Research

Author for correspondence:
Niall Rodgers
e-mail: NXR081@student.bham.ac.uk

Influence and influenceability: global directionality in directed complex networks

Niall Rodgers[1,3], Peter Tiňo[2] and Samuel Johnson[1,4]

[1]School of Mathematics, University of Birmingham, Birmingham, UK
[2]School of Computer Science, University of Birmingham, Birmingham, UK
[3]Topological Design Centre for Doctoral Training, University of Birmingham, Birmingham, UK
[4]The Alan Turing Institute, The British Library, London, UK

NR, 0000-0002-3936-2709; PT, 0000-0003-2330-128X; SJ, 0000-0002-8648-1735

Knowing which nodes are influential in a complex network and whether the network can be influenced by a small subset of nodes is a key part of network analysis. However, many traditional measures of importance focus on node level information without considering the global network architecture. We use the method of trophic analysis to study directed networks and show that both 'influence' and 'influenceability' in directed networks depend on the hierarchical structure and the global directionality, as measured by the trophic levels and trophic coherence, respectively. We show that in directed networks trophic hierarchy can explain: the nodes that can reach the most others; where the eigenvector centrality localizes; which nodes shape the behaviour in opinion or oscillator dynamics; and which strategies will be successful in generalized rock–paper–scissors games. We show, moreover, that these phenomena are mediated by the global directionality. We also highlight other structural properties of real networks related to influenceability, such as the pseudospectra, which depend on trophic coherence. These results apply to any directed network and the principles highlighted—that node hierarchy is essential for understanding network influence, mediated by global directionality—are applicable to many real-world dynamics.

1. Introduction

Influence in directed complex networks and the ability of the networks to be influenced is vitally important in many real-world systems, for example the spreading of opinions and epidemics or the synchronization of the brain in a seizure [1]. However, when we think of what makes nodes influential, the answer is very

disparate and depends on many factors. This question has been of great interest to the network science community and well studied in the case of undirected networks [2,3]. However, in the directed case insight can be gained by considering properties unique to directed networks. This is necessary as many real-world systems are intrinsically directed and the influence of nodes is strongly tied to the directionality of the edges [4]. Influence can be thought of as how well nodes control the dynamics of a network, how measures of centrality are distributed across the network, how many nodes can be reached from a set of nodes and how sensitive the network is to perturbations. We show that all of these properties can be understood and made intuitive by considering the hierarchical ordering and global directionality of the network as measured through the technique of trophic analysis [5–7].

When the nodes are easily ordered in a hierarchical fashion, as in a food web, it is clear that the network can be influenced by the nodes at the bottom of this hierarchy and conversely it is very difficult to influence the network from the top of the hierarchy. When there is little hierarchical structure, like in an Erdős–Rényi random graph, this effect is damped and influence over the network is more evenly distributed. This simple construction, detailed in the Background, aides in the interpretation of control over network dynamics, spreading processes, localization of centrality measures and sensitivity to perturbations. Hierarchical ordering can explain why nodes which locally look unimportant may be able to influence the entire network. This framework may provide an interpretation to observations made about influence in directed networks in a variety of literature settings such as the effect of heterogeneous centrality and directionality in opinion formation in real networks [8], the influence of peripheral nodes on dynamics of directed networks [9] or the asymmetry between paths up and down the network hierarchy [10]. Trophic analysis pairs a global measure of directionality with a local measure of hierarchy, making it different from a single centrality measure. This allows an intuition surrounding the variability in the importance of a node's position in the hierarchy and how this affects its role in the network. This differs from previous results on the relationship between hierarchy and influence which do not feature this pairing of local and global measures [11,12]. The fact that centrality measures play different roles dependent on the mesoscale structure of the network has been noted previously in the case of PageRank and clustering [13].

This paper is organized as follows. We first introduce the background and explain the principles underlying trophic analysis. We then introduce some dynamics and highlight how global directionality can affect the influence and influenceability of these processes. These are majority vote, Kuramoto oscillators, the voter model and the frequency of strategies in generalized rock–paper–scissors games. We then include some results on the relationship between structure and influence in real-world networks and how this can be shaped by hierarchy. We demonstrate the relationship between hierarchy and eigenvector localization, left and right eigenvector correlation (with specific real-world examples), sensitivity to structural perturbation through the pseudospectra and the size of the out-component of a node. The real network data used in this paper, also used in [7,14], are available at [15] and contain food webs, neural networks, social networks and more as well as the original sources of the network data used. This shows how many diverse notions of influence can be investigated in directed networks using trophic analysis and how in a wide range of systems the ability to exploit hierarchy to influence a system depends on the global directional organization.

2. Background

Real-world systems formed by many interacting elements can be represented using graphs. These complex networks are sets of nodes or vertices representing the elements of the system while the edges or links represent the interactions or connections between elements. Many real-world systems such as social networks, food webs, the Internet and more have interactions which are intrinsically directional and may represent the influence a node has over another [16]. This structure can be represented through a matrix, A, known as an adjacency matrix where the edges are represented by the non-zero entries of the matrix. For an unweighted directed network of N nodes this $N \times N$ matrix is defined as

$$A_{ij} = \begin{cases} 1 & \text{if there exists an edge } i \to j \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

In the case of directed networks this matrix is not symmetric, unlike the undirected case where edges go in both directions and the matrix is symmetric. Each node i then has an in-degree, $k_i^{\text{in}} = \sum_j A_{ji}$, and an out-degree, $k_i^{\text{out}} = \sum_j A_{ij}$. The matrix A can also be weighted to show the strength of interactions but here we focus on the simple unweighted case although it is possible to use trophic analysis in the weighted regime [5].

