density tends to a constant as N approaches infinity. In this situation, as the network grows in size, the proportion of non-zero entries in the adjacency matrix remains constant. The density of a sparse network, on the other hand, approaches zero as the network size increases, and so the average degree tends to remain constant. Almost all real-world systems are thought to be sparse networks, which has profound implications for our understanding of complex systems [90].

2.2 Percolation

A system is a set of nodes that are linked together through a chain of intermediate connectors. Nodes that have no connection to the rest of the connected group of species, genes, or people are almost certainly to be removed from the global structure. This raises the question of how the system components were first joined together.

A simple model provides a broad view of this problem (see Figure 3). Consider a square grid of $L \times L$ water pipes. This spatial network has three types of pipes: (1) pipes that are only open, (2) pipes that are closed, and (3) pipes that are open and may be linked to the left-hand side via a global route. We are interested in the circumstances that allow global routes to appear. Assume that a random pipe on this lattice is opened with probability p, with the other pipes remaining independent. Open pipes look separated from one another with low p values. Although the density of open pipes rises linearly with p (on average, we will have pL^2 open pipes), the global connectedness does not. The fraction of connected pipes is a non-linear function of the probability p. For low p, this fraction stays constant until the critical value of $p = 1/2$ (the so-called 'percolation threshold') is achieved, at which time the water percolates the grid and a global route connects its opposing sides [56].

Percolation transitions like these occur in all types of networks. The simplest model of a random network, the Erdős-Rényi graph [39], provides a null hypothesis for percolation phenomena in disordered systems. The number of nodes N in the random graph is fixed, and each edge is present with probability p and missing with probability $1 - p$. Consider the two extreme scenarios of this model. When $p = 0$, each vertex is an island: the network comprises N isolated components, each of which has precisely one vertex. When $p = 1$, the random network shows all possible edges, resulting in a single linked component. A 'clique' is the formal term for this configuration in which a network of n nodes is linked to every other node.

What happens when probability of connection continuously increases from $p = 0$ to $p = 1$ in a random network? We may hypothesize that the size of the largest connected component $S(p)$ grows linearly with p. But something much more interesting happens. The fraction S of nodes in the largest connected component changes abruptly as the system progresses from disconnected to fully connected. At some critical value of the average degree $z = p(N - 1)$, the random network exhibits a phase transition [87]. A large connected component emerges at the critical point $z = z_c = 1$. When $z > z_c$, most nodes are linked in a single graph; however, when $z < z_c$, the system breaks down into many little subgraphs. That is, the random network predicts the emergence of a percolation threshold separating a disconnected
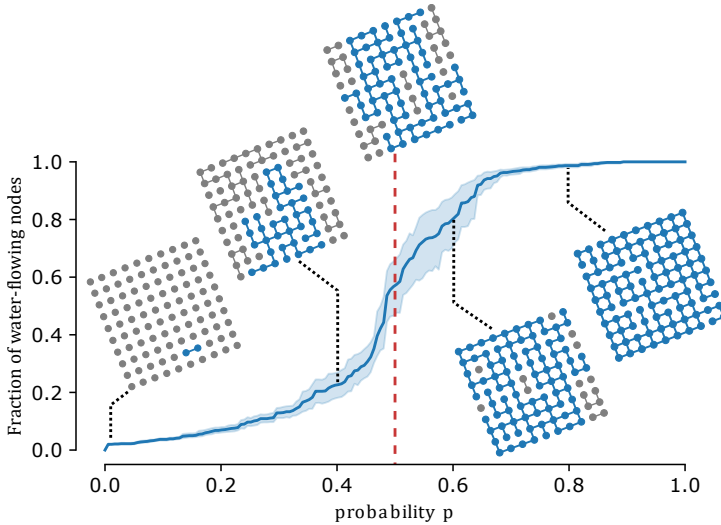
Fig. 3 Using a square regular lattice for water transportation. In a pipe grid, the probability p (horizontal axis) of opening a pipe (forming an edge) and the total fraction of nodes transporting water (vertical axis) are not linearly related. Nodes on opposing sides of the lattice are part of global transportation routes beyond a critical probability of $p > p_c = 1/2$, even when pipes are randomly opened. Blue and grey circles represent water-flowing and disconnected nodes, respectively.

phase from another configuration in which most (if not all) components are connected by at least one global route. This prediction applies to all types of networks, as the local structure at the percolation threshold is expected to be a tree with no loops [21]. However, in real systems, there is also a tendency to short loops [77], which are also very sparse structures.

2.3 Clustering

Random graphs explain the basic features of real-world networks, like the emergence of a large connected component associated with the percolation transition. The random network must be extended to account for other empirical features, particularly when examining local network structures. For example, in social networks [74], the notion of transitivity significantly influences the associations between nodes [49]. The ideal concept of transitivity suggests that connections follow a 'friend of a friend' pattern, often observed in well-connected, fully-linked networks (or cliques). Yet,

most real-world networks, being inherently sparse, operate under partial transitivity. For instance, the fact that 'A' is acquainted with 'B' and 'B' with 'C' does not guarantee that 'A' directly knows 'C' in social networks.

The clustering coefficient may be used to measure the degree of partial transitivity in real systems. The clustering coefficient c_i of a vertex i is the ratio between the number of edges e_i among its nearest neighbors and its maximum possible value [70]:

$$c_i = \frac{e_i}{\binom{k_i}{2}} = \frac{2e_i}{k_i(k_i - 1)} \tag{6}$$

This coefficient ranges from 0 to 1. $c_i = 1$ denotes complete transitivity, i.e., a network composed entirely of cliques. If $c_i = 0$, there are no closed triads. This can happen in a number of network types, including trees (which do not have any loops) and square lattices (which have closed loops with an even number of vertices but no closed triads). There is no transitivity or network clustering in random networks because the chance that any two vertices are neighbors is the same, no matter what the nodes are. As a result, in a random network, the average clustering coefficient is $\langle c \rangle = p = \langle k \rangle / (N - 1)$, which goes to zero in the limit of high N. Empirical clustering values often range from 0.01 to 0.5, indicating considerable differences between random and real networks [118].

Many real-world networks display an approximate degree dependence, with higher degree vertices having a lower clustering coefficient on average [64]. Clustering may also be used to detect so-called "structural holes"[20] in the network, or missing linkages between unrelated nodes linked indirectly via a third vertex. Structured holes are negative for transmission efficiency because they restrict the number of alternate pathways through which information may travel. Structured holes, on the other hand, might be advantageous for a central node whose friends lack connections, granting control over information flow between those friends.

2.4 Motifs

Clustering coefficients are instrumental when looking at the local structure of social networks. When applied to directed networks, where imbalanced information, energy, or matter flow predominates, this coefficient is less effective. An alternative lens to explore network architecture involves network motifs [66]. These motifs denote compact subgraphs comprising M nodes, interconnected by a subset of links within a larger network with $N > M$ nodes. For each value of M, a finite set of variants exists

(see Figure 4a). Counting the occurrences of subgraphs within a network generates a frequency spectrum that provides key insights into their role as basic building blocks. Some of these subgraphs occur at considerably greater frequencies than expected in random graphs. To assess the significance of these occurrences, the Z-score provides a statistical metric that quantifies the prevalence of a particular subgraph Ω_i across the network [66]:

$$z(\Omega_i) = \frac{N_{real}(\Omega_i) - \langle N_{rand}(\Omega_i) \rangle}{\sigma(N_{rand}(\sigma_i))} \qquad (7)$$

where N_{real} is the number of appearances of the subgraph Ω_i in the network, $\langle N_{rand} \rangle$ is the average number of appearances of the subgraph Ω_i in a large number of randomised networks with the same number of nodes and degree distribution as the original network, and σ denotes the standard deviation. A subgraph's Z-score may be positive or negative; $z(\Omega_i) > 0$ indicates that it is greatly overrepresented (it is a motif) in the original network relative to randomized, and $z(\Omega_i) < 0$ indicates that it is severely underrepresented (anti-motif). There are some subgraphs that are overrepresented in the structural and functional networks of the human brain [95, 68].

