J Biosci (2022)47:55
DOI: 10.1007/s12038-022-00285-4

Review

Network approach to understand biological systems: From single to multilayer networks

Sayantoni Chaudhuri and Ashutosh Srivastava*

Discipline of Biological Engineering, Indian Institute of Technology, 382355 Gandhinagar, India

**Corresponding author (Email, ashutosh.s@iitgn.ac.in)*

MS received 18 October 2021; accepted 20 April 2022

Network theory has led to the abstraction of many real-world systems and enabled their modelling as simple networks comprising nodes and edges. In particular, in the field of biological sciences, network theory provides a robust framework to capture the complexity inherent to biological systems. Networks in biology have been modelled at different scales, starting from cells to population levels. These models have provided crucial insights into the evolution, mechanism, and functions of several biological systems. However, most natural and engineered systems are composed of multiple subsystems and layers of connectivity. A multilayer network paradigm has proven useful in understanding such systems. Here, we have briefly introduced the network formalism of modelling biological systems at various levels. This is followed by an introduction to multilayer networks. Multilayer networks have been utilized to model biological systems at multiple scales ranging from protein–protein interactions, transcription and metabolic networks, to ecological networks involving interactions between species. Recent advances in studying the structure and dynamics of such multilayer networks have enabled a better understanding of the complexity in these biological systems. Finally, we discuss the recent advances in studying the structure and dynamics of such multilayered networks followed by the challenges and future prospects.

Keywords. Biological networks; interdependent networks; multilayer networks

1. Introduction

The concept of networks has been used to model real systems for about three centuries now and can be traced to the famous Konisberg Bridge Problem by Euler in 1736, who reduced the problem to an analysis of the degrees of vertices on a graph. Since then, several theoretical advances in the understanding of networks have led to a well-established field of network science that finds application in almost every area of research.

A network, interchangeably called a graph, is a collection of vertices joined by edges. Mathematically, a graph consists of nodes (or vertices), edges, and a mapping function, represented as a 3-tuple:

$$G = (N, L, f)$$

where N is a set of nodes, L is a set of links, and f is a mapping function that maps links onto node pairs. Several different parameters are used to describe network properties, some of which have been described in box 1. Graphs can be classified in various ways based on specific properties of the nodes and edges (box 1). In an undirected graph G = (N, L, f), the mapping function describing the link between two vertices, f_1, is equal to a mapping function f_2 that reverses the link:

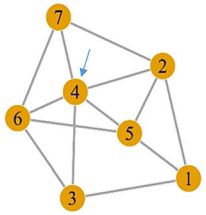
This article is part of the Topical Collection: Emergent dynamics of biological networks.

http://www.ias.ac.in/jbiosci
Published online: 06 September 2022

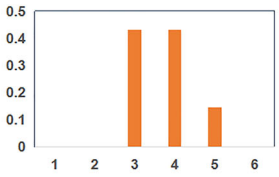
Network Measures

1. Degree

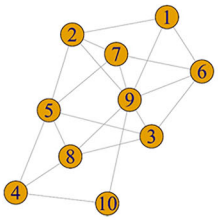
The degree of a node *k*, in a network is the number of edges adjacent to it.
Here *k* (node 4) = 5



2. Degree Distribution



The degree distribution $P(k)$ of a network is the fraction of nodes with degree *k*. Here, the degrees are k_1=3, k_2=4, k_3=5. The corresponding degree distribution is: $P(k_1) = 3/7 = 0.43$, $P(k_2) = 3/7 = 0.43$, and $P(k_3) = 1/7 = 0.14$



$G = (V, E)$

$n(V) = 10$

3. Clustering Coefficient

Clustering coefficient can be calculated as the fraction of triangles formed by a node's neighbours, given as: $C_i = T(i)/\binom{k_i}{2}^{-1}$ where k_i is the degree of node i and $T(i)$ is the number of triangles with node i as a vertex. In the graph G shown, $C_i\ (node\ 1) = 2/3 = 0.67$

4. Closeness Centrality

Closeness centrality indicates how close a node is to all other nodes in the network. It is given as: $CC(i) = \frac{(N-1)}{\sum_{j=1}^{n} d(i,j)}$ where $j = 1,2,..,10$ and N = degree(i). In the graph G shown, $CC(node\ 1) =$ 9/16 = 0.562

5. Betweenness Centrality

Betweeness centrality quantifies the extent to which a node falls on paths between other nodes. It is given as: $BC\ (i) = \frac{s_t(i)}{s_t}$ where s_{uv} IS the shortest path from one node u to another node v. In the graph G shown, $BC(node\ 1) =$ 0.33

6. Eigenvector Centrality

Eigenvector centrality measures the influence of a node, based on the assumption that a certain node is important if it is linked to other important nodes. It is given as: $C_e(i) = \frac{1}{\lambda} \sum_{j=1}^{n} A_{j,i}\ C_e(i)$ where $A_{j,i}$ is the adjacency matrix of graph G, j is the set of neighbours of i and λ is a constant. In the graph G shown, $C_e(node\ 1) =$ 0.55

Box 1. Network parameters commonly used to define networks.

$f_1 = [e_1 : v_1 \leftrightarrow v_2,\ e_2 : v_1 \leftrightarrow v_3,$
$e_3 : v_2 \leftrightarrow v_3]$ is equivalent to
$f_2 = [e_1 : v_2 \leftrightarrow v_1,\ e_2 : v_3 \leftrightarrow v_1,$
$e_3 : v_3 \leftrightarrow v_2]$

Directed graphs have edges assigned with a direction, and the mapping function f_1 of a directed graph G = [N, L, *f*] would not be equivalent to a mapping function f_2 with reversed links.

Random graphs described by Solomonoff and Rapoport (1951) and then later by Erdös and Rényi (1959) consist of *n* vertices, connected by *m* undirected edges chosen randomly from the set of all possible edges. This realization of the Erdös–Rényi (ER) model is represented as the G(N, L) model, N being the number of nodes and L being the edges. A second realization of random graphs is the G(N, p) model, where each pair out of N nodes is connected with a probability *p* (Gilbert 1959). Figure 1a shows an ER network comprising $n = 10$ vertices and $m = 16$ edges. The degree distribution of the ER random network follows a binomial distribution that becomes approximately Poissonian for large networks (Newman *et al.* 2001). The random network model has been applied extensively in understanding the theoretical basis of several complex systems. Real networks, however, follow a more complex connectivity pattern with several showing power law behaviour in their degree distribution. Networks with a power law degree distribution are essentially scale-free, which means that upon addition of new nodes, the overall structure of the network remains unchanged. Many biological networks follow a scale-free topology (Strogatz 2001; Albert 2005). Another type of network model generally observed in real systems, such as a power grid, account for the co-existence of high local clustering as well as the presence of a small average path length (Watts and Strogatz 1998). Watts and Strogatz found that rewiring a ring lattice of *n* vertices connected to their *k* nearest neighbours randomly with a probability of *p,* where $0<p<<1$, leads to a 'small-world' network with high clustering coefficient and small characteristic path length (Watts and Strogatz 1998). This network paradigm has been used extensively to study the structure and dynamics of real networks such as protein interaction networks and metabolic networks (Watts and Strogatz 1998; Telesford *et al.* 2011). To arrive at more realistic models, Barabási and Albert (1999) proposed a simple network model that gives rise to a scale-free degree distribution which is essentially based on two components: (a) *growth of a network's nodes*, instead

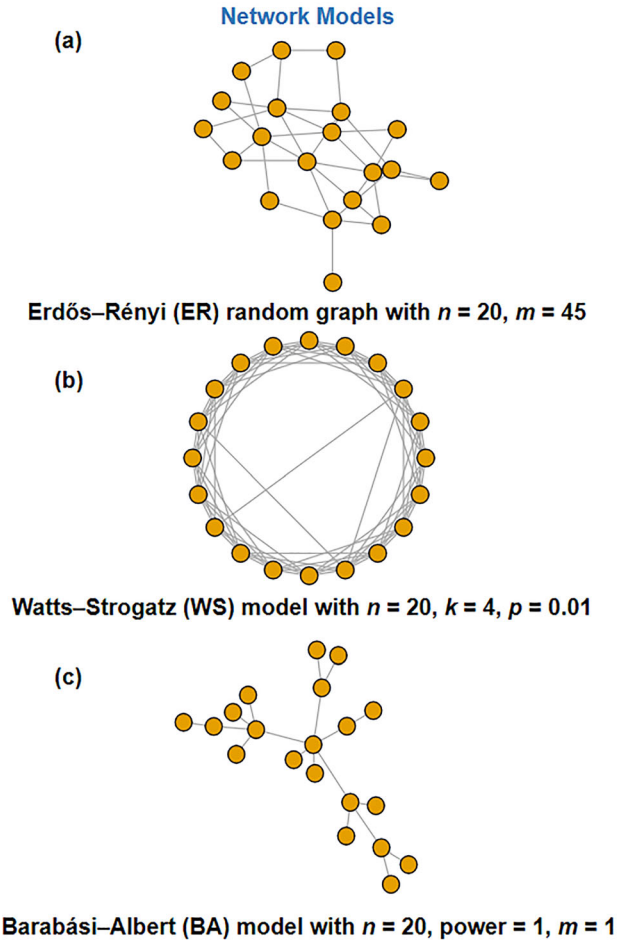


Figure 1. Three different types of network models. (**a**) ER random graph with number of nodes (n) = 20 and number of edges (m) = 45. (**b**) WS model with number of nodes (n) = 20, number of nearest neighbours in ring topology each node is connected to (k) = 4, and probability of rewiring each edge (p) = 0.01. (**c**) BA model with number of nodes (n) = 20 and number of edges added at each time step (m) = 1.

of an assumption of a fixed number of vertices, and (b) *preferential attachment,* which takes into account the fact that incorporation of new nodes depends on the degree of the current node. To generate such a 'growing network' based on these two features, the initial network starts with a small number of nodes m_0. New nodes are added to this network such that each newly added node creates m ($\leqq m_0$) links to different nodes already present in the network. This attachment is further constrained by a probability Π that an incoming node will connect to an already present node i based on its degree $d(i)$ such that $\Pi(d(i)) = d(i)/\sum_j d(j)$ (Barabási and Albert 1999). The three network models discussed here are built on different principles and have been able to describe real-world network features, both local and global, effectively. Figure 1 illustrates these three network models, namely, (a) the Erdös–Rényi (ER) random graph model, (b) the Watts–Strogatz (WS) model, and (c) the Barabási–Albert (BA) model. With the help of box 2, we summarize some attributes that are often used to characterize, analyse, and compare networks. These include the power law and scale freeness, modularity, assortativity and robustness.