2.1. Trophic analysis

Trophic analysis is a technique to quantify the global directionality inherent in real directed networks [5], and is applicable to any directed network. Trophic analysis was originally derived from ecology [17] where the original definition linked hierarchy to weighted steps from the basal nodes (vertices of in-degree zero), which is an intuitive way to view hierarchy but cannot be generalized to any directed network without basal nodes like the definition used here and in [5–7,18]. Much of the previous work which applied trophic level and incoherence used the previous definition [14,17,19–21]. This was successfully applied in a wide variety of settings including infrastructure [22,23], the structure of food webs [20], spreading processes such as epidemics or neurons firing [21] and organizational structuring [19].

Trophic analysis combines two parts: the node-level local information, trophic level, and a measure of the global network directionality, trophic incoherence. Trophic level is a node-level quantity which measures where a specific node sits in the network hierarchy. Trophic level is calculated by solving the $N \times N$ matrix equation, first proposed in [5], given by

$$\Lambda h = v, \tag{2.2}$$

where h is the vector of trophic levels for each node and v is the vector imbalance of the in- and out-degree of each node, where each element is defined as $v_i = k_i^{\text{in}} - k_i^{\text{out}}$. Λ is the Laplacian matrix:

$$\Lambda = \text{diag}(u) - A - A^T, \tag{2.3}$$

where u is the sum of the in- and out-degrees of each node, $u_i = k_i^{\text{in}} + k_i^{\text{out}}$, and A^T is the transpose of the adjacency matrix, A. The solution of equation (2.2), which provides the trophic levels of the nodes, is only defined up to a constant vector so we take the convention that the lowest level node is set to trophic level zero. The Laplacian matrix is singular by definition, yet a solution can be found either by choosing a node (say $i = 1$) and setting its value (e.g. $h_1 = 0$); or iteratively, which is convenient for very large networks [5].

In a balanced network, for instance a directed cycle, there is no hierarchical structure so every node has the same level. This is due to the dependence of equation (2.2) on the imbalance vector, $v_i = k_i^{\text{in}} - k_i^{\text{out}}$, which means that when the in- and out-degrees of all nodes are equal the right-hand side of the equation goes to zero. In a network with a perfect hierarchy like a directed line, the nodes are assigned integer levels with steps of one between connected nodes. The level distributions of real networks are more complex and lie somewhere between these extreme cases.

Trophic incoherence is a global parameter which measures how well hierarchically ordered the network is. It is related to the amount of feedback in the system, and thus to network properties such as the spectral radius, non-normality and strong connectivity [5,7]. It is quantified via the trophic incoherence parameter F, which is defined as

$$F = \frac{\sum_{ij} A_{ij}(h_j - h_i - 1)^2}{\sum_{ij} A_{ij}}. \tag{2.4}$$

In [5], the authors begin with equation (2.4) (which is the original definition of trophic incoherence [17]) and define the trophic levels, h, as those which minimize F. Hence, the linear form of equation (2.3) and dependence on the imbalance vector come from the minimization of the quadratic function in equation (2.4).

Trophic incoherence measures how far the mean square of the difference in trophic level, h calculated via equation (2.2), between start and end vertices of all the edges differs from one. The equation for trophic level, equation (2.2), can be derived by minimizing equation (2.4) with respect to h. The trophic incoherence takes values between 0 and 1 with real networks found on a spectrum between these extreme values. Networks with a perfect hierarchy are coherent and have $F = 0$. When the network is balanced like a directed cycle then $F = 1$. It is also possible to speak in terms of coherence instead of incoherence by using the quantity $1 - F$. When we talk about hierarchy we refer to the bottom of the hierarchy as nodes of low trophic level and the top of the hierarchy as the nodes of high trophic level. In a directed path the low trophic level nodes would be the start of the path and the high level nodes would be near the end of the path. In a food web the bottom nodes are plants and the top nodes are apex predators. This, however, is just a convention and the notion of top and bottom can be flipped by taking the transpose of the adjacency matrix (for example, when describing hierarchies of information flow from nodes to nodes).
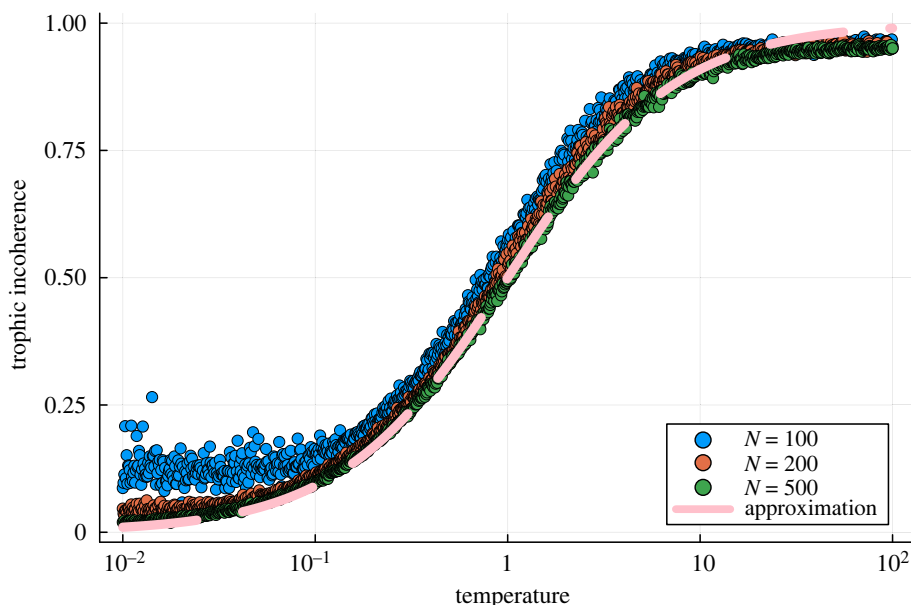
Figure 1. Relationship between trophic incoherence and generation temperature in the model used in this work, where each of the 1000 points, per network size, represents a distinct network generated by the model, with $N = 100$, $N = 200$ and $N = 500$, $\langle k \rangle = 20$ and no basal nodes. The generation temperature is logarithmically spaced between 10^{-2} and 10^{2}. We also plot with a dashed line an analytical approximation, given by equation (2.7), for the relationship between the trophic incoherence and temperature.