Although motif analysis is a valuable tool, e.g., when assessing the vulnerability of complex networks [28], its biological relevance in general remains unclear [91, 58]. However, synthetic biologists have made extensive use of the motif framework to understand and exploit design spaces in their biological designs. For instance, an exhaustive analysis of feedforward motifs in regulation has revealed that the most common 'band-pass filter' design in evolution is also the most robust one, both in terms of dynamical properties as well as resilience to mutations in their components [82]. Similarly, understanding biological motifs underpins our capacity to engineer self-organized criticality [117], associative learning [62] and more generally, both temporal [36] and spatial patterns [33] (see Figure 4b,c).

Still, it would be unreasonable to directly infer functionality from the presence or absence of specific motifs, given that a specific subgraph might fulfill a variety of roles in different systems. For example, flexible components that modify their mode of operation to replace damaged or absent components might achieve resilient functionality [101]. According to synthetic models, brain networks are formed to maximize the number and diversity of motifs, as local structural variability enables an extensive spectrum of functional states [95]. Computational research suggests that the degree of synchronization exhibited by a motif is influenced by its structure. Motifs with greater density of links are more likely to synchronize compared with those with fewer connections.
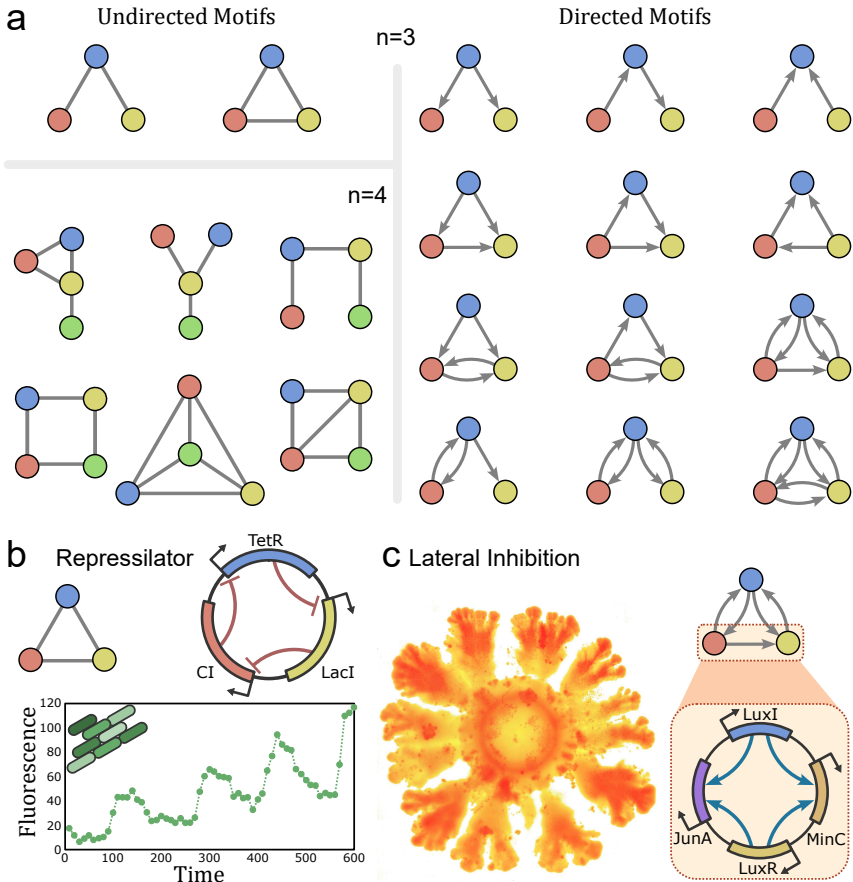
Fig. 4 Motif analysis in biology. (a) List of 3 and 4 node undirected network motifs. Directed motifs (right) display a wider range of patterns than undirected motifs (left), resulting in more complex networks of connections. Synthetic biologists have taken inspiration from these smaller subnetworks to design complex biological circuits. These motifs have proven useful in the implementation of temporal patterns, such as (b) oscillators, e.g., the repressilator circuit [36], which features a cascading triad of inhibiting transcriptional regulators (typically TetR, CI and LacI), yielding an oscillation every 150 minutes. The repressilator can be exogenously controlled and owes its robust behaviour despite its simplicity to the underlying transcriptional motif. Robust spatial pattern formation induced by lateral inhibition (c) also hinges on specific design motifs [33], in the example shown a communication-coordinated feedback loop (LuxR-LuxI) promotes cellular adhesion (JunA) while also arresting bacterial growth (MinC), creating regular patterns of cellular density with a characteristic wavelength of 0.2 cm in synthetic *E. coli* colonies.

3 Evolution

While random graphs offer insights into network structures, they possess a fundamental limitation in their static nature. Random theory depict networks where all nodes are simultaneously created and remain unaltered, overlooking the dynamic expansion processes prevalent in real-world systems. Take cities, for example. They exhibit a dynamic nature akin to living organisms [119, 120]. The complex development of urban networks has been studied from the perspective of fractals [15] and chaos theory [24], since both include self-similar structures as well as sensitivity to initial conditions. This complex city expansion often reflects a combination of distinct developmental patterns: an urban network core surrounded by a grid-like expansion that conforms to the principles of urban planning, while the periphery experiences a more organic and unpredictable growth [114]. A static depiction of networks overlooks the complex interplay among historical events, regulatory influences, and selection factors. Embracing dynamical models is crucial to understanding the evolving dynamics of cities [10] and other complex networks.

An open question is to what extent we can define a universal theoretical framework for evolving networks. While basic growth mechanisms, like preferential attachment, contribute to explaining common properties such as the power-law degree distribution, they fall short in capturing the distinctive nature of individual networks. Understanding the evolution of real-world networks necessitates a combination of structural traits beyond local features.

3.1 Growth

Many interesting examples of growing networks are related to evolutionary processes, whether in natural or artificial systems. For example, although technical breakthroughs are frequently the result of purposeful efforts, George Basalla observed that a broad range of inventions cannot be explained solely by intentional processes, implying deep analogies between biology and technology that have been the focus of extensive research [11, 72, 97]. Furthermore, networks may be used to explore the parallels (as well as the differences) across various evolutionary systems. Historical events both have a significant impact on the course of technology and biology, often resulting in unexpected detours on the path to innovation [5]. Biology and technology also depend on differential replication, in which accidental, seemingly random changes are inherited by descendants. Similar to selection forces in evolution, artifacts are also susceptible to external success measures [65], such as reliability,

performance, novelty [32, 69], and even aesthetic preferences.

Networks are a promising approach to study the evolution of technology, which lies at the intersection between statistical physics, computer science, and evolutionary biology [93]. The network of patent citations, for example, is a scale-free network, a property also observed in other physical, natural, and man-made systems in which the degree distribution $P(k)$ is a power-law rather than a normal distribution:

$$P(k) = Uk^{-\gamma} \tag{8}$$

where U is a constant, and $2 \leq \gamma \leq 3$ is the exponent of the power-law [8]. Power law distributions are different from normal distributions in that they do not have a peak at the average degree $\langle k \rangle$, and they are more likely to contain extreme degree values.

Assuming that the amount of citations is a reliable indicator of a technology's importance, the power-law distribution of citations may be interpreted as an indication of inequality between innovations. But what factors contribute to the popularity of a node? De Solla Price suggested the 'cumulative benefit' mechanism for citation networks in 1965, stating that 'there is a likelihood that the more an article is cited, the more probable it is to be mentioned later' [94]. This process is similar to the 'rich get richer phenomenon' proposed by Yule [125] and Simon [85], as well as the newer concept of 'preferential attachment' [9].

We can incorporate the mechanism of 'cumulative advantage' into a model of growing networks, where the rate of incoming links is proportional to the degree of any node:

$$\Pi(k) \sim k^{\beta} \tag{9}$$

where the attachment exponent β regulates the strength of link reinforcement. It is possible to demonstrate that the above equation produces a degree distribution that follows a power-law [125, 9]. Although this model is often simplistic for reproducing real-world degree distributions, it can be easily expanded. For example, this equation does not take into account existing temporal relationships between inventions. To understand the evolution of technology, we must consider that individual traits are not the unique criteria of success; the environment must also be adequate, since their role may change over time in evolution (e.g., an exaptation [48]). Taking into account this, we can define a more realistic model for the growth of patent citations that combines preferential attachment with temporal dependence [115]:

$$\Pi(k, \tau) \sim k^{\beta} \tau^{\alpha-1} \exp(\tau/\tau_0)^{\alpha} \tag{10}$$

where τ is the node age, the exponent $\alpha > 0$ weights how quickly the ageing influences the chance of attachment, and τ_0 is a scale parameter that determines the rightward extension of the ageing curve. We can verify these assumptions in a database of over 3 million patents [115], which suggests a complex interplay between structure and environment in the evolution of technology. However, a power-law degree distribution is a general feature that many systems have, hence it cannot identify what distinguishes a network from others.