Applications of networks now span across almost all domains of sciences. Particularly, the network paradigm provides a robust framework to capture the complexity inherent to biological systems. Networks at various spatial scales ranging from subcellular to ecological have provided crucial information about these systems. At the cellular level, metabolic and biochemical networks describing the conversion of various substrates through enzyme-driven reactions have been crucial in not only understanding the evolution of pathways and their function but also in determining drug targets (Fang *et al.* 2020). At the molecular level, the protein–protein interaction networks have been mapped for almost the whole proteome of humans, providing crucial insights regarding cellular function and diseases (Luck *et al.* 2020). At a much larger scale, interactions between various species in the natural environment are represented as ecological networks and provide crucial insights regarding keystone species and conservation (Bascompte 2010).

Although in the past years, network theory has successfully characterized the interaction among the constituents of a variety of systems, ranging from biological to technological, as well as social, it has been realized that most natural and engineered systems are composed of multiple subsystems and layers of connectivity. For instance, systems such as water and food supply, communications, fuel, financial transactions, and power stations exhibit a degree of interdependence and require a model that captures the interaction among various networks (Gao *et al.* 2014). Biological systems show similar interdependence, and hence, a 'multilayered networks' paradigm provides a feasible way to model the complexity of biological systems. In this review, we first describe various biological networks by grouping them into three broad categories, namely, (a) networks at the subcellular level, (b) networks at the tissue level and (c) networks at the population level, while briefly going over types, recent advances, and applications. We then provide a short introduction to the multilayer paradigm of complex system modelling, followed by recent advances in applying these concepts to biological systems.

Network Properties

1. Power Law and Scale-Freeness: Scale-free networks are those that do not depend on the number of nodes present in the network and the network "grows" when new nodes are added. For such networks, the distribution of edges to the nodes follows a power law given by $P(k) \sim k^{-\gamma}$ where γ is the degree component, and $2 < \gamma < 3$.

2. Six Degrees of Separation: This refers to a theory in a social network based on Milgram's experiment that no two people are five or less than five social connections away from each other.

3. Shortest Path and Mean Path Length: Shortest path is the path between two nodes that covers the minimum number of edges and the mean path length is the average of the shortest path lengths of all the node-pairs.

4. Modularity: Modularity refers to a property of networks wherein some nodes and edges form communities that are highly inter-connected among each other than to the rest of the network.

5. Assortative Mixing: It refers to the preference to associate with nodes that have similar characteristics, for instance assortative mixing by degree refers to the tendency of high degree nodes to connect to similar high degree nodes, and low degree nodes to connect to similar low degree nodes.

6. Robustness: It refers to the ability of networks to resist failures upon perturbations of nodes an edges when removed from a network. Mathematically, it can be defined as

$$r = \frac{N_a}{N_b}$$

where N_a and N_b are the size of the largest connected component after and before removal of node.

Box 2. Network properties often used to characterize, analyze, and compare networks.

2. Networks in biological systems

2.1 *Networks at the sub-cellular level*

All the structural and functional processes in the cell are governed by interactions among genes, proteins, and other molecules. Network approach is a popular paradigm to investigate, model, and understand these cellular processes. In the following subsections, we have briefly described some of these networks. Figure 2 illustrates a schematic representation of how some biological systems can be modelled as simple network systems.

2.1.1 *Protein contact networks:* Proteins play a crucial role in and are the building blocks of most biological processes that takes place inside a cell. They comprise a sequence of amino acids that pack together into a three-dimensional structure with the help of various covalent and non-covalent interactions that include disulphide bonds, ionic interactions, hydrogen bonds, hydrophobic interactions, covalent interactions, and van der Waals interactions. The amino acids and their interactions within the tertiary structure can be modelled as networks, referred to as protein contact networks (PCNs) to provide additional insights into the structural and functional roles of interacting residues as well as to understand structure–function relationships (figure 2a) (Doncheva *et al.* 2011). PCN can be defined at a coarse-grained level with nodes as the Cα or the Cβ atoms of the amino acids while the edges can be defined based on the physical distance between these elements (Dokholyan *et al.* 2002; Vendruscolo *et al.* 2002; Atilgan *et al.* 2004; Bagler and Sinha 2005; Di Paola *et al.* 2013). This simple representation of protein structures as unweighted, undirected networks has provided some important insights into the properties of protein structures in terms of their connectivity and folding (Dokholyan *et al.* 2002; Vendruscolo *et al.* 2002; Bagler and Sinha 2007). Coarse-grained Cα atom based PCNs were recently used to understand subtle conformational changes caused by the emergence of drug resistance in HIV-1 reverse transcriptase (Srivastava *et al.* 2020). Analysis of contacts and network parameters such as degree, betweenness centrality and eigenvector centrality (box 1) has been used to understand allostery (Di Paola and Giuliani 2015; Srivastava and Sinha 2017). The information regarding side-chain interactions is implicit in this simplistic representation of protein structures by way of assuming a longer cut-off (7–8Å) in defining the interactions between residues. However, the information pertaining to specific interactions between side chains can be explicitly modelled in the construction of PCNs (Brinda and Vishveshwara 2005). Such networks have been extensively used to understand protein structure–function relationships in various systems such as effect of ligand binding and

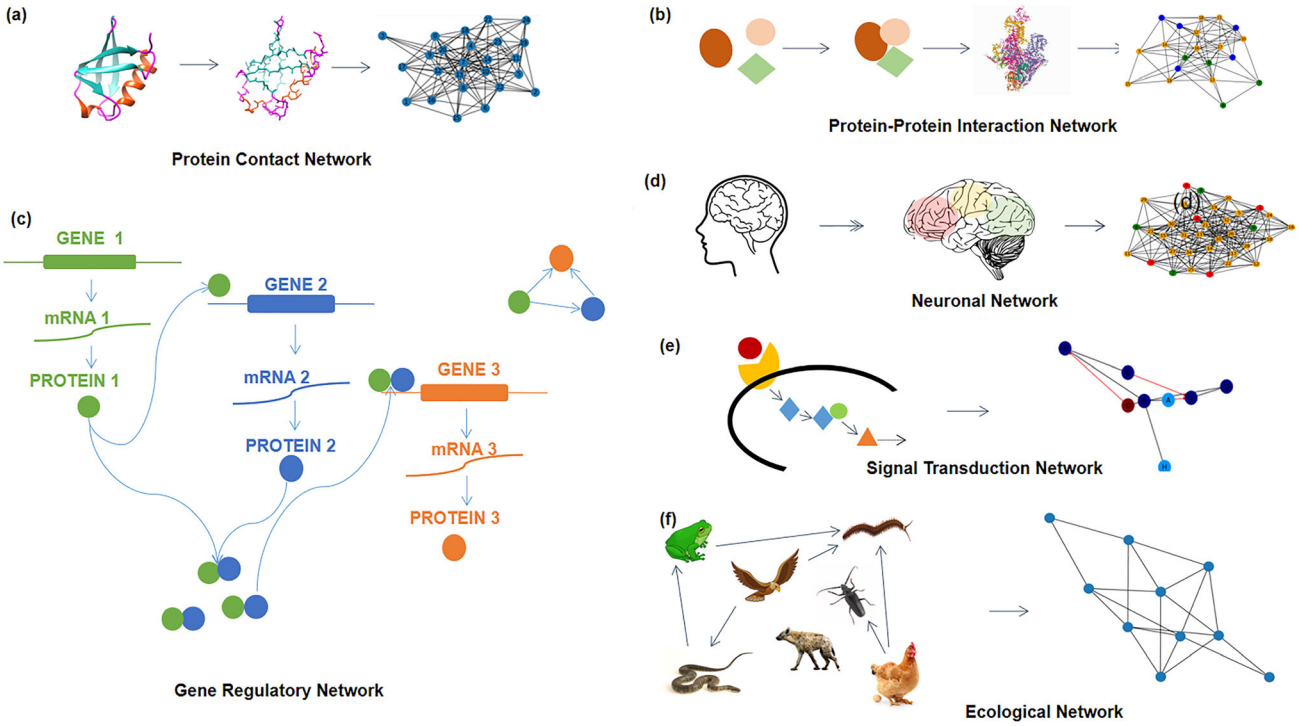


Figure 2. A schematic representation of some biological networks as simple network systems: (**a**) Protein contact network, (**b**) protein–protein interaction network, (**c**) gene regulatory network, (**d**) signal transduction network, (**e**) neuronal network, and (**f**) ecological network.

allostery in tRNA synthetases (Ghosh *et al.* 2007, 2011; Sathyapriya and Vishveshwara 2007) and protein–DNA interactions (Sathyapriya *et al.* 2008). Using the interactions between all atoms of the residues, Greene and Higman (2003) defined two types of networks exhibiting slightly different properties, across a variety of protein folds. The network with all the interactions showed a bell-shaped Poisson distribution of degrees, whereas the network formed of only long-range interactions (between residues distant on primary sequence) showed a power law distribution of degree, suggesting it to be a scale-free network. Taking a different approach towards defining the protein structure networks, Vijaybhaskar and Vishveshwara (2010) used interaction energies between residues, obtained from equilibrium molecular dynamics simulations, to define the edges between nodes represented by the residues. These protein energy networks were used to understand the stability of the TIM barrel fold despite the divergence in sequence (Vijayabaskar and Vishveshwara 2012). Representing protein structures as networks has led to the prediction of functional residues (Amitai *et al.* 2004), active sites (Sol *et al.* 2006), and hot-spot residues for protein–protein interaction (Sol and O'Meara 2005). Chakraborty and Parekh (2014, 2020) used spectral analysis of protein structure networks to identify tandem repeats in protein structures. Using modularity (box 2) measures on Cα-based protein contact networks, Yalamanchili and Parekh (2009) identified domains in the protein structures.