Using directionality and hierarchy to analyse the structure of directed networks in this way is quite natural and as such alternative similar formulations exists which rank nodes and measure the global directionality. SpingRank [24] views the ranking problem as minimizing the energy of a network of directed springs, leading to a similar minimization problem to the one used in trophic incoherence. However, these authors focus on the node-level ranks rather than the global directionality and add a regularization term to remove the invariance of their rank equation to addition of a constant vector. There is also a methodology based on Helmholtz–Hodge decomposition for measuring 'circularity' in directed networks [25,26] which has been shown to be equivalent to trophic analysis [5].

2.2. Network generation

In order to study how the properties of networks depend on their trophic structure it is necessary to be able to numerically generate networks which span the full range of trophic incoherence, while keeping the number of nodes and edges fixed. We adapt the generalized preferential preying model (GPPM) [21] in an identical way to Rodgers *et al.* [6] to the definition of trophic level which does not require basal nodes (vertices of zero in-degree) and use that to generate the networks we require. This works by generating an initial configuration of N nodes and a small number of random edges, calculating the initial trophic levels of this setup, adding edges up to the required amount using a probability determined by the trophic level and a temperature-like control parameter which can affect the spread of level differences. In order to generate a network with no basal nodes the initial configuration is that each node has in-degree one with the source vertex for that edge chosen uniformly at random from the other nodes in the network. Once the initial trophic levels $\tilde{h}$ are calculated via equation (2.2) then new edges, up to the required amount, can be added with probability defined as

$$P_{ij} = \exp\left[-\frac{(\tilde{h}_j - \tilde{h}_i - 1)^2}{2T_{\text{Gen}}}\right], \tag{2.5}$$

where P_{ij} is the probability of connecting nodes i to j. The parameter T_{Gen} controls this process. When this parameter is small it is likely that edges connect only between nodes where the level difference is 1 or near to it. When T_{Gen} is very large then the edge-addition probability goes towards 1 irrespective of the trophic level difference between the end and start nodes.

The relationship between the generation temperature and the trophic incoherence is displayed in figure 1, which shows how the model can be used to create a sample of networks of varying

incoherence. In order to efficiently sample the whole spectrum of the trophic incoherence we use logarithmic spacing of T_{Gen} throughout the paper, except for in the sections on the generalized rock–paper–scissors games and the out-component analysis. In these two sections, we compare networks of low, intermediate and high trophic incoherence, for which we selected the generation temperatures to be 0.02, 1 and 100, respectively. As shown in figure 1 this picks out each of the broad regimes of the generative model: the low temperature regime of very coherent networks, the intermediate regime where we see more variability with the temperature, and the high temperature regime of incoherent networks.

We also show in figure 1 how the behaviour of this model can be analytically approximated. In previous work on trophic analysis, where the ecological definition of trophic levels was used, the trophic incoherence, q, was defined as the standard deviation of the trophic level differences spanned by edges [17]. Under this definition the mean trophic level is always one, and a network is more incoherent the more heterogeneously the trophic differences are distributed around this value. Given a set of trophic levels (by whichever definition), it is possible to convert between the two measures of incoherence via the formula defined in [5]:

$$F = \frac{\eta^2}{1 + \eta^2},$$ (2.6)

where $\eta = \sigma/\mu$, σ is the standard deviation of the level differences and μ is the mean level difference. Hence, η is the appropriate replacement for the incoherence parameter q used in [14,17] and subsequent papers, when using the MacKay level definition [5]. If we assume that the trophic level difference distribution across the edges in the generated networks is distributed as in the edge addition probability, we can replace η^2 with T_{Gen}, if we assume that iteration of the attachment probability given by equation (2.5) leads to an approximately Gaussian distribution of trophic differences. The differences in trophic level have been found roughly to follow a Gaussian distribution in real networks. Also, analytical results which hold for real networks have been based on the assumption that the level difference distribution is Gaussian [7]. The assumption that the trophic levels follow the same Gaussian as implied by the edge-addition probability leads to the approximation that

$$F \approx \frac{T_{\mathrm{Gen}}}{1 + T_{\mathrm{Gen}}}.$$ (2.7)

As shown by figure 1 this fits quite well for the larger networks sizes. We plot the data with a logarithmic scaling on the x-axis which allows the different regimes of low, intermediate and high incoherence to be visualized and leads to the sigmoid shape, as writing $a = \log_{10} T_{\mathrm{Gen}}$ leads to the alternative form

$$F \approx \frac{1}{1 + 10^{-a}},$$ (2.8)

which is an equation of a sigmoid curve. This approximation fails at low temperatures for the small very dense networks, $N = 100$, as the generative model struggles to make networks which are very coherent at such a high edge density—it becomes difficult to place edges in locations where the coherence is maximized. We note that at high T_{Gen} we obtain $\eta^2 < T_{\mathrm{Gen}}$; however, the approximation works fairly well thanks to the sigmoid function, which is close to one for all large arguments. Hence, η^2 and T_{Gen} are not directly equivalent, but relatable when used to calculate F in this way. Nevertheless, this approximation provides an illustration of how the model works. Incoherence increases with temperature by augmenting the probability that edges are placed with edge difference further from one.