3.2 Modularity

A modular system is formed by quasi-independent parts that appear integrated within themselves, but also exhibit a certain degree of interdependency among them. Modularity is considered a prerequisite for the adaptation of complex organisms and their evolvability [83]. This is particularly evident in cellular networks, where modularity may be identified at the topological level [79, 92]. In many situations, modular architectures seem to be linked to functional traits: for example, a group of closely similar proteins may all be involved in cell division or communication [53]. While empirical research, e.g. [38], suggests a potential relationship between modularity and functional traits, no universal laws or consistent patterns have emerged in relation to modular structure. A main obstacle is the need to get reliable functional data, which involves the detailed and costly inspection of a large number of nodes.

An active area of research combines statistical physics with computational approaches to develop community detection algorithms that solve this problem [70]. The approach takes into account a decomposition of the graph $G = (V, E)$ into a collection of s subgraphs $C = \{C_i \subset V : 1 \leq i \leq s\}$ that defines a partition. Because there are so many possible partitions, it is critical to have a reliable index for assessing the degree of connectivity between alike nodes in the network. The optimal partition is determined by maximizing the modularity index $Q = Q_{\max} = \max(Q)$ [71]. The modularity score is:

$$Q = \sum_{i=1}^{s} \frac{e_i}{m} - \left(\frac{d_i}{2m}\right)^2 \qquad (11)$$

where e_i is the number of edges in module i, d_i is the total degree of nodes in module i and m is the number of edges in the full network. This method is widely used, but it comes with several limitations (see Figure 5). The so-called 'resolution limit' [42] is a significant constraint that forces the algorithm to integrate small modules whenever the merging gives a positive gain. Let's define the gain ΔQ as follows:

$$\Delta Q_{ij} = \frac{e_{ij}}{m} - 2\left(\frac{d_i}{2m}\right)\left(\frac{d_j}{2m}\right) \tag{12}$$

where e_{ij} is the number of edges between modules i and j. A positive gain $\Delta_Q > 0$ occurs when the number of edges between two subsets of nodes is greater than the expected number in a random graph [46]:

$$e_{ij} > \frac{d_i d_j}{2m} \tag{13}$$

In sparse or large networks, this value can fall below one ($e_{ij} < 1$), prompting the merging of weakly linked modules, contrary to our intuition that they should remain separate [46]. The resolution limit results from the modularity equation's null model, which assumes equal probabilities of node connections. This assumption may not hold valid in large or geographically embedded networks, as each node may only have a finite "horizon" within which it may communicate with other nodes [111]. For example, each connection in an intercommunication network has a cost that lowers the likelihood of connecting with another one separated by a large distance [25].

Modularity maximization presents an inherent challenge: maximizing the first term of the modularity equation often requires including numerous edges within modules, while reducing the second term necessitates dividing the graph into smaller, low-degree modules. Finding the optimal balance is a hard computational problem. And as a result, all practical modularity algorithms are based on heuristics, which yield suboptimal partitions [18]. Furthermore, it is unclear how to tune the parameters of the maximization algorithm to obtain the optimal solution. One method involves sweeping the parameter space by varying the temperature in a simulated annealing search [84]. The network gets clustered into an increasing number of groups and modularity values when the temperature is lowered. The shift from low to high modularity is not smooth, but consists of multiple phases separated by abrupt discontinuities (see Figure 5a). The presence of multiple potential solutions, each with slightly different modularity scores for a given network (Figure 5b), reveals a landscape of local maxima [46] (see Figure 5c). To address this problem of degeneration, consensus-based solutions that identify resilient communities are necessary.

Evolution cannot invariably produce maximally modular networks when functional and cost constraints are present, which is a more significant limitation than the constraints associated with a mathematical definition of modularity. The so-called "breakdown of modularity" involves a transition from modular to well-connected networks [105, 43], and it could be related with neurological disorders including Alzheimer's disease or schizophrenia [27, 2]. An extensive search of the space of

3-input logical functions reveals that not all minimal feed-forward Boolean networks
(FFBNs) are modular [105] (see Figure 5d). Looking at this phenotypic network, we
can observe that evolution of circuits such as the multiplexer and the majority func-
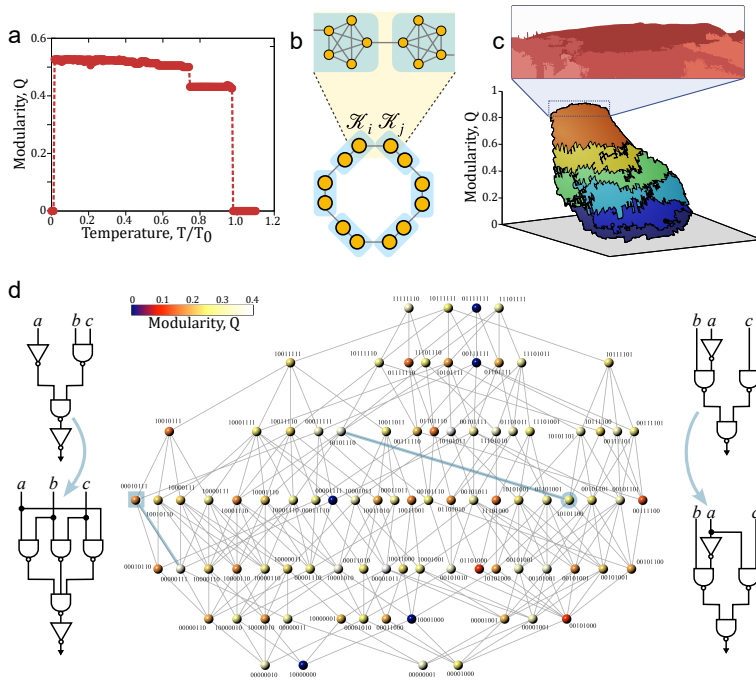

Fig. 5 Limits to modularity. (a) A simulated annealing search across a broad temperature range
finds abrupt modularity transitions separating homogeneous phases. The presence of stable phases
suggests how a partial (heuristic) search is likely to provide sub-optimal modularity values. (b)
Pairs $(\mathcal{K}_i, \mathcal{K}_j)$ of 5-node cliques are connected by a single link (top) to create a 24-clique ring
network (bottom). The best modularity partition, which combines two neighboring cliques, has
slightly larger modularity $Q_2 = 0.8712$ than the intuitive partition, which arranges individual
cliques on their own, with modularity $Q_1 = 0.8674$. (c) The modularity landscape for this ring
network displays a high-modularity plateau of degenerated solutions (inset). (d) A phenotype network
(center) for each possible three-input Boolean function, with each node representing one logic
function. The functions that are most similar are linked together. The color of a node represents the
modularity of the most cost-effective (wire-wise) feed-forward Boolean circuit for implementing
that function. A light blue square (left) and a circle (right) highlight two important functions:
"majority" and "multiplexor." Evolving a circuit while simultaneously minimizing the number of
wires and implementing these target functions inevitably reduces the modularity of its predecessors.
The "breakdown of modularity" refers to the transition from modular ancestors to less modular
target circuits (both sides of the panel), as depicted in the phenotypic network (light blue arrows).

tion, which are important in both electrical circuits and synthetic biology, is linked to a modularity reduction from an ancestral Boolean function. In these circumstances, the only way to retain modularity is to raise the cost of the circuit, for example, by increasing structural redundancy.

3.3 Tinkering

Complex systems emerge from the nonlinear interactions among numerous components, leading to unpredictable behaviors. Modular structures are a frequent component of many complex systems, both in nature and in engineering. In technology, modularity provides significant benefits as it establishes distinct task divisions, aiming to reduce costs and enhance overall reliability [7]. Herbert Simon postulated that a modular system can adapt more rapidly than one that is not modular [86]. But whether modularity is a reflection of general evolutionary principles or is exclusive to certain functional features is still an open question. The prevalence of modular designs in both natural and human-made contexts suggests the existence of universal processes, while current heuristic methods for identifying modularity paint a different picture.