Considering the utility of network representations of protein structures and their varied applications, dedicated software such as RINalyzer (Doncheva *et al.* 2011), AMINONET (Aftabuddin and Kundu 2010) and webservers such as NAPS (Chakrabarty and Parekh 2016) have been developed to model and visualize network models of protein structures.

2.1.2 *Protein–protein interaction networks:* The functionally active form of a protein rarely exists in isolations, but rather in close association with other biomolecules. These interactions can be transient or stable depending on multiple parameters including function and environment. Several biophysical as well as biochemical methods have been used to determine protein–protein interactions (Zhou *et al.* 2016). These include affinity chromatography, co-immunoprecipitation, yeast two-hybrid system, cross-linking mass spectrometry, and fluorescence resonance energy transfer. Conventionally, the protein–protein interaction network (PPIN) can be modelled via graphs whose nodes

represent proteins and whose edges connect pairs of interacting proteins (figure 2b). Edge weights may be used to incorporate reliability information associated with the corresponding interactions. PPINs provide a lucid means to understand the functional organization of the proteome as well as complex cellular phenomena (Stelzl *et al.* 2005; Gandhi *et al.* 2006). With the application of experimental techniques such as yeast two-hybrid in a high-throughput manner, it has been possible to map protein–protein interactions on a genome wide scale, leading to the creation of complete protein–protein interaction maps, often referred to as the 'interactome', for several organisms from bacteria to higher eukaryotes. Recently, a reference map of the human binary interactome has been defined using yeast two-hybrid experiments and other assays that includes more than 90% of the human protein-encoding genes (Luck *et al.* 2020). The analyses of such PPINs have revealed their crucial topological features. Yook *et al.* (2004) analysed large PPINs from *Saccharomyces cerevisiae* and found that their degree distribution followed a power law, suggesting a scale-free organization of the PPINs. It directly follows that the most highly connected proteins in the cell are the most important for its survival and PPINs largely remain insensitive to random removal of single vertices, but are particularly sensitive to the targeted removal of hubs; this is also known as the centrality–lethality rule (Jeong *et al.* 2001). PPINs have a wide range of applications in modern day biology: discovering novel protein functions (Sharan *et al.* 2007), identifying functionally coherent modules (Spirin and Mirny 2003; Dittrich *et al.* 2008), and identifying conserved molecular interaction patterns (Sharan *et al.* 2005; Jaeger *et al.* 2010). Since proteins play a major role in biological functions, their interactions determine the molecular mechanisms that control the healthy and diseased states in organisms. Therefore, the molecular basis of diseases can be deciphered through protein interaction networks, which can further help in their prevention, diagnosis, and treatment (Kann 2007; Barabási *et al.* 2011). The experimentally derived as well as predicted information on protein–protein interactions has been organized in several databases. Some of these include BioGRID (the Biological General Repository for Interaction Datasets), which currently holds ∼1,740,000 interactions curated from both high-throughput datasets and individual focused studies (Stark *et al.* 2006); the STRING database, which currently contains information regarding 2,45,84,628 proteins from 5090 organisms (Szklarczyk *et al.* 2021); IntAct, a comprehensive database of individual interactions comprising over 4,00,000 interactions and interactomes of 15 different organisms (Orchard *et al.* 2013), and HPRD (Human Protein Reference Database), which that provides information on human protein interactions (Keshava Prasad *et al.* 2009).

2.1.3 *Gene regulatory networks (GRNs):* All cellular functions are dependent on the well-regulated expression of genes. This regulation of gene expression happens at multiple levels including transcriptional, translational, as well as post translational. The transcription factors (TFs), including activators and repressors along with their target binding sites, form intricate networks to control gene expression. Regulatory RNAs such as micro-RNAs or long non-coding RNAs also participate in the regulation of gene expression through various mechanisms (Statello *et al.* 2021). In the graph-theoretical representation of these GRNs, the regulators such as TFs, regulatory RNAs, and target binding sites are represented as nodes and the interactions by directed edges. The computational investigation of these GRNs involves inference of the networks of interactions from gene expression data and then understanding the phenotype or diseased state as an emergent property of these networks (Mercatelli *et al.* 2020). By studying and analyzing the reconstructed GRNs of four model organisms, *Saccharomyces cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster* and *Arabidopsis thaliana,* it was reported that GRNs exhibit scale-free degree distribution (Ouma *et al.* 2018). Complex disorders such as cancer often exhibit altered gene regulation, and analysis of the perturbation in GRNs can provide insights for diagnosis or prognosis (Emmert-Streib *et al.* 2014). A database comprising gene regulatory networks across various human conditions has been developed recently and comprises 12,468 GRNs from 36 human tissues and 28 cancers (Guebila *et al.* 2022).

2.1.4 *Metabolic networks:* Metabolic networks refer to a class of networks that represent the pathways and reactions modifying various metabolites within organisms. Metabolic networks can be defined in different ways based on the nodes and the type of interactions. In the substrate graph model, the metabolites or compounds are the nodes and the reactions converting one substrate to a product are represented by edges. In the reaction graph model, the reactions are the nodes and they are connected if the same compound is a substrate in one and a product in another reaction. In an enzyme graph, the nodes represent enzymes and the edges represent a common compound present in the reactions catalysed by the two enzymes. Metabolic networks can also be represented as bipartite graphs (with edges

strictly between two sets of nodes), with the two sets of nodes represented by substrates and enzymes. Large-scale sequencing efforts over the course of past several years have enabled the curation of gene, protein, and reaction data for several organisms with the possibility of construction of global metabolic networks (Ma *et al.* 2007; Herrgard *et al.* 2008; Orth *et al.* 2011). Substrate graphs of large metabolic networks were found to show small-world and scale-free characteristics (Wagner and Fell 2001). In an analysis of microbial metabolic networks as directed reaction centric graphs, the degree distribution was found to follow a power law distribution, suggesting a scale-free nature (Kim *et al.* 2019). Several studies have attempted to model metabolic interactions in order to determine the effect of knockout of certain enzymes crucial in metabolic pathways of pathogenic organisms, microbial interactions and their ecological properties (Roume et al. 2015; Song *et al.* 2017; Muller *et al.* 2018), and host–microbe interactions and their relation to disease phenotypes (Heinken and Thiele 2015).

2.1.5 *Signal transduction networks:* Cells respond to environmental cues by activating a cascade of signalling that leads to change in the transcriptional landscape of the cells. This involves sequential interaction of receptors, enzymes, and transcription factors This network of molecules and interactions, referred to as a signal transduction network (STN), can be represented as a graph where the vertices are genes and the directed edges denote activating or repressing effects on transcription (figure 2d). STNs can be constructed from gene expression or protein–protein interaction data. Algorithms have been developed that utilize these data to construct STNs (Zhao *et al.* 2008; Supper *et al.* 2009; Wang *et al.* 2011). Modelling the signal transduction as networks has provided crucial insights into some emergent properties such as signal integration, bistability, robustness, and ultrasensitivity, owing to the topology and motifs in the network (Bhalla and Iyengar 1999; Azeloglu and Iyengar 2015).

2.2 *Networks at the tissue level*

In recent years, a vast majority of research has been done on various cellular-level networks such as protein interactions networks, gene regulatory networks, and signalling networks, however, relatively little is known about how all of these cellular-level networks come together to link higher-level systems such as the immune system and the brain. Such systems are conventionally evaluated as a hierarchy comprising different types of components, both on a microscopic and a macroscopic scale.

2.2.1 *Networks of the immune system:* The immune system involves multiple types of cells that work together in a coordinated manner to defend the host against infections. Although a vast amount of information is available about individual cells, few studies have addressed the immune system as a network or collection of network modules. A systems-level approach to understand the immune system has led to the emergence of 'systems immunology' to better understand the immune system and its disorders (Yu *et al.* 2019a). Hao Shi *et al.* (2020) have recently reviewed the network approach to study and analyse the immune system with a focus on studies of the signalling and transcriptional landscape, as well as cell–cell communication in hematopoiesis, adaptive immunity, and tumor immunology.

2.2.2 *Neuronal networks:* The brain is a collection of a range of different components including molecules, receptors, ion channels, synapses, and neurons that can be represented as networks. Neuronal networks can be defined as consisting of 'nodes', represented by neurons that are interconnected by a set of 'edges' which can represent functional, structural, or effective connections between different regions of the brain (figure 2e) (Friston 1994; Sporns *et al.* 2004). The information to derive these connections can be based on analysis of neuroimaging data. Neuronal networks have shown to exhibit power law and scale-free behaviour, as confirmed in several studies (Young 1992; Hilgetag *et al.* 2000; Sporns 2003; Sporns and Zwi 2004). Neuronal networks generally have heterogeneous or broad-scale degree distributions, meaning that the probability of a highly connected hub is higher than in a comparable random network. Bassett and Bullmore (2006) argued that a small-world topology can support both segregated/specialized and distributed/integrated information processing, and is economical, tending to minimize wiring costs while also supporting high dynamical complexity, which is why brain functional networks tend to assume small-world properties.