Additional detail of how to efficiently generate these networks by more efficiently sampling the space of edge addition probabilities so that each draw of a random number results in an edge being added is given in [6]. All graph manipulations in this work were carried out using the Julia package Graphs.jl [27].

3. Influence and influenceability of dynamics

In this section, we present a wide range of dynamics and show how trophic level and incoherence can be used to provide insight into disparate and unrelated processes. The idea is to use simple dynamics to highlight the phenomena being picked out by trophic analysis. In this case, we define 'influence' to mean the ability of a targeted perturbation or modification of the nodes to affect the state of the system some time after it is applied. In each of the dynamics we show how the nodes which can be regarded as 'influential' are the ones with low trophic level—i.e. those at the bottom of the hierarchy.

6

We also find that the 'influenceability' of the whole network by the low trophic level nodes depends on the network's trophic coherence. The goal of this section is to highlight the broad range of dynamics to which trophic analysis can be applied. In this section, we use numerically generated networks only—as opposed to the structural section where we provide results on numerically generated networks and a dataset of real-world networks [15]. We generate networks numerically with the model described above, which allows us to set the numbers of nodes and edges and to vary the trophic coherence. We can also set the number of basal nodes (nodes with zero in-degree). The dynamics of basal nodes are not affected by the states of other nodes, since they have no inputs. However, they are obvious candidates for nodes with a high influence on the whole network, and this has been demonstrated in previous work [9,16]. While basal nodes will tend to have low trophic levels, trophic analysis is not needed in order to identify them. We, therefore, fix the number of basal nodes in our generated networks to be zero, so as to highlight the importance of trophic level even in networks where all nodes can be affected by the rest of the system.

Our hypothesis is that the nodes at the lowest trophic levels (regardless of whether they are basal nodes) are the most influential across a broad range of different dynamics. We test this in the following way. We set up initial conditions such that all the nodes begin in one state or frequency, except for a 'perturbed fraction' of nodes which start in a different state or frequency. We choose the nodes with lowest trophic level for this perturbation, and observe the subsequent dynamics of the system. Whenever the system evolves to the state or frequency initially given to, say, only 5% of the nodes, we can conclude that: the network is highly 'influenceable', and these are the most 'influential' nodes. We go on to find that influenceability depends on trophic coherence, and influence is indeed determined by trophic level, across several dynamics.

We present the results as scatter plots of individual networks, but also show the averages and standard deviations of the sets of networks we generate. This allows the variability of the dynamics for a small number of influential nodes to be visualized as well as showing the general trend. For each of the dynamics we influence 5%, 20% and 40% of the nodes with the lowest trophic level to show how varying the size of this set affects the influence and the spread of the results.

3.1. Majority vote

Majority vote dynamics is the first dynamics to be studied as it is very simple to define as well as being widely used. This allows the primary focus to be on the network structure. Each agent (node) in the model is given an 'opinion' and then updates their opinion such that they select the opinion held by the majority of their neighbours. This model is very simplistic yet shares some properties with other common agent-based discrete dynamics widely used in complex networks, such as the SIS epidemic model [28], the Hopfield neural network model [6], the Ising model [29] and the Moran process from biology [30]. Each of these models has additional details which separate it from the basic majority vote, but the dynamics is governed by similar principles.

Majority vote dynamics can be defined as a system where there are two opinions, denoted as +1 and −1. An agent holds either of the two options and updates their state using an update rule defined as

$$S_i(t + \Delta t) = \text{sgn}\left(\sum_{j=1}^{N} A_{ji} S_j(t)\right). \tag{3.1}$$

The updates can be done asynchronously or in parallel and in this work we chose parallel updates. This specific update rule is deterministic but it can be modified in a range of ways to add stochastic behaviour to the system. Additionally, more than two possible states could be added to model a larger number of opinions. An absorbing state is reached if all nodes have the same opinion. However, it is also possible for the system to remain in a fixed point or in a cycle involving nodes with different opinions. In practice, we allow the system to run for a specified number of updates before recording the proportion of nodes in the +1 opinion.