According to [54], networks evolved under "modularly varying goals" (MVG) should also be modular. A network evolves under MVG when it adapts to changing objectives over time, with each goal having the same subproblems as the preceding one. An example is a logic circuit with multiple parts (modules), each implementing a sub-function needed to perform the overall objective. When objectives change, mutations that rewire these modules rapidly become permanent in the population to meet the new objective. Computational investigations have shown that MVG may greatly speed evolution (resulting in a modular network) as compared to evolution with fixed or fluctuating objectives (leading to a non-modular network). However, it is unclear how many biological settings vary in this modular fashion and if they change often enough to produce modular systems.

Modularity has also been suggested to evolve, not because it promotes evolvability in fluctuating environments, but as a by-product of selection that limits the number of links [25]. A more intriguing possibility is that mutation mechanisms, like gene duplication, could create a bias towards modular structures [92]. In the field of technology, makers employ previously accessible innovations as an integral aspect of their work, such as by replicating designs and reusing components acquired from others. Technology, like evolution, evolves mostly by tinkering [52], that is, by combining prior components in novel ways [6]. Even in the absence of selection, a basic network

model based on duplication and rewiring can predict not only the empirical degree distribution but also the frequency of motifs in the real systems [112].
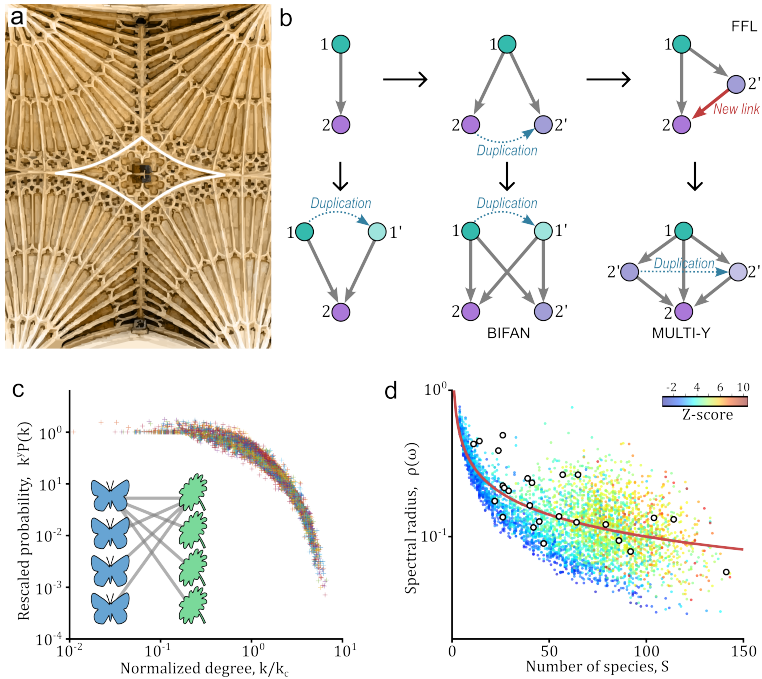

Fig. 6 Networks as evolutionary spandrels. Some network properties might be "spandrels," or byproducts or unintended consequences of tinkering rules, rather than the outcome of selection forces. (a) The ceiling of this cathedral shows decorated structures in the middle of four branches (surrounded by a white line). These decorations may be understood as functional essential pieces, although the space (an architectural spandrel) is just an unavoidable result of two arches. (b) We can generate several motifs by copying nodes (blue arrow) or adding new links (red arrow) starting from the simplest motif (top-left). This set of tinkering rules can generate three common motifs (FFL, BIFAN, and MULTI-Y) in directed networks, the (c) truncated power-law behaviour of mutualistic networks (inset) and the (d) relationship between the spectral radius ρ (a robust measure of nestedness, see [96]) and the number of species (S) (color nodes). Open circles correspond to a total of 25 real mutualistic networks [109]. The red line represents the random network prediction for nestedness, which is based on Wigner's semi-circle rule [4] and describes the average behavior but not the dispersion of real systems.

Agreement between data and generative models lead us to propose that motif distribution is equivalent to the concept of evolutionary spandrel discussed by Gould

and Lewontin [91]. A popular biological argument is that structural features (such as modularity) have been selected for the advantages they provide. The logic goes as follows: certain features improve stability, improving the odds of survival; hence, these traits must be evolutionary adaptations. Some features, however, are little more than a 'spandrel,' as Gould and Lewontin call them [47, 59]. For example, we may observe ornamented structures in the midst of four branches on the ceiling of a cathedral (see Figure 6a). Although these decorations look like important functional elements, they really reside in triangular spaces known in architecture as spandrels, which are an inevitable outcome of having two arches. Similarly, even in the absence of selection, the overabundance of certain subgraphs in networks may be an unavoidable result of tinkering processes (see Figure 6b).

3.4 Nestedness

A widespread and intriguing pattern in ecological networks is nestedness, which is characterized by the tendency of low-degree species to interact with a subset of highly connected species [13]. The canonical interpretation of nestedness in ecological networks is linked to the likelihood observing a particular pattern of interactions compared to a null model [121]. This analysis is traditionally carried out on topological matrices (meaning that they only account for the presence or absence of links). Such matrices are typically nested, but it has been pointed out [96], that a better characterization of these webs using link weights reveals that only a small fraction of them exhibit nestedness [100], suggesting that a truly meaningful assessment of nested patterns requires a weighted interaction matrix.

Edges in a nested network are organized in such a way that specialists interact with a subset of the species whom generalists interact with. This definition has been recently extended to quantitative networks using spectral graph theory [96, 109] (see Figure 6d). Formally, nestedness is a systematic arrangement of non-zero entries in the adjacency matrix. This pattern can be clearly identified by computing the so-called overlap and declining fill (NODF[3]). For the bipartite graph $G = (P, M, E)$ with biadjacency matrix $B = [B_{ij}]$:

$$NODF(G) = \frac{1}{K}\left[\sum_{i,j=1}^{N_P}\left(\theta(q_i - q_j)\frac{\sum_{k=1}^{N_M} B_{ik}B_{jk}}{q_j}\right) + \sum_{k,l=1}^{N_M}\left(\theta(r_k - r_l)\frac{\sum_{i=1}^{N_P} B_{ik}B_{il}}{r_l}\right)\right]$$

where $K = [N_P(N_P - 1) + N_M(N_M - 1)]/200$ is a normalisation constant to ensure that $0 \leq NODF \leq 100$, $N_P = |P|$, $N_M = |M|$, θ is the Heaviside function with $\theta(0) = 0$, and q_i and r_j are the degrees of nodes $i \in P$ and $j \in M$, respectively. A high NODF value implies that certain species' interactions are a subset of other generalist species' interactions (and so the network demonstrates nestedness), whereas a low value indicates clustering (thus implying modular structures).

Following an adaptationist view of naturally evolved systems, it has been argued that the presence of nested patterns is a consequence of underlying selection processes that reduce competition relative to the benefits of facilitation, increasing biodiversity [12] and food web persistence [61]. Supporting evidence for this view are analytical approaches to generalized Lotka–Volterra equations with different functional responses. However, recent papers have challenged this view [100]. Instead, it has been argued that nestedness is likely to be a consequence (rather than a causative property) of biodiversity, in particular of the heterogeneous distributions of connections [29, 96] (see Figure 6c). Furthermore, null models [109] have demonstrated that nestedness can be easily generated from generative processes; this suggests that tinkering models, which rely on basic copy and rewire processes, generate nested structures, thereby implying that nestedness is an evolutionary spandrel [47, 109]. Thus, validation of empirical nested patterns should take into account the predictions of tinkering models as a baseline to assess whether the observed patterns are meaningful, i.e. deviating from null expectations (see Figure 6d).

4 Beyond Networks

Networks offer an efficient means of capturing both local and global structural features, serving as a valuable framework for investigating the dynamic properties of systems. However, their scope is limited to such pairwise interactions. In contrast, hypergraphs provide a solution to this limitation by allowing edges to connect any number of vertices (see Figure 7b), rendering them exponentially more powerful and expressive in describing real-world interactions. Following previous notation, a hypergraph $H = (V, E)$ is a graph consisting of nodes (or vertices) $v_i \in V$ and and hyperlinks (or hyperedges) which can span any number of nodes $(v_i...v_j) \in E$.