2.3 *Networks at the population level*

Networks are a convenient means of abstraction of complex emergent properties, and have been widely adopted in the studies of ecological systems. Ecological

systems are often represented as networks, typically as food webs, which are networks of consumer–resource interactions between groups of organisms (Montoya *et al.* 2006; Bascompte 2010). In terms of network theory, the nodes are represented by individual species and the edges are inter-species interactions (figure 2f) that can be antagonistic such as in the case of competition and predation or mutualistic, representing cooperation between species (Bascompte 2010). Modelling of interactions between species as networks has enabled deeper understanding of the biodiversity (Pascual-García and Bastolla 2017) and robustness of ecological systems under perturbations (Lu *et al.* 2016), and the role of keystone species in the overall stability of ecosystems.

3. From single networks to multilayer networks

Most real-world systems include multiple subsystems and layers of connectivity, and it is important to take such 'multilayer' features into account in order to improve our understanding of complex systems. With advances in research on complex systems, it has become increasingly important to move beyond simple graphs and investigate more complicated and realistic frameworks (Kivelä *et al.* 2014). Although multilayer networks have been studied for many years, a general lack of convention in terminology is observed. Kivela *et al.* (2014) described 26 different types of multilayer networks based on the properties and constraints that they had in their representation. To represent systems that consist of networks at multiple levels or with multiple types of edges, we consider structures that have layers in addition to nodes and edges. In other words, a multilayer network has a set of N nodes and each node is assigned a type from a set of M types. An in-depth mathematical representation of the same can be found in Boccaletti *et al.* (2014). In this review, we shall highlight a few popularly used multilayer networks to study real-world systems, especially biological systems, namely, multiplex networks, interdependent networks, and networks of networks (figure 3). *Multiplex networks* are a special type of multilayer network comprising a fixed set of nodes connected by different types of links. They are conventionally used to describe the interactions between the same set of nodes, and in which each 'layer' represents a different type of interaction, as notably seen in social networks, where the same set of people can be related by different sets of interactions such as friendships, collaborations, or co-authorships, or they can be connected via different modes such as phone, email, online forums, conferences, or chats. Gene co-expression networks, protein interaction networks can also be expressed in 'layers' depending on the many cellular interactions possible. *Interdependent networks* are a collection of different networks whose nodes are interdependent on each other, and often, the functional properties of nodes of one layer depend on another. For example, real-world interdependencies such as electric power, natural gas and petroleum production and distribution, telecommunications (information and communications) characterized by multiple connections between infrastructures, feedback and feed-forward paths, and intricate, branching topologies have been modelled as interdependent networks and studied for their resilience, cascading failures, and other dynamic properties (Rinaldi *et al.* 2001; Buldyrev *et al.* 2010; Duan *et al.* 2019). In *interconnected networks*, the edges that connect different networks need not indicate dependency relations as in interdependent networks. For *networks of networks,* as generalized by Gao *et al.* (2014), each node is itself a network itself and each link represents a partially dependent pair of networks. Or, in other words, they are multilayer networks with a 'super-network'. Such a super-network can be very useful to determine uniquely which set of layers can be connected by interlinks. Modelling of complex systems as multilayer networks also brought forward certain challenges with respect to analysis and visualization of such networks. A software tool named MuxViz has been developed to overcome some of these challenges (De Domenico *et al.* 2015).

3.1 *Structure of multilayer networks*

In order to extend the existing metrics designed for single networks to multilayer networks, we first try to mathematically define them. A simple network can be represented by a graph which is a 3-tuple, as discussed in section 1. This simple notation can be extended to introduce *layers*. Thus, a multilayer network then becomes a 'quadruple' and is defined as

$$M = (V_M, E_M, W, L)$$

where V_M is the set of layers in which a node is present, E_M is the set containing the set of all pairs of possible combinations of nodes and layers, and L is the set of layers defined by d aspects. (Kivelä *et al.* 2014). Some of the common terminologies and metrics for multilayer networks can be defined as follows:

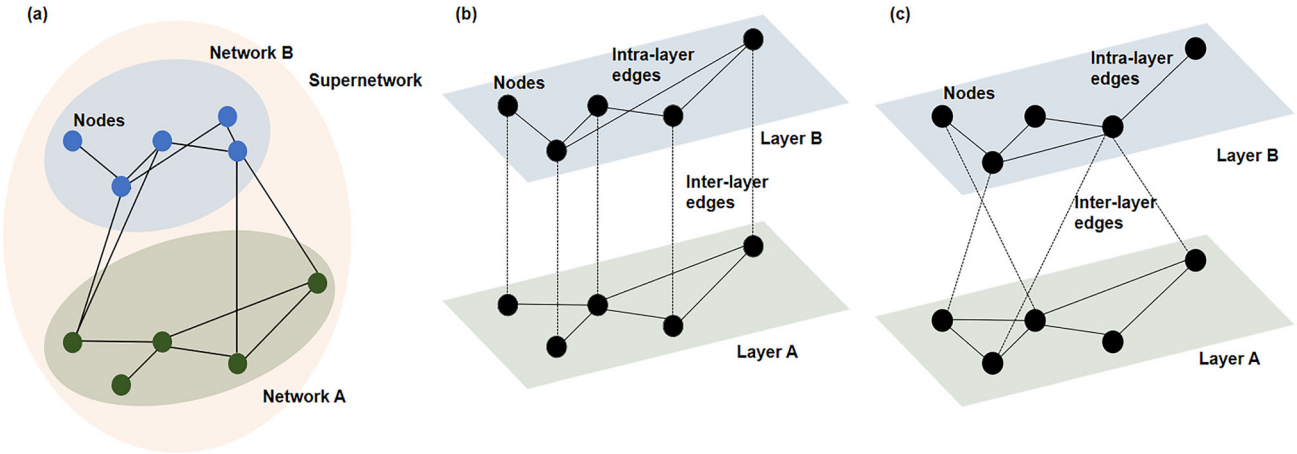


Figure 3. A schematic representation of some common types of multilayer networks studied in biological systems: (**a**) a network of two network systems A and B, (**b**) a multiplex network of two layers A and B, and (**c**) an interconnected network of two layers A and B.

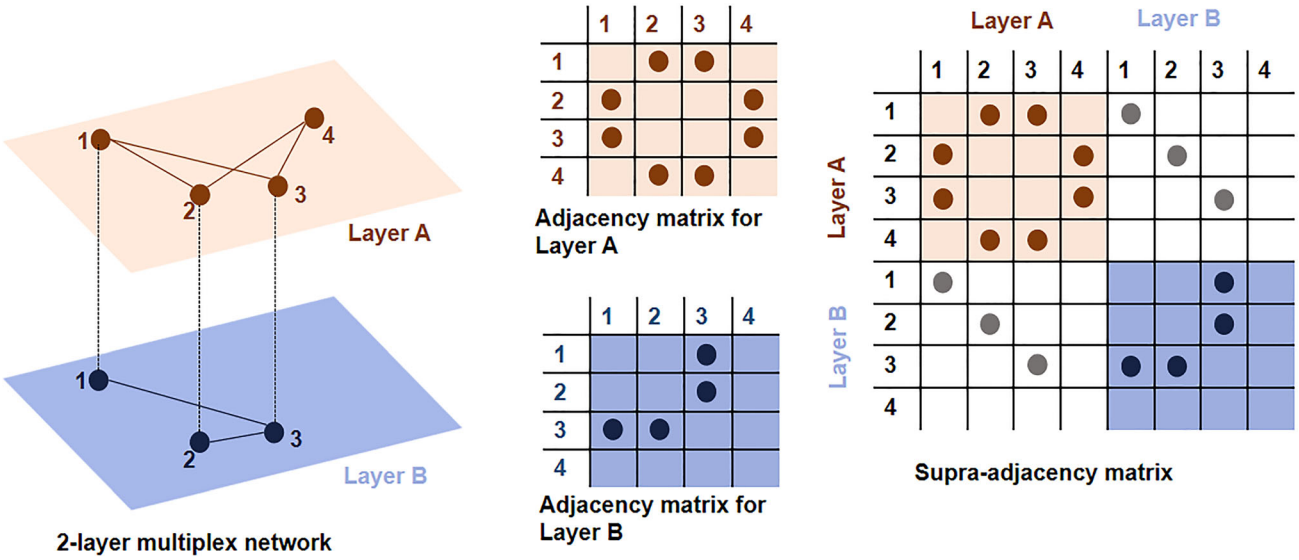


Figure 4. Schematic representation of a supra-adjacency matrix for a two-layer multiplex network.

(i) *Supra-adjacency matrices:* The adjacency-matrix for simple networks shows the pairs of nodes connected by an edge in a square matrix representation. This can be extended to multilayer networks by using supra-adjacency matrices (figure 4). A supra-adjacency matrix is a block-diagonal structure, where the interior block-diagonal structures correspond to intra-layer adjacency matrices.

(ii) *Degree of a node in a layer:* The degree of a node, as in a simple network, can also be defined for a multilayer network by aggregating data from all the layers and then applying the original definition of node to the resulting single network. In a multilayer network comprising nodes and edges in M layers, each layer has its own adjacency matrix $\mathbf{A}$, and the degree of a node can be defined for each layer. The degree of a specific node i on a specific layer m can be formulated as

$$k_{i,[m]} = \sum_j a_{ij[m]}$$

Therefore, the degree of any node i can be written as a vector

$$k_i = \left\{k_{i,[1]}, k_{i,[2]}, \ldots, k_{i,[M]}\right\}$$

(iii) *Transitivity:* The tendency to form cycles involving three nodes is known as *transitivity* or *clustering coefficient.* It indicates how well the neighbours of a node are connected to each other (box 1). Since a multilayer network has several layers in consideration, the definitions of network clustering coefficients can be used to characterize the abundance of triangles on each layer.

(iv) *Centrality measures:* Unlike the single network, the connections between the nodes are of different types based on the layers and the relationships between them. This makes the definition of centrality, or importance of nodes, complicated. However, centrality measures have been defined keeping different types of interactions between layers in consideration. *PageRank,* for simple networks, defines the importance of a node based on the frequency of visits to that node during a random walk on the network. *MultiRank* can be defined on multiplex networks where a random walker can either take steps inside a layer or change layers (Halu *et al.* 2013). In this, the rank of a node in one layer affects the rank in an other layer. Based on the effect on one layer on the other Halu *et al.* defined four different types of PageRanks: Additive, Multiplicative, Combined, and Neutral. Similarly, variations of *eigenvector centrality* have been deduced to determine the importance of nodes in a multiplex network with heterogeneous interaction paradigms between various layers (Sol *et al.* 2013).