The effectiveness of using trophic level to identify influential nodes is shown in figure 2. When a small number of nodes at the bottom of the hierarchy (nodes with low trophic level) are placed in the −1 opinion, most or all the rest of the nodes can evolve towards this state despite the majority of the nodes being initialized in the +1 opinion. This scenario holds except when the networks become too incoherent for the perturbed nodes to have a decisive influence. Viewed as a function of trophic incoherence, there is a transition from an influenceable regime (in which the low trophic level nodes dominate) to one which is less influenceable. Where this transition occurs varies with the fraction of perturbed nodes. Moreover, for a given trophic incoherence and perturbed fraction, some network
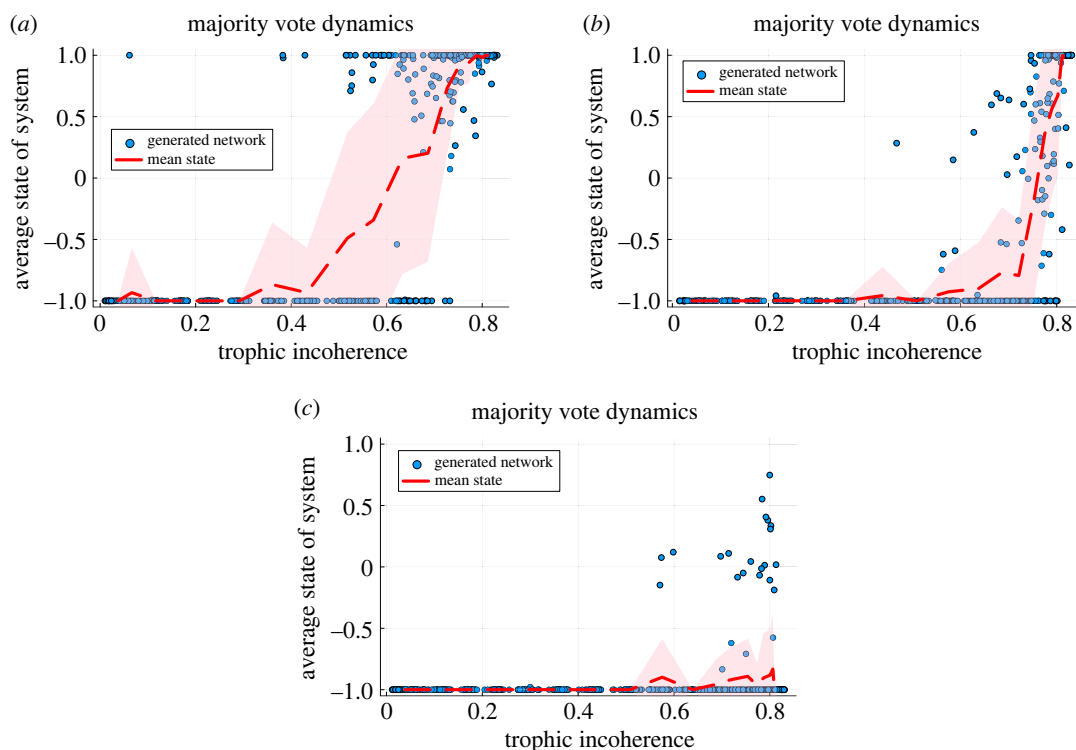
Figure 2. Majority vote dynamics after 1000 updates when a specified fraction of nodes are initially set to -1 with the rest of the nodes set to $+1$. This fraction is 5%, 20% and 40% in (a–c), respectively. The nodes set to -1 are those with lowest trophic level. The average state of the system (which can be related to the fraction in each state) is shown against trophic incoherence, with each point a distinct network generated by the model in §2.2 with $N = 500$ and $\langle k \rangle = 5$ and no basal nodes. Temperatures, T_{Gen}, are logarithmically spaced between 10^{-1} and 10^2 with 20 distinct temperatures used. At each temperature we generate 30 networks and compute the mean and standard deviation of this set with the mean plotted with the red dashed line and the standard deviation shown with the shaded area. (a) Influence by 5% of nodes with lowest trophic level. (b) Influence by 20% of nodes with lowest trophic level. (c) Influence by 40% of nodes with lowest trophic level.

realizations are highly influenceable and others are not, suggesting that there are other relevant network properties to be determined.

This is the expected behaviour in networks of lower incoherence as there is less feedback and cycles so the majority opinion at the bottom of the network moves up through the hierarchy, eventually changing the opinion of the whole system. In a network which is more incoherent this does not happen, for two reasons: there are fewer nodes connected only to the perturbed fraction; and there is more feedback (from directed cycles), which allows the unperturbed nodes to reinforce their initial opinion.