For instance, hypergraphs are indispensable when pathogens exhibit varied infection behaviors based on the presence or absence of multiple disease vectors [80] (see Figure 7d), or in cases of symbiotic relationships involving more than two species [67, 34]. In other ecological systems, conventional depictions of predatory

interactions, such as those in food webs, typically involve only two species. However, research has repeatedly demonstrated that many ecological interactions are of higher order [1]. This means that they can involve more than two species [44], incorporate external ecological factors that influence pairwise interactions [61, 116] (see Figure 7e), or are driven by density-dependent behavior (such as when prey change their behaviour leading to their predators switching to other more suitable preys). In these cases, a network formalism may fall short in capturing the essential structural properties that drive emergent dynamical features. Hypergraphs become a necessary tool for making accurate predictions in systems with higher order interactions, due to their enhanced expressiveness and flexibility, rendering them highly valuable in complex systems research.

Given that the structure of ecological networks has been linked to patterns of extinction and robustness [17], how does a hypergraph perspective address these views? It has long been proposed that a modular network structure prevent the spread of perturbations (such as secondary extinctions) in ecological networks, with modules acting as firewalls against the propagation of extinction cascades [98, 73]. This suggests that modular patterns should be more resilient and, consequently, preserved in evolutionary processes. Conversely, mutualistic networks tend to exhibit a nested rather than modular pattern, which has also been associated with increased resilience and productivity in various ecological systems [12]. This assessment goes beyond mere conjecture; Sanders and colleagues demonstrated through field experiments with plant-insect communities that more complex food webs can indeed buffer against the effects of species loss [81].

Traditionally, modularity and nestedness were considered mutually exclusive [96, 41] and imbued ecological networks with different dynamical properties [12]. Various network models have emerged, attempting to account for their independent origins. However, recent research in hypergraphs has revealed that when modularity and nestedness inhabit distinct dimensions of high-order interaction graphs, or hypergraphs, they can coexist [116]. In this way, a projection of a tripartite hypergraph into a network (by collapsing the adjacency matrix from 3 to 2 dimensions) can recover either structural pattern (see Figure 7a).

Fig. 7 Hypergraphs examples and bipartite projections. A tripartite hypergraph of host-parasites-habitat can be decomposed or projected into host-parasite or host-plant bipartite networks (a). These projections can display distinct structural patterns such as the host-plant network being modular (left) and the host-parasite being nested (right). ...
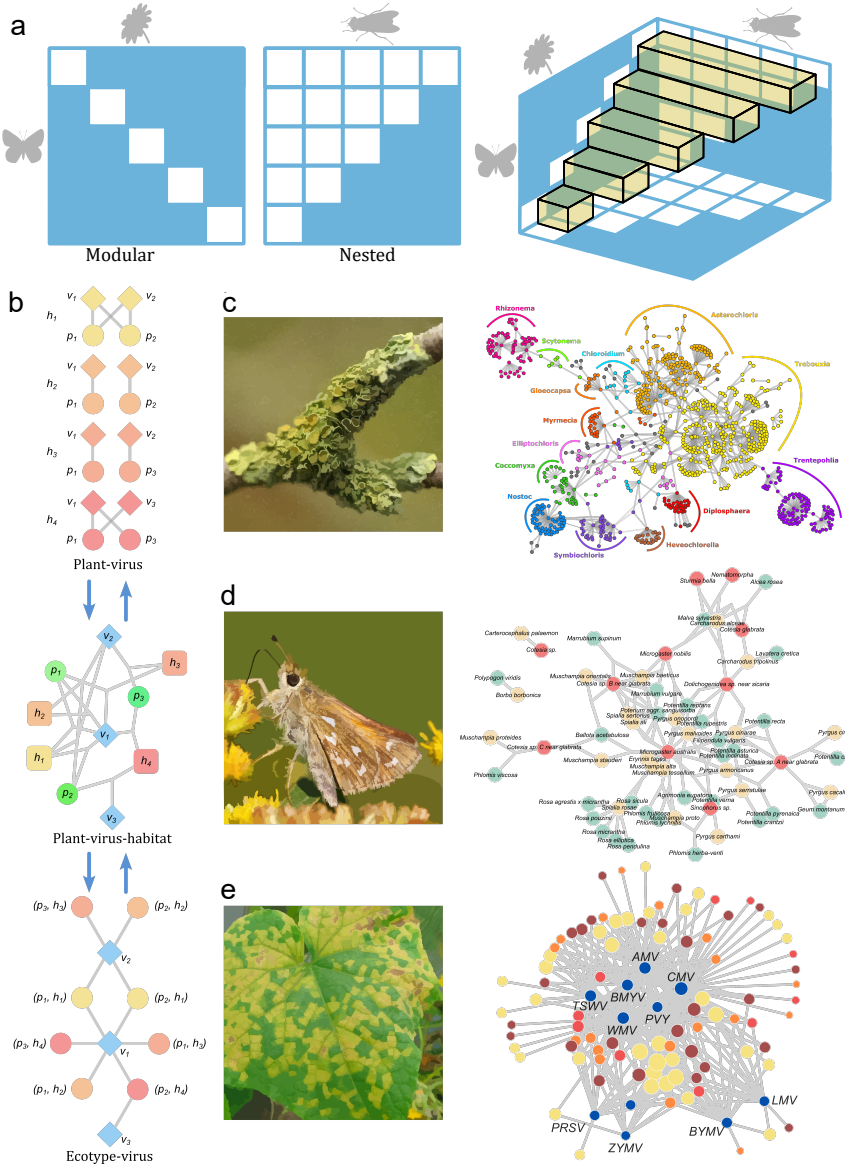
Fig. 7 ... Hypergraphs can also be reconstructed from 'partial' information, such as separate bipartite networks (b). Here, two bipartite networks are integrated into a hypergraph containing plants, viruses and biomes [116]. Many systems are amenable to a hypergraph representation: from lichen symbionts that contain many fungi-algal-bacterial relations [34] (c), plant-butterfly-parasite interactions [102] (d) and plant ecotype-parasite infection [116] (e).

Hypergraphs can also help researchers integrate different types of interactions in a single graph object. For instance, to better understand the compounding effect of human laws in biodiversity and biological interactions [122, 37]; the complex relationships between landscape, species and ecosystem services [40]; how anthropogenic climate change drives the reorganization of ecological networks [45]. With hypergraphs it also becomes possible to incorporate spatial or temporal dimensions to ecological networks [116], allowing us to address when structural patterns are driven by seasonal variation, or when interactions are guided by 'ecotypes' (see Figure 7e, i.e. when species change their morphological features depending the biome they inhabit [51], thus altering the suite of interactions they partake in). However, although many of the structural metrics mentioned in this chapter also exist for hypergraphs [14], the field of empirical research in hypergraphs is still pretty much in its infancy, and further examples and mechanistic models will be necessary to bring them up to par to the ubiquity of network analysis.

5 Discussion

Biological complexity results from adaptations and shifts to novel features, taking place at various scales. Not all network characteristics, however, are the outcome of direct selection processes; rather, they can unexpectedly emerge from internal dynamics, akin to evolutionary spandrels. Scale-free degree distributions, motif frequencies, and even nestedness may have emerged as accidental outcomes of non-adaptive processes, such as tinkering rules. Simple growth models, however, are insufficient to explain how network complexity has evolved over time. Exogenous and endogenous mechanisms both contribute to the emergence of complexity.