The parameters defining multilayer networks have been comprehensively reviewed by Boccaletti *et al.* (2014) and Kivela *et al.* (2014).

3.2 *Dynamic processes on multilayer networks*

Dynamical processes have been extensively studied in single networks (Boccaletti *et al.* 2006). Epidemic spreading in complex networks has been particularly relevant in studying the spread of infectious diseases in a population and to design strategies to control such spreads.

Another widely studied phenomenon is synchronization, which is an emergent property of networks with interacting units (Arenas *et al.* 2008).

Dynamic processes have also been studied in multilayer networks, especially in interdependent networks, a subclass of multilayer networks. The interdependency of these networks is such that if an attack or failure occurs in either, the other dependent node stops functioning, which makes these nodes critical in a network (Kenett *et al.* 2015). In interdependent networks, failure in nodes of one network can cause failure of dependent nodes in other networks, leading to a cascade of failures and ultimately the collapse of the network system (Buldyrev *et al.* 2010). Several studies have described conditions under which this catastrophic collapse of the interdependent networks can be prevented (Reis *et al.* 2014; Radicchi and Bianconi 2017; Min and Zheng 2018). Redundant connections play a crucial role in making such networks robust. Synchronization in multilayer networks has been characterized only recently. A generalized framework for synchronization in multilayer networks has been developed by Del Genio *et al.* (2016) and Rossa *et al.* (2020). Further details can be obtained in comprehensive reviews by Boccaletti *et al.* (2014) and Kivelä *et al.* (2014). Having briefly introduced the theoretical developments for multilayer paradigm of complex system modelling, we will now describe recent advances in the application of this paradigm to biological systems.

4. Multilayer networks in biological systems

4.1 *Multilayer network modelling of transcriptional and post-transcriptional processes*

Multilayer network modelling of biological systems enables integration of information at multiple scales and evinces the emergent properties at systems level. In a recent study, Azevedo *et al.* (2021) modelled the human transcriptome as a multilayer network at the intra- and inter-tissue level. Five different multilayer networks were generated and analyzed, comprising (a) all human tissues with each layer representing a tissue type, (b) brain and gastrointestinal tissues with 16 layers each corresponding to the brain tissues and gastrointestinal tissues, giving insights into the gut–brain axis, (c) brain tissues and whole blood, comprising 14 layers that helped to study brain-derived communities, (d) non-brain tissues comprising 36 layers, and (e) brain tissues comprising 13 layers each corresponding to the various brain regions. Network analysis of such detailed multilayer networks helps to gain insights into how perturbations affect biological pathways and mediate disease processes (Azevedo *et al.* 2021).

A multilayer framework was used to identify functional communities in colon adenocarcinoma by considering a three-layered network comprising transcriptional, post-transcriptional, and physical interaction levels. A seed-centric community detection approach was used to identify hub nodes and genes that showed a different expression pattern in tumour progression and therefore could be potential biomarkers in the detection of colon adenocarcinoma (Pournoor *et al.* 2020). In another application, a multilayer network modelling comprising protein–protein interaction, shared domain information and shared protein complex information as three different layers was used to accurately predict protein function (Zhao *et al.* 2016). Multilayer networks have also been used to model the interdependence of gene regulation and cellular metabolism (figure 5a). In this context, the regulation and metabolism in *Escherichia coli* was studied as a three-layer network comprising the gene regulatory layer, the protein interaction layer, and the metabolic layer. This study investigated the spreading of internal and external perturbations through the three layers, and found that the interdependent network was particularly robust against metabolic perturbations (Klosik *et al.* 2017).

4.2 *Human physiology and diseases as multilayer networks*

In terms of networks, one can imagine the human body as a complex network of vertices representing distinct physiological systems and edges representing various interactions among them. Such reduced descriptions were previously studied by Kivelä *et al.* (2014). Halu *et al.* (2019) modelled 779 human diseases as a multiplex network comprising a genotype-based layer and a phenotype-based layer, and reported multiple communities that bridged the gap between the genotype and phenotype layers, thereby discovering new disease–disease interactions, which would not have been possible in the case of single network analysis (figure 5b). In another study, a weighted four-layer disease–disease similarity network comprising disease network based on protein interactions, disease symptoms, gene ontology, and disease ontology was built and nine conserved disease modules were observed using a tensor-based computational framework. Their results showed that diseases of the same type were mostly grouped together and that this method would be useful in identifying potential disease–disease relationships (Yu *et al.* 2019b). Berenstein *et al.* (2016) used a multilayer network approach for drug repositioning by considering three separate layers of bioactive compounds, target proteins, and their functional associations. This enabled identification of novel drug targets for orphan bioactive compounds.

4.3 *Modelling brain structure and function, a multilayer approach*

In addition to understanding the structure of the nervous system, it is important to understand the relationship between structure and function, where a multilayer representation becomes very useful (Crofts *et al.* 2016). Brain networks vary across time, frequency, subjects, conditions, and connectivity, thus necessitating its study as a multilayer network. Multilayer networks allow us to analyse brain networks comprising both structural and functional layers (figure 5c). Battiston *et al.* (2017) modelled the brain as a multilayer network comprising two layers: (a) the structural or anatomical layer where nodes are usually putative brain regions and the links represent the physical connections among them; and (b) the functional layer where nodes represent an area of the brain usually consisting of neural assemblies, and the links represent functional interactions such as electrical, magnetic, hemo-dynamic, and metabolic activities between two regions. An analysis of motifs and subgraphs in these layers provided crucial insights regarding relationships between the structure and function of the brain (Battiston *et al.* 2017). The structural network of *Caenorhabditis elegans* has been modelled as a four-layer multilayer neural communication network, with each layer representing a different mode of connection between neurons, namely, synaptic, gap junction, monoamine, and neuropeptide coupling (Battiston *et al.* 2014). In a similar way, the connectomics of the *C. elegans* nervous system at the level of individual modes of connection can serve as a prototype to understand more complex nervous systems such as that of the human brain and the behaviourally relevant communication within the brain, and to study the topological properties and dynamics of such networks.

Several mathematical concepts of multilayer networks have been actively used to model changing brain dynamics over time and have proven to be a powerful tool in describing the complex organization and evolution of the human brain and its relationship to cognition. A multilayer network framework can be utilized to integrate information from different methods, where each layer of the network corresponds to a network obtained from different neuroimaging techniques, and various graph metrics can be