The standard deviation of the points is largest at the incoherence which is between the two regimes. When 40% of the nodes are perturbed, figure 2c, almost all nodes are influenced by the low level nodes apart from in the most incoherent networks, which may remain in an intermediate state. Similar behaviour is seen when 20% of nodes are perturbed, figure 2b, although there is now a clearer regime of networks which are too incoherent to be influenced by this fraction of nodes. When only 5% of nodes are perturbed, figure 2a, despite the number of nodes being small, they are able to influence the state of many of the networks of low coherence and the average state only increases as the incoherence becomes intermediate and there are more cycles and feedback in the system. There are a few outlier points at low incoherence. This is to be expected as we perturb such a small fraction of the nodes that there may be features of the network which make it difficult for them to influence the system. For example the number of edges which leave the set which is perturbed may be small or the trophic levels may be organized in such a way that the small set does not fill the lowest level of the network [6]. However, these simulations still demonstrate the usefulness of this method. Indeed, in very coherent networks specific targeted influence can overcome an opinion held by 95% of the network. This example provides perhaps the simplest demonstration of the insight trophic analysis can give into network dynamics. This may have potential applications in social networks and organizations, particularly in the promoting of cooperation in prisoner dilemma-like games on networks [31].
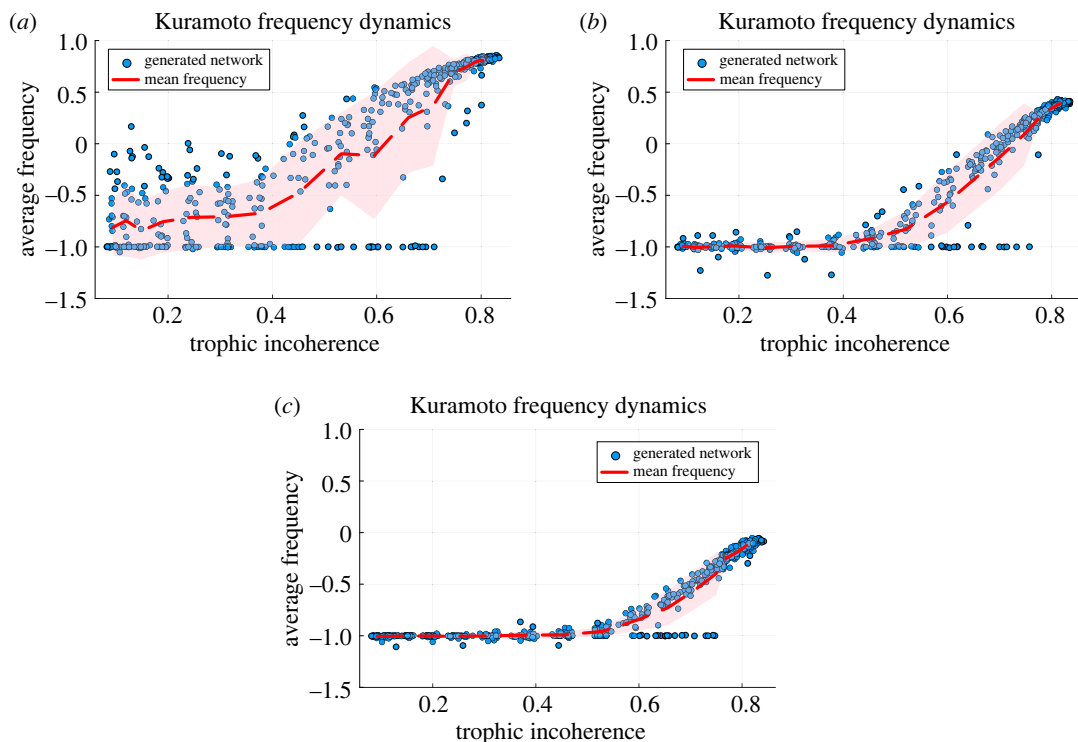
Figure 3. Kuramoto model dynamics after simulating for 20 time units when different fractions of nodes, chosen by lowest trophic level, are given intrinsic frequency −1 while the rest of the nodes are given intrinsic frequency +1; for varying trophic incoherence, where each point is a distinct network generated by the model in §2.2 with $N = 500$, $\langle k \rangle = 5$ and no basal nodes. Temperatures, T_{Gen}, are logarithmically spaced between 10^{-1} and 10^2 with 20 distinct temperatures used. At each temperature we generate 30 networks and compute the mean and standard deviation of this set, with the mean plotted with the red dashed line and the standard deviation shown with the shaded area. The coupling constant, $K = 100$, is twenty times the average degree to ensure synchronization. Nodes are initially synchronized with phase 0. (a) Influence by 5% of nodes with lowest trophic level. (b) Influence by 20% of nodes with lowest trophic level. (c) Influence by 40% of nodes with lowest trophic level.

3.2. Synchronization phenomena and Kuramoto oscillators

A very similar result to that observed in the discrete majority vote dynamics can be observed in the frequency synchronization of continuous Kuramoto oscillators, where the low level nodes can influence the frequency of the whole system. The Kuramoto model is a well-known model of synchronization with a wide range of applications, from neuroscience to power grids [32–36]. We use NetworkDynamics.jl [37] to solve the system of differential equations used in our variant of the model. Each oscillator (node) i has a phase, θ_i, which evolves according to the equation

$$\frac{d\theta_i}{dt} = \omega_i + \frac{K}{k_i^{in}} \sum_{j=1}^{N} A_{ji} \sin(\theta_j - \theta_i), \tag{3.2}$$

where K is the coupling constant, k_i^{in} is the in-degree of the node and ω_i is the natural frequency of the node. We use the form normalized by in-degree so that the oscillators update at similar rates regardless of the number of input nodes. For the system used we set the coupling constant to $K = 20\langle k \rangle$, which is well above the critical threshold for maintaining frequency and phase synchronization when the system is initialized at a single phase. This equation could be modified by adding noise or delays to the updates to make the system more realistic. However, we focus here on the simplest case. Phase synchronization can be measured by the order parameter

$$r = \frac{1}{N} \left| \sum_{i=1}^{N} e^{i\theta_i} \right|, \tag{3.3}$$

which reaches one when all the oscillators share the same phase but is zero when the phases are uniformly randomly distributed from 0 to 2π. How the low trophic level nodes influence the frequency in networks with varying trophic coherence is shown in figure 3, where a fraction of

the low trophic level nodes have natural frequency $\omega = -1$ and the remainder have $\omega = +1$. Average frequency is measured by sampling differences in node phase during the latter half of the simulation and dividing by the time differences, averaging over samples and nodes. Since we are interested in which frequency the nodes synchronize to, we set the initial condition so that the nodes begin with the same phase and the dynamics are driven by the topology and natural frequencies only, and not the varying initial conditions. The networks maintain phase synchronization, r close to 1, and adequate frequency synchronization, standard deviation of node frequencies is generally very small and always less than 1 (equivalent to standard error less than 0.045), as the simulation evolves due to the initial conditions and strong coupling. Networks with low incoherence propagate the frequency of the low-level nodes through the whole system and the network synchronizes at that frequency. In networks of high trophic incoherence the feedback from the higher level nodes back to the lower level nodes dilutes this effect, so the final average frequency of the system is further from that at the low level nodes and closer to that of the higher level nodes.

This transition is different from the majority vote transition since both phase and frequency are continuous, rather than the discrete states available to the majority vote dynamics. Additionally, in this case we are permanently changing an intrinsic property of the nodes which is maintained throughout the process, whereas previously the perturbation was to the initial conditions. However, the fact that applying the same principle to two very different systems results in qualitatively similar behaviour highlights the usefulness of trophic analysis as a tool. When only 5% of the low level nodes are selected in the Kuramoto case, figure 3a, the trend is noisy and the standard deviation is large. However, even in this regime we see the trend of the more coherent networks being more easily influenced by the low level nodes and only the most incoherent networks resistant to this. When 20% of the low level nodes are set, figure 3b, the trend is much clearer. The most coherent networks are all very strongly influenced by the low level nodes, and there is a clear transition as the network becomes more incoherent, with a lower standard deviation. When 40% of the nodes are perturbed, figure 3c, the system is more strongly influenced by the low level nodes and the frequency is reduced away from one, even in the most incoherent networks. In all cases, even at high incoherence there are some networks which synchronize with the 'influenced' (low level) nodes.