Not all innovations lead to more complexity; the choice of which ones persist is ultimately based on environmental conditions and cost-benefit analyses. In technology, software engineers try to manage complexity by rationally subdividing a system into modules that interact in a neat and clear way [7]. External sources of complexity, like performance constraints, can entangle the original design, which soon becomes much less modular [105]. Software developers actively oppose complexity in this situation because unwanted complexity makes the system much more challenging to understand and maintain, often leading to entangled and monolithic systems [35]. Another source of complexity is the design principle that maximizes reuse of existing components. The merging of an existing system with an external one, e.g., a third-party library, re-

sults in complex co-evolutionary dynamics between different software projects [103].
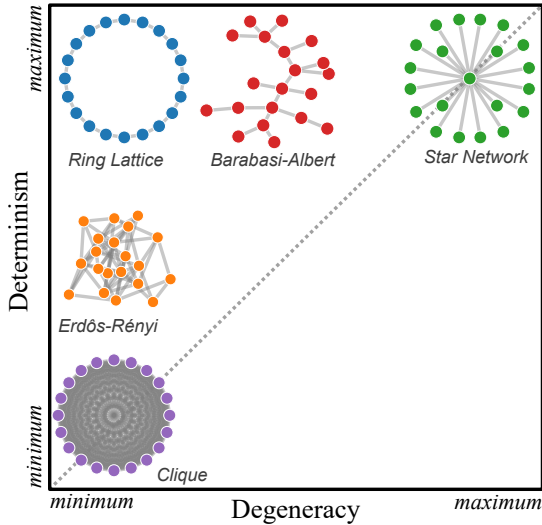

Fig. 8 Network morphospaces. Degeneracy $\langle H(W_i^{out}) \rangle$ and determinism $H(\langle W_i^{out} \rangle)$ display a fundamental trade-off which confines networks to a specific domain of this design space [57]. Here, $H(X) = -\sum_n P(x_i) \log_2 P(x_i)$ represents the Shannon's entropy of a given distribution X, and the distribution $W_i^{out} = \{w_{ij}\}/\sum_j w_{ij}$ is the normalized set of weights from node i to every other node j. These weights can be understood as the distribution of probabilities a random walker might take when visiting node i. Regular structures (star, clique), lattices (ring) and typical networks produced by standard generative models (Erdős-Rényi and Barabási-Albert) are shown in this space. A forbidden region in the bottom right corner of the morphospace is uninhabited by any real network.

We can use network morphospaces to characterise the vast landscape of network designs and how external forces and internal constraints drive structural diversity. These are representations of design spaces, where different networks are located according to some well defined metrics. Typically, the morphospace dimensions are chosen because they reflect some deeper insight into the evolutionary process or constrains that limit what is possible in these systems, i.e. they often display trade-offs between them. One such example is the network degeneracy-determinism morphospace (see Figure 8). This space quantifies the information contained in the topology of a network and can be used to signal the presence of higher informative scales. Here, determinism corresponds to the degree of determinism or certainty associated to a random walker traversing the network in every node, and the second

dimension captures the system's degeneracy or entropy in weight distributions [57]. The features of this morphospace have implications for the dynamical properties of brain activity [50].

In this chapter we have explored a wealth of tools, metrics and morphospace representations that allow us to map and navigate complex systems. Understanding complexity not only gives us insights into the generative processes that underlie biological and technological systems, but also informs the software and biological engineers that want to manage the complexity of their designs. Stephen Hawking's prediction of a 'century of complexity' seems to be spot on with our current reality, which is distinguished by extraordinary technological advances contrasted against profound social and planetary challenges. In this age of complexity, our analysis of network architectures barely scratches the surface. Underlying there is a vast, undiscovered terrain, just like the bigger complexities awaiting discovery and comprehension in the world around us.

References

[1] Abrams, P.A., 1983. Arguments in favor of higher order interactions. The American Naturalist 121, 887–891.

[2] Alexander-Bloch, A.F., Gogtay, N., Meunier, D., Birn, R., Clasen, L., Lalonde, F., Lenroot, R., Giedd, J., Bullmore, E.T., 2010. Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia.

Frontiers in systems neuroscience 4, 147.

[3] Almeida-Neto, M., Ulrich, W., 2011. A straightforward computational approach for measuring nestedness using quantitative matrices. Environmental Modelling & Software 26, 173–178.

[4] Arnold, L., 1971. On wigner's semicircle law for the eigenvalues of random matrices. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 19, 191–198.

[5] Arthur, W.B., 1994. Increasing returns and path dependence in the economy. University of michigan Press.

[6] Arthur, W.B., 2009. The nature of technology: What it is and how it evolves. Simon and Schuster.

[7] Baldwin, C.Y., Clark, K.B., 2000. Design rules: The power of modularity. volume 1. MIT press.

[8] Barabasi, A.L., 2016. Network science: The scale-free property, in: Network science: The scale-free property. Cambridge University Press, pp. 1–15.

[9] Barabási, A.L., Albert, R., 1999. Emergence of scaling in random networks. science 286, 509–512.

[10] Barthelemy, M., 2018. Morphogenesis of spatial networks. Springer.

[11] Basalla, G., 1988. The evolution of technology. Cambridge University Press.

[12] Bascompte, J., Jordano, P., 2007. Plant-animal mutualistic networks: the architecture of biodiversity. Annual Review of Ecology, Evolution, and Systematics 38, 567–593.

[13] Bascompte, J., Jordano, P., Melián, C.J., Olesen, J.M., 2003. The nested assembly of plant–animal mutualistic networks. Proceedings of the National Academy of Sciences 100, 9383–9387.

[14] Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., Young, J.G., Petri, G., 2020. Networks beyond pairwise interactions: Structure and dynamics. Physics Reports 874, 1–92.

[15] Batty, M., Longley, P.A., 1994. Fractal cities: a geometry of form and function. Academic press.

[16] Bonner, J.T., 1988. The evolution of complexity by means of natural selection. Princeton University Press.

[17] Bornholdt, S., Sneppen, K., 2000. Robustness as an evolutionary principle. Proceedings of the Royal Society of London. Series B: Biological Sciences 267, 2281–2286.

[18] Brandes, U., Delling, D., Gaertler, M., Görke, R., Hoefer, M., Nikoloski, Z., Wagner, D., 2006. Maximizing modularity is hard. arXiv preprint physics/0608255 .

[19] Brunet, T., King, N., 2017. The origin of animal multicellularity and cell differentiation. Developmental cell 43, 124–140.

[20] Burt, R.S., 2004. Structural holes and good ideas. American journal of sociology 110, 349–399.

[21] Callaway, D., Newman, M., Strogatz, S., Watts, D., 2000. Exact solution of percolation on random graphs with arbitrary degree distributions. Phys. Rev. Lett. 85, 5468–5471.

[22] Carbonell-Ballestero, M., Duran-Nebreda, S., Montanez, R., Solé, R., Macía, J., Rodríguez-Caso, C., 2014. A bottom-up characterization of transfer functions for synthetic biology designs: lessons from enzymology. Nucleic acids research 42, 14060–14069.

[23] Carbonell-Ballestero, M., Garcia-Ramallo, E., Montañez, R., Rodriguez-Caso, C., Macía, J., 2016. Dealing with the genetic load in bacterial synthetic biology circuits: convergences with the ohm's law. Nucleic acids research 44, 496–507.

[24] Cartwright, T.J., 1991. Planning and chaos theory. Journal of the American Planning Association 57, 44–56.

[25] Clune, J., Mouret, J.B., Lipson, H., 2013. The evolutionary origins of modularity. Proceedings of the Royal Society b: Biological sciences 280, 20122863.

[26] Corominas-Murtra, B., Valverde, S., Solé, R., 2009. The ontogeny of scale-free syntax networks: phase transitions in early language acquisition. Advances in Complex Systems 12, 371–392.

[27] David, A.S., 1994. Dysmodularity: a neurocognitive model for schizophrenia. Schizophrenia bulletin 20, 249–255.

[28] Dey, A.K., Gel, Y.R., Poor, H.V., 2019. What network motifs tell us about resilience and reliability of complex networks. Proceedings of the National Academy of Sciences 116, 19368–19373.

[29] Dunne, J.A., Williams, R.J., Martinez, N.D., 2002. Network structure and biodiversity loss in food webs: robustness increases with connectance. Ecology letters 5, 558–567.

[30] Duran-Nebreda, S., Bassel, G.W., 2017. Bridging scales in plant biology using network science. Trends in plant science 22, 1001–1003.

[31] Duran-Nebreda, S., Bonforti, A., Montanez, R., Valverde, S., Solé, R., 2016. Emergence of proto-organisms from bistable stochastic differentiation and adhesion. Journal of The Royal Society Interface 13, 20160108.

[32] Duran-Nebreda, S., O'Brien, M.J., Bentley, R.A., Valverde, S., 2022. Dilution of expertise in the rise and fall of collective innovation. Humanities and Social Sciences Communications 9, 1–10.