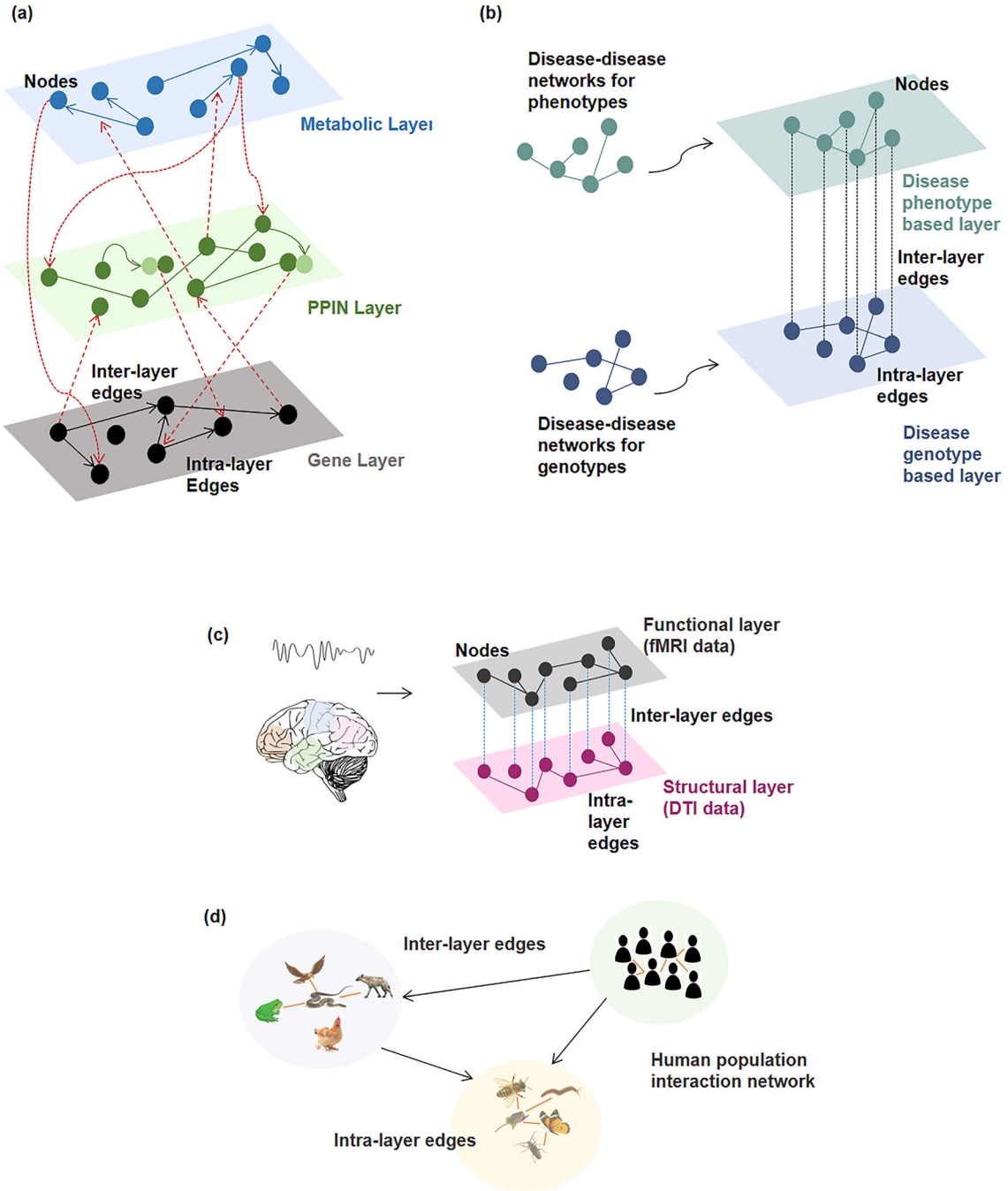


Figure 5. A schematic representation of how some biological networks can be modelled as multilayer network systems. **(a)** An interdependent network of gene regulation and metabolism comprising three layers: gene regulatory layer, protein interaction layer, and metabolic layer (Brazhnik *et al.* 2002). The components (nodes) are organized in three levels: mRNAs, proteins, and metabolites. The dashed red directed edges represent the inter-layer connections such as activation, repression, and catalysis. The solid lines represent intra-layer connections such as gene regulation in the 'gene layer' (shown in grey), protein–protein interaction in the PPIN layer (shown in green), and metabolite conversion in the metabolic layer (shown in blue). **(b)** Disease network represented as two-layer network comprising a phenotype layer and a genotype layer. The phenotype layer (shown in teal) comprises diseases connected by symptoms, and the genotype layer (shown in dark blue) comprises diseases connected by common genes. The inter-layer edges (shown in dashed lines) represent connections between diseases. **(c)** A multiplex network of the human brain, comprising two layers: the functional and the structural layer, derived from fMRI and DTI data, respectively. **(d)** A network of networks representation of an ecological network. Intra-layer edges (shown as orange lines) denote trophic interactions, and inter-layer edges (shown as black arrows) represent species dispersal between communities.

applied (Mandke *et al.* 2018). Community detection in the case of brain networks can be challenging, given that brain networks are usually time-varying. In one such work that discusses various mathematical techniques to quantify brain networks, Puxeddu *et al.* (2019) analysed the performances of community detection algorithms as employed for brain networks, thereby providing a guide to choosing the optimal temporal resolution according to network features. Most of our current understanding about functional brain networks assumes the brain to be a static entity; however, in reality, optimal brain function involves frequent switching of brain networks. Pedersen *et al.* (2018) showed that brain regions transit between different network configurations at a high rate, which is a predictor of the performance of cognitive functions including working memory, planning, and reasoning.

Multilayer analysis of data from healthy and diseased patients have recently been employed to study a variety of disorders. For instance, the fMRI data representing functional networks with the nodes as brain regions and the edges as coherence between regional activity, and diffusion tensor imaging (DTI) data representing the structural network, can be used to build a multilayer network comprising two layers (Vaiana and Muldoon 2020). Furthermore, several kinds of neural data arising from fMRI, magnetoencephalography (MEG), and electroencephalography (EEG) can be decomposed into frequency bands for each brain region and then studied in layers as multilayer frequency-based brain networks (Buldú and Porter 2018). This provides insights into studying disease patterns such as schizophrenia, their overall detailed brain connectomics, and how disruptions in both the structural and functional network layers lead to diseases (Heuvel and Fornito 2014).

4.4 *Cancer complexome as a multilayer network*

The cancer complexome comprises components that belong to various organizational levels, such as components involved in cytology, physiology, signalling mechanisms, and response. The entire complexome of cancer in seven different human cancers, namely, breast, oral, ovarian, cervical, lung, colon, and prostate, was studied in order to understand cancer development, progression and treatment response by using a combined framework of spectral graph theory, and network theory along with a multilayer analysis (Rai *et al.* 2017). The multilayer framework highlights the proteins that are common in all cancers and have structural importance in individual networks. In the context of cancer therapy, it can help elucidate the genetic mechanisms that are altered, thereby causing disease.

4.5 *Understanding the dynamics of infection transmission using multilayer network approach*

The transmission of pathogens in nature cannot be accurately quantified via only a single mode of transmission. Applying the techniques of complex network theory to epidemiology gives us valuable insights into understanding pathogen transmission dynamics in populations with multiple transmission modes. In this context, several studies have used multilayer networks to study pathogen transmission through multiple modes and pathways of infection (Tsai *et al.* 2011; Mari *et al.* 2012; Buono and Braunstein 2015). A multilayer network approach that assumes each transmission mode as a layer can help us to understand how a novel pathogen spreads through a population and in what ways it affects the dynamics of the population. Silk *et al.* (2019) defined three possible ways of modelling disease transmission using multilayer networks by considering habitat connectivity, social networks, transmission of disease to multiple hosts via the environment and finally co-infection. In another study, Silk *et al.* (2018a) used a multilayer network to model the transmission of infection between European badgers and domestic cattle. Using the multilayer approach enabled them to understand the role of direct and indirect transmission of infection between different hosts. Using a multiplex network approach with two layers, namely, ecological and transmission layers, Stella *et al.* (2017) studied the spread of *Trypanosoma cruzi* in its natural habitat across different mammalian species through different modes of transmission. Multilayer networks have also been used to model the simultaneous propagation of two coexisting pathogens through a network. It helps to understand spreading dynamics, for instance, if the infection of a node by one pathogen alters the susceptibility to the second pathogen, or if co-infection of a node influences its ability to transmit either pathogen (Azimi-Tafreshi 2016).

4.6 *Modelling ecological interactions as multilayer networks*

Although ecological networks have been extensively studied as discrete networks, with a multilayer approach to model ecological networks, the layers

can be defined by interaction type, group identity, and levels of organization (Pilosof *et al.* 2017). An example of networks layered according to levels of organizations is a three-level network comprising population, community, and meta-community, studied and analysed by Scotti *et al.* (2013) It was shown that changes at the population level could impact the meta-community at the top level, thereby illustrating the role of hierarchical structural mechanisms in ecological networks (figure 5d). A multilayer network has also been used to link animal social behaviour with their respective physical environment. Habitats and resources influence animal movement patterns, which in turn influence multiple types of social interactions such as predation, competition, mating choice, and their time-dependent dynamics. Integrating such spatial, temporal, and social dynamics using multilayer networks can uncover novel ecological impacts on animal social behaviour (Silk *et al.* 2018b).

5. Future directions and challenges

The use of multilayer networks to study complex systems is a promising direction, especially in a biological context. At the cellular level, proteins and genes are often involved in many processes such as signalling, and transcriptional or metabolic regulations, depending on various aspects such as the external environment and type of tissue that necessitates the use of a multilayer framework to describe them. At the population level, a domain that could greatly benefit from a multilayer framework is that of disease spreading, especially in cases where two or more diseases propagate together in the host population, giving rise to complex dynamical interdependencies like cross-immunization or cooperation between the diseases, as is the case with tuberculosis and HIV. These can be accounted for by considering multiple networks of contacts. Multilayer networks, thus, have provided a promising direction across all domains. The application of multilayer network models to real-world networks has been instrumental in understanding the robustness (or lack thereof) of these complex systems. Considering the robustness inherent to the biological systems to be a result of evolution, it would be interesting to understand the molecular underpinnings of this robustness in biological systems by considering various layers of interactions and cross-talks. This also provides an opportunity to understand complex diseases such cancer and metabolic disorders in the context of various interacting networks and to identify novel targets. Another area where the multilayer network formalism can provide useful insights is in understanding brain networks by modelling various regions and their connections with respect to their functions and comparing them in various diseased states.

Although the multilayer paradigm provides several advantages in terms of understanding the complexities inherent in real-world complex systems, it comes with its own challenges. One of the most significant challenges is understanding or delineating the layers to be modelled within the system. This becomes even more relevant when redundant information could be present in different layers (Hasenjager *et al.* 2021). Another challenge is dealing with the scarcity of data and formalism to model real systems as multilayer networks. Identifying and assigning the intra- and inter-layer edges is another crucial task that presents challenges in multilayer networks. Furthermore, calculating the appropriate network parameters that can help in understanding the system is another challenge in modelling complex systems as multilayer networks. However, with advances being made, both, in terms of the development of theory as well as in the availability of data at various levels (subcellular, cellular, and organism), modelling biological systems as multilayer networks will provide deeper insights into their mechanisms.

Acknowledgements

SC thanks DST-INSPIRE for the fellowship. AS thanks DBT for the Ramalingaswami Fellowship and IIT Gandhinagar for support.