The oscillators begin in a synchronized state, so we find similar results when the end time of the simulation is varied. If a different oscillator function were used it may be possible to exploit the network hierarchy to create more complex states. It has recently been shown that oscillators on non-normal networks can lead to chimaera states [38], and it is known that non-normality is related to trophic coherence [5,16].

3.3. Voter model

A voter model is another way to model opinion formation for which similar results around influence and influenceability can be demonstrated. Similar to the majority vote dynamics the agents can be given states of either +1 or −1 and updated according to an update rule. However, instead of taking the majority opinion of the neighbours the agent can update its state by choosing at random one of its incoming neighbours and adopting the selected state. This model can be adapted, increased in complexity and then applied to study real voting processes [39], economics [40] and chemistry [41,42]. The model has two absorbing states where all nodes share the same opinion (consensus). However, after a finite time the system can find itself in an intermediate state which includes both opinions (and on a directed network the absorbing states can be inaccessible).

Similar results to the previously obtained ones can be seen, where a new opinion introduced at the lowest trophic level nodes spreads well in coherent networks but is less likely to establish itself in more incoherent networks. This is demonstrated in figure 4. In the low incoherence networks, the standard deviation is quite small and the ability of the low level nodes to influence the system is clear. As we increase incoherence this is no longer always the case and we pass through a region of higher variability as we transition to the incoherent regime, where the dynamics cannot be as easily shaped by the low level nodes. When 40% of the nodes are set to −1, figure 4c, almost all the networks flip to the minority state and only a small number of very incoherent networks maintain the majority state. Very similar behaviour is found when 20% of the nodes are set to −1, figure 4b, with most of the networks being easily influenced and only some of the high incoherence networks being resilient to the perturbation. When only 5% are set to −1, figure 4a, we see the same trend but a smaller fraction of the networks are influenced, with more networks being outliers from the trend.
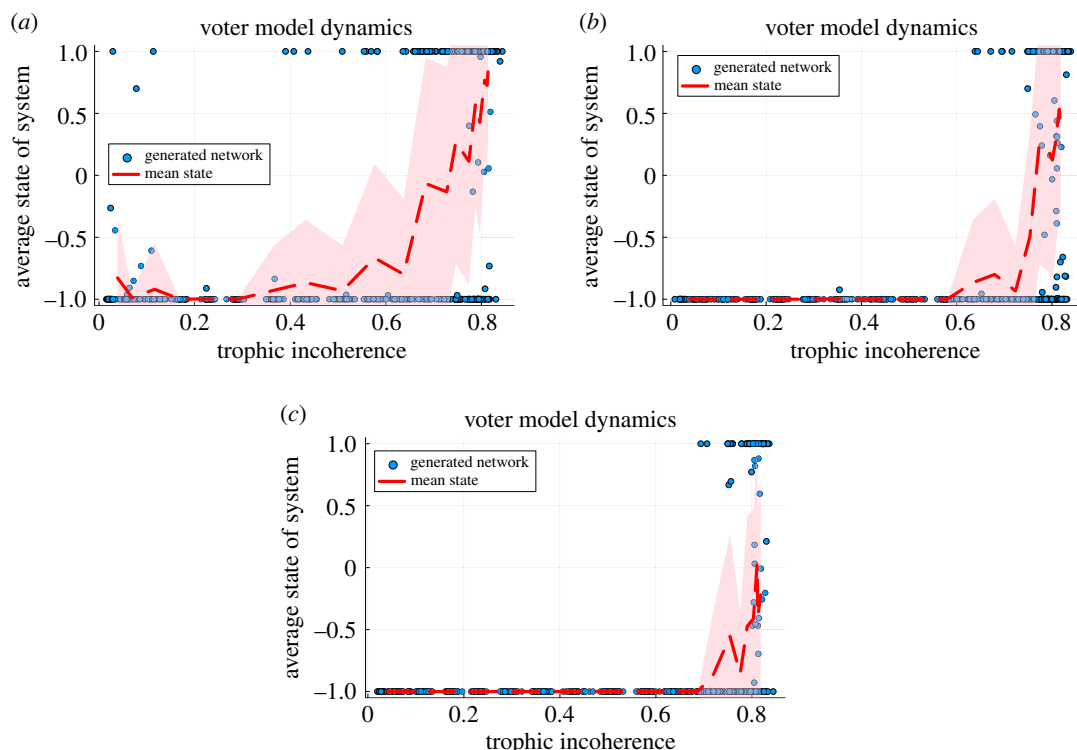
Figure 4. Voter model dynamics after 1000 updates when different fractions of nodes, chosen by lowest trophic level, are set to the initial condition −1 while the rest of the nodes are initially +1; for varying trophic incoherence, where each point is a distinct network generated by the model in §2.2 with $N = 500$, $\langle k \rangle = 5$ and no basal nodes. Temperatures, T_{Gen}, are logarithmically spaced between 10^{-1} and 10^2 with 20 distinct temperatures used. At each temperature we generate 30 networks and compute the mean and standard deviation of this set, with the mean plotted with the red dashed line and the standard deviation shown with the shaded area. (a) Influence by 5% of nodes with lowest trophic level. (b) Influence by 20% of nodes with lowest trophic level. (c) Influence by 40% of nodes with lowest trophic level.