[33] Duran-Nebreda, S., Pla, J., Vidiella, B., Pinero, J., Conde-Pueyo, N., Solé, R., 2021. Synthetic lateral inhibition in periodic pattern forming microbial colonies. ACS Synthetic Biology 10, 277–285.

[34] Duran-Nebreda, S., Valverde, S., 2023. Composition, structure and robustness of lichen guilds. Scientific Reports 13, 3295.

[35] Eick, S.G., Graves, T.L., Karr, A.F., Marron, J.S., Mockus, A., 2001. Does code decay? assessing the evidence from change management data. IEEE transactions on software engineering 27, 1–12.

[36] Elowitz, M.B., Leibler, S., 2000. A synthetic oscillatory network of transcriptional regulators. Nature 403, 335–338.

[37] Epstein, Y., López-Bao, J.V., Chapron, G., 2016. A legal-ecological understanding of favorable conservation status for species in europe. Conservation Letters 9, 81–88.

[38] Érdi, P., Makovi, K., Somogyvári, Z., Strandburg, K., Tobochnik, J., Volf, P., Zalányi, L., 2013. Prediction of emerging technologies based on analysis of the us patent citation network. Scientometrics 95, 225–242.

[39] Erdős, P., Rényi, A., 1960. On the evolution of random graphs. Publ. math. inst. hung. acad. sci 5, 17–60.

[40] Felipe-Lucia, M.R., Soliveres, S., Penone, C., Fischer, M., Ammer, C., Boch, S., Boeddinghaus, R.S., Bonkowski, M., Buscot, F., Fiore-Donno, A.M., et al., 2020. Land-use intensity alters networks between biodiversity, ecosystem functions, and services. Proceedings of the National Academy of Sciences 117, 28140–28149.

[41] Fortuna, M.A., Stouffer, D.B., Olesen, J.M., Jordano, P., Mouillot, D., Krasnov, B.R., Poulin, R., Bascompte, J., 2010. Nestedness versus modularity in ecological networks: two sides of the same coin? Journal of Animal Ecology , 811–817.

[42] Fortunato, S., Barthelemy, M., 2007. Resolution limit in community detection. Proceedings of the national academy of sciences 104, 36–41.

[43] Godwin, D., Barry, R.L., Marois, R., 2015. Breakdown of the brain's functional network modularity with awareness. Proceedings of the National Academy of Sciences 112, 3799–3804.

[44] Golubski, A.J., Westlund, E.E., Vandermeer, J., Pascual, M., 2016. Ecological networks over the edge: hypergraph trait-mediated indirect interaction (tmii) structure. Trends in ecology & evolution 31, 344–354.

[45] Gómez-Ruiz, E.P., Lacher Jr, T.E., 2019. Climate change, range shifts, and the disruption of a pollinator-plant complex. Scientific reports 9, 14048.

[46] Good, B.H., De Montjoye, Y.A., Clauset, A., 2010. Performance of modularity maximization in practical contexts. Physical review E 81, 046106.

[47] Gould, S.J., Lewontin, R.C., 1979. The spandrels of san marco and the panglossian paradigm: a critique of the adaptationist programme. Proc. R. Soc, Lond. B 205, 581–598.

[48] Gould, S.J., Vrba, E.S., 1982. Exaptation—a missing term in the science of form. Paleobiology 8, 4–15.

[49] Handcock, M.S., Raftery, A.E., Tantrum, J.M., 2007. Model-based clustering for social networks. Journal of the Royal Statistical Society: Series A (Statistics in Society) 170, 301–354.

[50] Hoel, E.P., Albantakis, L., Marshall, W., Tononi, G., 2016. Can the macro beat the micro? integrated information across spatiotemporal scales. Neuroscience of Consciousness 2016, niw012.

[51] Jackson, M.D., Xu, H., Duran-Nebreda, S., Stamm, P., Bassel, G.W., 2017. Topological analysis of multicellular complexity in the plant hypocotyl. Elife 6, e26023.

[52] Jacob, F., 1977. Evolution and tinkering. Science 196, 1161–1166.

[53] Jonas, K., Chen, Y.E., Laub, M.T., 2011. Modularity of the bacterial cell cycle enables independent spatial and temporal control of dna replication. Current biology 21, 1092–1101.

[54] Kashtan, N., Alon, U., 2005. Spontaneous evolution of modularity and network motifs. Proceedings of the National Academy of Sciences 102, 13773–13778.

[55] Kay, A.C., 1993. The early history of smalltalk. ACM SIGPLAN Notices 28, 69–95.

[56] Kesten, H., et al., 1980. The critical probability of bond percolation on the square lattice equals 1/2. Communications in mathematical physics 74, 41–59.

[57] Klein, B., Hoel, E., 2020. The emergence of informative higher scales in complex networks. Complexity 2020, 1–12.

[58] Knabe, J.F., Nehaniv, C.L., Schilstra, M.J., 2008. Do motifs reflect evolved function?—no convergent evolution of genetic regulatory network subgraph topologies. Biosystems 94, 68–74.

[59] Koonin, E.V., 2016. Splendor and misery of adaptation, or the importance of neutral null for understanding evolution. BMC biology 14, 1–8.

[60] Lacasa, L., Luque, B., Ballesteros, F., Luque, J., Nuno, J.C., 2008. From time series to complex networks: The visibility graph. Proceedings of the National Academy of Sciences 105, 4972–4975.

[61] Levine, J.M., Bascompte, J., Adler, P.B., Allesina, S., 2017. Beyond pairwise mechanisms of species coexistence in complex communities. Nature 546, 56–64.

[62] Macia, J., Vidiella, B., Solé, R.V., 2017. Synthetic associative learning in engineered multicellular consortia. Journal of The Royal Society Interface 14, 20170158.

[63] Madsen, C., Goñi Moreno, A., P, U., Palchick, Z., Roehner, N., Atallah, C., Bartley, B., Choi, K., Cox, R.S., Gorochowski, T., et al., 2019. Synthetic biology open language (sbol) version 2.3. Journal of integrative bioinformatics 16, 20190025.

[64] Maslov, S., Sneppen, K., 2002. Specificity and stability in topology of protein networks. Science 296, 910–913.

[65] Mesoudi, A., 2021. Cultural selection and biased transformation: two dynamics of cultural evolution. Philosophical Transactions of the Royal Society B 376, 20200053.

[66] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., Alon, U., 2002. Network motifs: simple building blocks of complex networks. Science 298, 824–827.

[67] Miransari, M., 2011. Interactions between arbuscular mycorrhizal fungi and soil bacteria. Applied Microbiology and Biotechnology 89, 917–930.

[68] Mirzasoleiman, B., Jalili, M., 2011. Failure tolerance of motif structure in biological networks. PLoS One 6, e20512.

[69] Newberry, M.G., Plotkin, J.B., 2022. Measuring frequency-dependent selection in culture. Nature Human Behaviour 6, 1048–1055.

[70] Newman, M., 2018. Networks. Oxford university press.

[71] Newman, M.E., Girvan, M., 2004. Finding and evaluating community structure in networks. Physical review E 69, 026113.

[72] O'Brien, M.J., Lyman, R.L., 2000. Applying evolutionary archaeology: A systematic approach. Springer Science & Business Media.

[73] Olesen, J.M., Bascompte, J., Dupont, Y.L., Jordano, P., 2007. The modularity of pollination networks. Proceedings of the National Academy of Sciences 104, 19891–19896.

[74] Otte, E., Rousseau, R., 2002. Social network analysis: a powerful strategy, also for the information sciences. Journal of information Science 28, 441–453.

[75] Perna, A., Jost, C., Couturier, E., Valverde, S., Douady, S., Theraulaz, G., 2008. The structure of gallery networks in the nests of termite cubitermes spp. revealed by x-ray tomography. Naturwissenschaften 95, 877–884.

[76] Poisot, T., Gravel, D., 2014. When is an ecological network complex? connectance drives degree distribution and emerging network properties. PeerJ 2, e251.

[77] Radicchi, F., Castellano, C., 2016. Beyond the locally treelike approximation for percolation on real networks. Physical Review E 93, 030302.

[78] Ratcliff, W.C., Denison, R.F., Borrello, M., Travisano, M., 2012. Experimental evolution of multicellularity. Proceedings of the National Academy of Sciences 109, 1595–1600.