References

Aftabuddin M and Kundu S 2010 AMINONET – a tool to construct and visualize amino acid networks, and to calculate topological parameters. *J. Appl. Crystallogr.* **43** 367–369

Albert R 2005 Scale-free networks in cell biology. *J. Cell Sci.* **118** 4947–4957

Amitai G, Shemesh A, Sitbon E, *et al.* 2004 Network analysis of protein structures identifies functional residues. *J. Mol. Biol.* **344** 1135–1146

Arenas A, Díaz-Guilera A, Kurths J, Moreno Y and Zhou C 2008 Synchronization in complex networks. *Phys. Rep.* **469** 93–153

Atilgan AR, Akan P and Baysal C 2004 Small-world communication of residues and significance for protein dynamics. *Biophys. J.* **86** 85–91

Azeloglu EU and Iyengar R 2015 Signalling networks: information flow, computation, and decision making. *Cold Spring Harb. Perspect. Biol.* **7** a005934

Azevedo T, Dimitri GM, Lió P and Gamazon ER 2021 Multilayer modelling of the human transcriptome and biological mechanisms of complex diseases and traits. *NPJ Syst. Biol. Appl.* **7** 24

Azimi-Tafreshi N 2016 Cooperative epidemics on multiplex networks. *Phys. Rev. E*. https://doi.org/10.1103/PhysRevE.93.042303

Bagler G and Sinha S 2005 Network properties of protein structures. *Phys. A Stat. Mech. Appl.* **346** 27–33

Bagler G and Sinha S 2007 Assortative mixing in protein contact networks and protein folding kinetics. *Bioinformatics* **23** 1760–1767

Barabási AL and Albert R 1999 Emergence of scaling in random networks. *Science* **286** 509–512

Barabási A-L, Gulbahce N and Loscalzo J 2011 Network medicine: a network-based approach to human disease. *Nat. Rev. Genet.* **12** 56–68

Bascompte J 2010 Structure and dynamics of ecological networks. *Science* **329** 765–766

Bassett DS and Bullmore E 2006 Small-world brain networks. *Neuroscientist* **12** 512–523

Battiston F, Nicosia V, Chavez M and Latora V 2017 Multilayer motif analysis of brain networks. *Chaos* **27** 1–20

Battiston F, Nicosia V and Latora V 2014 Structural measures for multiplex networks. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys*. **89** 032804

Berenstein AJ, Magariños MP, Chernomoretz A and Agüero F 2016 A multilayer network approach for guiding drug repositioning in neglected diseases. *PLoS Negl. Trop. Dis.* **10** e0004300

Bhalla US and Iyengar R 1999 Emergent properties of networks of biological signalling pathways. *Science* **283** 7

Boccaletti S, Latora V, Morenod Y, Chavez M and Hwang DU 2006 Complex networks: Structure and dynamics. *Phys. Rep.* **424** 175–308

Boccaletti S, Bianconi G, Criado R, *et al.* 2014 The structure and dynamics of multilayer networks. *Phys. Rep.* **544** 1–122

Brazhnik P, de la Fuente A and Mendes P 2002 Gene networks: how to put the function in genomics. *Trends Biotechnol.* **20** 467–472

Brinda KV and Vishveshwara S 2005 A network representation of protein structures: implications for protein stability. *Biophys. J.* **89** 4159–4170

Buldú JM and Porter MA 2018 Frequency-based brain networks: From a multiplex framework to a full multilayer description. *Netw. Neurosci.* **2** 418–441

Buldyrev SV, Parshani R, Paul G, Stanley HE and Havlin S 2010 Catastrophic cascade of failures in interdependent networks. *Nature* **464** 1025–1028

Buono C and Braunstein LA 2015 Immunization strategy for epidemic spreading on multilayer networks. *Europhys. Lett.* **109** 16–20

Chakrabarty B and Parekh N 2014 PRIGSA: Protein repeat identification by graph spectral analysis. *J. Bioinform. Comput. Biol.* **12** 1442009

Chakrabarty B and Parekh N 2016 NAPS: network analysis of protein structures. *Nucleic Acids Res.* **44** W375–W382

Chakrabarty B and Parekh N 2020 PRIGSA2: Improved version of protein repeat identification by graph spectral analysis. *J. Biosci.* **45** 95

Crofts JJ, Forrester M and O'Dea RD 2016 Structure-function clustering in multiplex brain networks. *EPL* **116** 1–7

De Domenico M, Porter MA and Arenas A 2015 MuxViz: a tool for multilayer analysis and visualization of networks. *J. Complex Netw.* **3** 159–176

Del Genio CI, Gomez-Gardenes J, Bonamassa I and Boccaletti S 2016 Synchronization in networks with multiple interaction layers. *Sci. Adv.* **2** 11

Della Rossa F, Pecora L, Blaha K, *et al.* 2020 Symmetries and cluster synchronization in multilayer networks. *Nat. Commun.* **11** 3179

Di Paola L and Giuliani A 2015 Protein contact network topology: a natural language for allostery. *Curr. Opin. Struct. Biol.* **31** 43–48

Di Paola L, De Ruvo M, Paci P, Santoni D and Giuliani A 2013 Protein contact networks: an emerging paradigm in chemistry. *Chem. Rev.* **113** 1598–1613

Dittrich MT, Klau GW, Rosenwald A, Dandekar T and Müller T 2008 Identifying functional modules in protein–protein interaction networks: an integrated exact approach. *Bioinformatics* **24** 223–231

Dokholyan NV, Li L, Ding F and Shakhnovich EI 2002 Topological determinants of protein folding. *Proc. Natl. Acad. Sci. USA* **99** 8637–8641

Doncheva NT, Klein K, Domingues FS and Albrecht M 2011 Analyzing and visualizing residue networks of protein structures. *Trends Biochem. Sci.* **36** 179–182

Duan D, Lv C, Si S, et al. 2019 Universal behavior of cascading failures in interdependent networks. *Proc. Natl. Acad. Sci. USA* **116** 22452–22457

Emmert-Streib F, Dehmer M and Haibe-Kains B 2014 Gene regulatory networks and their applications: understanding biological and medical problems in terms of networks. *Front. Cell Dev. Biol.* **2** 38

Erdös P and Rényi A 1959 On random graphs. *Publicationes Mathematicae* **6** 290–297

Fang X, Lloyd CJ and Palsson BO 2020 Reconstructing organisms in silico: genome-scale models and their emerging applications. *Nat. Rev. Microbiol.* **18** 731–743

Friston KJ 1994 Functional and effective connectivity in neuroimaging: A synthesis. *Hum. Brain Mapp.* **2** 56–78

Gandhi TK, Zhong J and Mathivanan S 2006 Analysis of the human protein interactome and comparison with yeast,

worm and fly interaction datasets. *Nat. Genet.* **38** 285–293

Gao J, Li D and Havlin S 2014 From a single network to a network of networks. *Natl. Sci. Rev.* **1** 346–356

Ghosh A, Brinda KV and Vishveshwara S 2007 Dynamics of lysozyme structure network: probing the process of unfolding. *Biophys. J.* **92** 2523–2535

Ghosh A, Sakaguchi R, Liu C, Vishveshwara S and Hou YM 2011 Allosteric communication in cysteinyl tRNA synthetase: a network of direct and indirect readout. *J. Biol. Chem.* **286** 37721–37731

Gilbert EN 1959 Random graphs. *Ann. Math. Stat.* **30** 1141–1144

Greene LH and Higman VA 2003 Uncovering network systems within protein structures. *J. Mol. Biol.* **334** 781–791

Guebila MB, Lopes-Ramos CM, Weighill D, *et al.* 2022 GRAND: a database of gene regulatory network models across human conditions. *Nucleic Acids Res.* **50** D610–D621

Halu A, Mondragon RJ, Panzarasa P and Bianconi G 2013 Multiplex PageRank. *PLoS One* **8** e78293

Halu A, De Domenico M, Arenas A and Sharma A 2019 The multiplex network of human diseases. *NPJ Syst. Biol. Appl.* **5** 15

Hao Shi, Yan KK, Ding L, *et al.* 2020 Network approaches for dissecting the immune system. *iScience* **23** 101354

Hasenjager MJ, Silk M and Fisher DN 2021 Multilayer network analysis: new opportunities and challenges for studying animal social systems. *Curr. Zool.* **67** 45–48

Heinken A and Thiele I 2015 Systems biology of host-microbe metabolomics. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **7** 195–219

Herrgård MJ, Swainston N, Dobson P, *et al.* 2008 A consensus yeast metabolic network reconstruction obtained from a community approach to systems biology. *Nat. Biotech.* **26** 1155–1160

Hilgetag CC, Burns GAPC, O'Neill MA, Scannell JW and Young MP 2000 Anatomical connectivity defines the organization of clusters of cortical areas in the macaque monkey and the cat. *Philos. Trans. R. Soc. B Biol. Sci.* **355** 91–110

Jaeger S, Sers CT and Leser U 2010 Combining modularity, conservation, and interactions of proteins significantly increases precision and coverage of protein function prediction. *BMC Genomics* **11** 717

Jeong H, Mason SP, Barabási AL and Oltvai ZN 2001 Lethality and centrality in protein networks. *Nature* **411** 41–42

Kann MG 2007 Protein interactions and disease: computational approaches to uncover the etiology of diseases. *Brief. Bioinform.* **8** 333–346

Kenett DY, Perc M and Boccaletti S 2015 Networks of networks–an introduction. *Chaos Solitons Fractals* https://doi.org/10.1016/j.chaos.2015.03.016

Keshava Prasad TS, Goel R, Kandasamy K, *et al.* 2009 Human Protein Reference Database—2009 update. *Nucleic Acids Res.* **37** D767–D772

Kim EY, Ashlock D and Yoon SH 2019 Identification of critical connectors in the directed reaction-centric graphs of microbial metabolic networks. *BMC Bioinform.* **20** 328

Kivelä M, Arenas A, Barthelemy M, *et al.* 2014 Multilayer networks. *J. Complex Net.* **2** 203–271

Klosik DF, Grimbs A, Bornholdt S and Hütt MT 2017 The interdependent network of gene regulation and metabolism is robust where it needs to be. *Nat. Commun.* **8** 534

Lu X, Gray C, Brown L, *et al.* 2016 Drought rewires the cores of food webs. *Nat. Clim. Change* **6** 875–878

Luck K, Kim DK, Lambourne L, *et al.* 2020 A reference map of the human binary protein interactome. *Nature* **580** 402–408

Ma H, Sorokin A, Mazein A, *et al.* 2007 The Edinburgh human metabolic network reconstruction and its functional analysis. *Mol. Sys. Biol.* **3** 135

Mandke K, Meier J, Brookes MJ, *et al.* 2018 Comparing multilayer brain networks between groups: Introducing graph metrics and recommendations. *Neuroimage* **166** 371–384

Mari L, Bertuzzo E, Righetto L, *et al.* 2012 Modelling cholera epidemics: the role of waterways, human mobility and sanitation. *J. R. Soc. Interface* **9** 376–388

Mercatelli D, Scalambra L, Triboli L, Ray F and Giorgia FM 2020 Gene regulatory network inference resources: A practical overview. *Biochim. Biophys. Acta Gene Regul. Mech.* **1863** 194430

Min B and Zheng M 2018 Correlated network of networks enhances robustness against catastrophic failures. *PLoS One* **13** e0195539