[79] Rives, A.W., Galitski, T., 2003. Modular organization of cellular networks. Proceedings of the national Academy of sciences 100, 1128–1133.

[80] Rizzoli, A., Tagliapietra, V., Cagnacci, F., Marini, G., Arnoldi, D., Rosso, F., Rosà, R., 2019. Parasites and wildlife in a changing world: The vector-host-

pathogen interaction as a learning case. International Journal for Parasitology: Parasites and Wildlife 9, 394–401.

[81] Sanders, D., Thébault, E., Kehoe, R., Frank van Veen, F., 2018. Trophic redundancy reduces vulnerability to extinction cascades. Proceedings of the National Academy of Sciences 115, 2419–2424.

[82] Schaerli, Y., Munteanu, A., Gili, M., Cotterell, J., Sharpe, J., Isalan, M., 2014. A unified design space of synthetic stripe-forming networks. Nature communications 5, 4905.

[83] Schlosser, G., Wagner, G.P., 2004. Modularity in development and evolution. University of Chicago Press.

[84] Schülke, C., Ricci-Tersenghi, F., 2015. Multiple phases in modularity-based community detection. Physical Review E 92, 042804.

[85] Simon, H.A., 1955. On a class of skew distribution functions. Biometrika 42, 425–440.

[86] Simon, H.A., 1962. The architecture of complexity. Proceedings of the American philosophical society 106, 467–482.

[87] Solé, R., 2011. Phase transitions. volume 3. Princeton University Press.

[88] Solé, R., Amor, D.R., Duran-Nebreda, S., Conde-Pueyo, N., Carbonell-Ballestero, M., Montañez, R., 2016. Synthetic collective intelligence. Biosystems 148, 47–61.

[89] Solé, R., Ollé-Vila, A., Vidiella, B., Duran-Nebreda, S., Conde-Pueyo, N., 2018. The road to synthetic multicellularity. Current Opinion in Systems Biology 7, 60–67.

[90] Solé, R.V., Valverde, S., 2004. Information theory of complex networks: on evolution and architectural constraints, in: Complex networks. Springer, pp. 189–207.

[91] Solé, R.V., Valverde, S., 2006. Are network motifs the spandrels of cellular complexity? Trends in ecology & evolution 21, 419–422.

[92] Solé, R.V., Valverde, S., 2008. Spontaneous emergence of modularity in cellular networks. Journal of The Royal Society Interface 5, 129–133.

[93] Solé, R.V., Valverde, S., Casals, M.R., Kauffman, S.A., Farmer, D., Eldredge, N., 2013. The evolutionary ecology of technological innovations. Complexity 18, 15–27.

[94] de Solla Price, D., 1965. Networks of scientific papers. Science (New York, NY) 149, 510–515.

[95] Sporns, O., Kötter, R., 2004. Motifs in brain networks. PLoS biology 2, e369.

[96] Staniczenko, P.P., Kopp, J.C., Allesina, S., 2013. The ghost of nestedness in ecological networks. Nature Communications 4, 1–6.

[97] Steadman, P., 2008. The evolution of designs: biological analogy in architecture and the applied arts. Routledge.

[98] Stouffer, D.B., Bascompte, J., 2011. Compartmentalization increases food-web persistence. Proceedings of the National Academy of Sciences 108, 3648–3652.

[99] Szathmáry, E., 2015. Toward major evolutionary transitions theory 2.0. Proceedings of the National Academy of Sciences 112, 10104–10111.

[100] Timóteo, S., Albrecht, J., Rumeu, B., Norte, A.C., Traveset, A., Frost, C.M., Marchante, E., López-Núñez, F.A., Peralta, G., Memmott, J., et al., 2022. Tripartite networks show that keystone species can multitask. Functional Ecology .

[101] Tononi, G., Sporns, O., Edelman, G.M., 1999. Measures of degeneracy and redundancy in biological networks. Proceedings of the National Academy of Sciences 96, 3257–3262.

[102] Toro-Delgado, E., Hernández-Roldán, J., Dincă, V., Vicente, J.C., Shaw, M.R., Quicke, D.L., Vodă, R., Albrecht, M., Fernández-Triana, J., Vidiella, B., et al., 2022. Butterfly–parasitoid–hostplant interactions in western palaearctic hesperiidae: a dna barcoding reference library. Zoological Journal of the Linnean Society 196, 757–774.

[103] Valverde, S., 2007. Crossover from endogenous to exogenous activity in open-source software development. Europhysics Letters 77, 20002.

[104] Valverde, S., 2016. Major transitions in information technology. Philosophical Transactions of the Royal Society B: Biological Sciences 371, 20150450.

[105] Valverde, S., 2017. Breakdown of modularity in complex networks. Frontiers in physiology 8, 497.

[106] Valverde, S., 2021. The long and winding road: Accidents and tinkering in software standardization. Metode Science Studies Journal 0, 91–97. URL: https://revistas.uv.es/index.php/Metode/article/view/16112, doi:10.7203/metode.11.16112.

[107] Valverde, S., Ferrer-i Cancho, R., Solé, R.V., 2002. Scale-free networks from optimal design. Europhysics Letters 60, 512.

[108] Valverde, S., Corominas-Murtra, B., Perna, A., Kuntz, P., Theraulaz, G., Solé, R.V., 2009. Percolation in insect nest networks: evidence for optimal wiring. Physical Review E 79, 066106.

[109] Valverde, S., Piñero, J., Corominas-Murtra, B., Montoya, J., Joppa, L., Solé, R., 2018. The architecture of mutualistic networks as an evolutionary spandrel. Nature ecology & evolution 2, 94–99.

[110] Valverde, S., Solé, R., 2005a. Logarithmic growth dynamics in software networks. Europhysics Letters 72, 858.

[111] Valverde, S., Solé, R.V., 2004. Internet's critical path horizon. The European Physical Journal B 38, 245–252.

[112] Valverde, S., Solé, R.V., 2005b. Network motifs in computational graphs: A case study in software architecture. Physical Review E 72, 026107.

[113] Valverde, S., Solé, R.V., 2007. Hierarchical small-worlds in software architecture. Dynamics of Continuous Discrete Impulsive Systems Series B: Applications and Algorithms (Special Issue on Software Engineering and Complex Networks, eds K. L. Teo and X. Liu) 14, 1–11.

[114] Valverde, S., Solé, R.V., 2013. Networks and the city. Architectural Design 83, 112–119.

[115] Valverde, S., Solé, R.V., Bedau, M.A., Packard, N., 2007. Topology and evolution of technology innovation networks. Physical Review E 76, 056118.

[116] Valverde, S., Vidiella, B., Montañez, R., Fraile, A., Sacristán, S., García-Arenal, F., 2020. Coexistence of nestedness and modularity in host–pathogen infection networks. Nature Ecology & Evolution 4, 568–577.

[117] Vidiella, B., Guillamon, A., Sardanyés, J., Maull, V., Pla, J., Conde, N., Solé, R., 2021. Engineering self-organized criticality in living cells. Nature communications 12, 4415.

[118] Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world'networks. nature 393, 440–442.

[119] Weinstock, M., 2010. The Architecture of Emergence: The Evolution of Form in Nature and Civilisation. Wiley.

[120] Weinstock, M., 2011. The metabolism of the city: the mathematics of networks and urban surfaces. Architectural Design 81, 102–107.

[121] Weitz, J.S., Poisot, T., Meyer, J.R., Flores, C.O., Valverde, S., Sullivan, M.B., Hochberg, M.E., 2013. Phage–bacteria infection networks. Trends in microbiology 21, 82–91.

[122] Wells, E., Wilkinson, M., Wood, P., Scanlan, C., 2007. The use of macroalgal species richness and composition on intertidal rocky seashores in the assessment of ecological quality under the european water framework directive. Marine pollution bulletin 55, 151–161.

[123] Wilson, E.O., Hölldobler, B., 2009. The superorganism: the beauty, elegance, and strangeness of insect societies. New York: W. W Norton & Company.

[124] Woodger, J.H., Tarski, A., Floyd, W.F., 1940. The axiomatic method in biology. Journal of Unified Science (Erkenntnis) 8.

[125] Yule, G.U., 1925. II.—A mathematical theory of evolution, based on the conclusions of dr. jc willis, fr s. Philosophical transactions of the Royal Society of London. Series B, containing papers of a biological character 213, 21–87.