Montoya JM, Pimm SL and Solé RV 2006 Ecological networks and their fragility. *Nature* **442** 259–264

Muller EEL, Faust K, Widder S, *et al.* 2018 Using metabolic networks to resolve ecological properties of microbiomes. *Curr. Opin. Syst. Biol.* **8** 73–80

Newman MEJ, Strogatz SH and Watts DJ 2001 Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* **64** 026118

Orchard S, Ammari M, Aranda B, *et al.* 2013 The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* **42** D358–D363

Orth JD, Conrad TM, Na J, *et al.* 2011 A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. *Mol. Sys. Biol.* **7** 535

Ouma WZ, Pogacar K and Grotewold E 2018 Topological and statistical analyses of gene regulatory networks reveal unifying yet quantitatively different emergent properties. *PLoS Comput. Biol.* **14** e1006098

Pascual-García A and Bastolla U 2017 Mutualism supports biodiversity when the direct competition is weak. *Nat. Commun.* **8** 14326

Pedersen M, Zalesky A, Omidvarnia A and Jackson GD 2018 Multilayer network switching rate predicts brain performance. *Proc. Natl. Acad. Sci. USA* **115** 13376–13381

Pilosof S, Porter MA, Pascual M and Kéfi S 2017 The multilayer nature of ecological networks. *Nat. Ecol. Evol.* **1** 0101

Pournoor E, Mousavian Z, Dalini AN and Masoudi-Nejad A 2020 Identification of key components in colon adenocarcinoma using transcriptome to interactome multilayer framework. *Sci. Rep*. **10** 4991

Puxeddu MG, Petti M, Mattia D and Astolfi L 2019 The Optimal setting for multilayer modularity optimization in multilayer brain networks. *41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* pp. 624-627

Radicchi F and Bianconi G 2017 Redundant interdependencies boost the robustness of multiplex networks. *Phys. Rev. X* **7** 011013

Rai A, Pradhan P, Nagraj J, *et al.* 2017 Understanding cancer complexome using networks, spectral graph theory and multilayer framework. *Sci. Rep.* **7** 41676

Reis SDS, Hu Y, Babino A, *et al.* 2014 Avoiding catastrophic failure in correlated networks of networks. *Nat. Phys.* **10** 762–767

Roume H, Heintz-Buschart A, Muller E, *et al.* 2015 Comparative integrated omics: identification of key functionalities in microbial community-wide metabolic networks. *NPJ Biofilms Microbiomes* **1** 15007

Rinaldi SM, Peerenboom JP and Kelly TK 2001 Identifying, understanding, and analyzing critical infrastructure interdependencies. *IEEE Control Syst. Mag.* **21** 11–25

Sathyapriya R and Vishveshwara S 2007 Structure networks of *E. coli* glutaminyl-tRNA synthetase: Effects of ligand binding. *Proteins Struct. Funct. Bioinform.* **68** 541–550

Sathyapriya R, Vijayabaskar MS and Vishveshwara S 2008 Insights into protein–DNA interactions through structure network analysis. *PLoS Comput. Biol.* **4** e1000170

Scotti M, Ciocchetta F and Jordan F 2013 Social and landscape effects on food webs: a multi-level network simulation model. *J. Complex Networks* **1** 160–182

Sharan R, Suthram S, Kelley RM, *et al.* 2005 Conserved patterns of protein interaction in multiple species. *Proc. Natl. Acad. Sci. USA* **102** 1974–1979

Sharan R, Ulitsky I and Shamir R 2007 Network-based prediction of protein function. *Mol. Syst. Biol.* **3** 88

Silk MJ, Drewe JA, Delahay RJ, *et al.* 2018a Quantifying direct and indirect contacts for the potential transmission of infection between species using a multilayer contact network. *Behaviour* **155** 731–757

Silk MJ, Finn KR, Porter MA and Pinter-Wollman N 2018b Can multilayer networks advance animal behavior research? *Trends Ecol Evol.* **33** 376–378

Silk MJ, Hodgson DJ, Rozins C, *et al.* 2019 Integrating social behaviour, demography and disease dynamics in network models: applications to disease management in declining wildlife populations. *Philos. Trans. R. Soc. B Biol. Sci*. **374** 20180211

Sol AD and O'Meara P 2005 Small-world network approach to identify key residues in protein–protein interaction. *Proteins* **58** 672–682

Sol AD, Fujihashi H, Amoros D and Nussinov R 2006 Residue centrality, functionally important residues, and active site shape: Analysis of enzyme and non-enzyme families. *Protein Sci.* **15** 2120–2128

Sol L, Romance M, Criado R, *et al.* 2013 Eigenvector centrality of nodes in multiplex networks. *Chaos* **23** 033131

Solomonoff R and Rapoport A 1951 Connectivity of random nets. *Bull. Math. Biophys.* **13** 107–117

Song HS, Nelson WC, Lee JY, *et al.* 2017 Metabolic network modelling for computer-aided design of microbial interactions; in HN Chang (ed) *Emerging Areas in Bioengineering* (Wiley) pp 793–801

Spirin V and Mirny LA 2003 Protein complexes and functional modules in molecular networks. *Proc. Natl. Acad. Sci. USA* **100** 12123–12128

Sporns O and Zwi JD 2004 The small world of the cerebral cortex. *Neuroinformatics* **2** 145–162

Sporns O, Chialvo DR, Kaiser M and Hilgetag CC 2004 Organization, development and function of complex brain networks. *Trends Cogn. Sci.* **8** 418–425

Sporns O 2003 Graph theory methods for the analysis of neural connectivity patterns; in R Kötter (ed) *Neuroscience databases* (Springer, Boston, MA) pp 171–185

Srivastava A and Sinha S 2017 Uncoupling of an ammonia channel as a mechanism of allosteric inhibition in anthranilate synthase of *Serratia marcescens*: dynamic and graph theoretical analysis. *Mol. Biosyst.* **13** 142–155

Srivastava A, Birari V and Sinha S 2020 Small conformational changes underlie evolution of resistance to NNRTI in HIV reverse transcriptase. *Biophys. J.* **118** 2489–2501

Stark C, Breitkreutz BJ, Reguly T, *et al.* 2006 BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* **34** D535–D539

Statello L, Guo CJ, Chen LL and Huarte M 2021 Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell. Biol.* **22** 96–118

Stella M, Andreazzi CS, Selakovic S, Goudarzi A and Antonioni A 2017 Parasite spreading in spatial ecological multiplex networks. *J. Complex Netw.* **5** 486–511

Stelzl U, Worm U, Lalowski M, *et al.* 2005 A human protein–protein interaction network: a resource for annotating the proteome. *Cell* **122** 957–968

Strogatz SH 2001 Exploring complex networks. *Nature* **410** 268–276

Supper J, Spangenberg L, Planatscher H, *et al.* 2009 BowTieBuilder: modelling signal transduction pathways. *BMC Syst. Biol.* **3** 67

Szklarczyk D, Gable AL, Nastou KC, *et al.* 2021 The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **49** D605–D612

Telesford QK, Joyce KE, Hayasaka S, Burdette JH and Laurienti PJ 2011 The ubiquity of small-world networks. *Brain Connect.* **1** 367–375

Tsai YS, Huang CY, Wen TH and Yen MY 2011 Integrating epidemic dynamics with daily commuting networks: building a multilayer framework to assess influenza A (H1N1) intervention policies. *Simulation* **87** 385–405

Vaiana M and Muldoon SF 2020 Multilayer brain networks. *J. Nonlinear Sci.* **30** 2147–2169

van den Heuvel MP and Fornito A 2014 Brain networks in schizophrenia. *Neuropsychol. Rev.* **24** 32–48

Vendruscolo M, Dokholyan NV, Paci E and Karplus M 2002 Small-world view of the amino acids that play a key role in protein folding. *Phys. Rev. E* **65** 061910

Vijayabaskar MS and Vishveshwara S 2010 Interaction energy based protein structure networks. *Biophys. J.* **99** 3704–3715

Vijayabaskar MS and Vishveshwara S 2012 Insights into the fold organization of TIM barrel from interaction energy based structure networks. *PLoS Comput. Biol.* **8** e1002505

Wagner A and Fell DA 2001 The small world inside large metabolic networks. *Proc. R. Soc. B Biol. Sci.* **268** 1803–1810

Wang K, Hu F, Xu K, *et al*. 2011 CASCADE_SCAN: mining signal transduction network from high-throughput data based on steepest descent method. *BMC Bioinform.* **12** 164

Watts DJ and Strogatz SH 1998 Collective dynamics of 'small-world' networks. *Nature* **393** 440–442

Yalamanchili HK and Parekh N 2009 Graph spectral approach for identifying protein domains. *Lect. Notes Comp. Sci.* **5462** 437–448

Yook SH, Oltvai ZN and Barabási AL 2004 Functional and topological characterization of protein interaction networks. *Proteomics* **4** 928–942

Young MP 1992 Objective analysis of the topological organization of the primate cortical visual system. *Nature* **359** 167–169

Yu J, Peng J and Chi H 2019a Systems immunology: Integrating multi-omics data to infer regulatory networks and hidden drivers of immunity. *Curr. Opin. Syst. Biol.* **15** 19–29

Yu L, Yao S, Gao L and Zha Y 2019b Conserved disease modules extracted from multilayer heterogeneous disease and gene networks for understanding disease mechanisms and predicting disease treatments. *Front. Genet.* **9** 745

Zhao XM, Wang RS, Chen L and Aihara K 2008 Uncovering signal transduction networks from high-throughput data by integer linear programming. *Nucleic Acids Res.* **36** e48

Zhao B, Hu S, Li X, *et al*. 2016 An efficient method for protein function annotation based on multilayer protein networks. *Hum. Genom.* **10** 1–15

Zhou M, Li Q and Wang R 2016 Current experimental methods for characterizing protein–protein interactions. *ChemMedChem* **11** 738–756

Corresponding editor: Susmita Roy