In a similar way to the majority rule case, targeting a very small number of nodes makes it more likely that an insufficient number of edges leave the perturbed set to influence the whole network [6].

This section has shown how in different dynamical processes the nodes which can be regarded as influential can be determined by using trophic levels, and the ability of the network to be influenced is heavily shaped by the trophic coherence.

3.4. Generalized rock–paper–scissors dynamics

We show that trophic level can also be used as an analysis tool for the states of generalized rock–paper–scissors games where the interactions between strategies are defined by a complex network. A standard rock–paper–scissors game can be described with a directed cycle, in which each strategy has an edge to the strategy it defeats. In this case, there is no overall best choice of strategy thanks to the symmetry of the situation. This, however, may not be the case when the strategies interact through a more complex 'strategy network' which does not feature this cyclic structure (e.g. interactions between strategies and characters in video-games or ecological interactions such as competition between bacteria [43], as shown in figure 5). In such cases, there may be strategies which perform better than others.

The interaction between individual agents can be defined on a spatial structure like a lattice or by allowing every agent to interact with every other as in a complete graph. For simplicity we allow all the agents to interact with each other and ignore spatial structure. This setup shares many similarities with mean-field analysis of Lotka–Volterra systems, which have been studied with respect to trophic levels and coherence in [17]. It has also been much more widely studied in the regime of large random systems [44,45]. Most studies on non-random systems have looked at smaller systems, quite often with spatial aspects to them [46]. There has been some work looking at networks of generalized rock–paper–scissors games but it mostly focuses on varying the strength of parameters rather than looking at large varied structures. This work has mostly been numerical [47–51], with some looking at
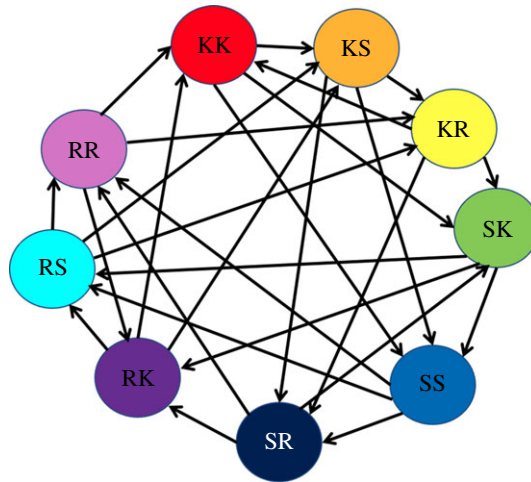
Figure 5. Example of a network of agents competing with each other through a complex interaction topology where an edge represents that an agent type is dominant over another. This is a network of bacteria competing through toxin warfare, drawn based on network structure data from [43]. Each node represents a type of bacteria (a strategy), which dominates over those types it has directed edges to, and is dominated by those it receives edges from. Hence, the network represents the relationship between bacteria rather than a spatial structure.

the behaviour of smaller systems with weighted interactions and weaker species. There has been some analytical work on rock–paper–scissors–lizard–spock networks which have a more complex topology than the basic rock–paper–scissors game [52].

The setup used here is a graph of 100 nodes representing the possible strategies. Then 1000 players per graph play each other at random with the interactions determined by the strategy graph. The initial strategy of each node is randomly assigned from the set of 100 possible strategies. If an edge points from strategy i to strategy j then if a player playing strategy j meets a player playing strategy i they will be beaten and switch from strategy i to strategy j. This definition could easily be flipped and then successful strategies would be those of high trophic level. However, we use this convention to match the earlier results. These dynamics can be quite complicated to analyse for several reasons. 'Weaker' strategies may survive better than expected if the strategies that they are weak to are made extinct before them. The presence of cycles and neutral interactions means that one strategy taking over everywhere is not the norm and instead the system moves to a dynamic equilibrium where many of the strategies coexist and 'defensive alliances' are possible [43]. What makes a strategy successful is a complex problem—it depends on the balance of how many strategies it beats or is beaten by, as well as where those strategies fit into the global meta-game. For example, it is better to be weak to a strategy which is overall not widely used and then strong against a strategy which is widely played. In this setup a strategy which is influential is one which becomes widely played. As a test for the ability of trophic level to predict influence we compare the probability of a strategy being played with the trophic level ranking of a node and then compare this to ranking by traditional centrality metrics such as PageRank [53,54].

The results are shown in figure 6 for networks with low, intermediate and high trophic incoherence. We consider the following centralities: trophic level, PageRank, out-degree and imbalance (the difference between in-degree and out-degree). In the case of PageRank and imbalance, we apply the centrality to the adjacency matrix such that influence should correlate with a high out-degree, which may require transposing the matrix depending on convention used. The higher the area under the curve for a given centrality, the better it performs at identifying the most successful strategies.

For the most coherent networks, the best predictor of influence is PageRank, closely followed by trophic level, figure 6a. In these networks PageRank correlates with trophic level but also reveals information about a node's reach, and we hypothesize that both these features contribute to a strategy's success. For networks of intermediate incoherence both trophic level and PageRank again are the best predictors of a strategy being successful, figure 6b, although imbalance is also now a good measure. However, this changes for networks with high incoherence, where the imbalance between in- and out-degrees starts to matter more than PageRank. Trophic level and the degree imbalance are now the best predictors, with PageRank performing similarly to out-degree, figure 6c. This demonstrates how trophic level is a good measure of influence as it compares well to the best centrality measure in each case. It also highlights the utility of trophic coherence as a measure to
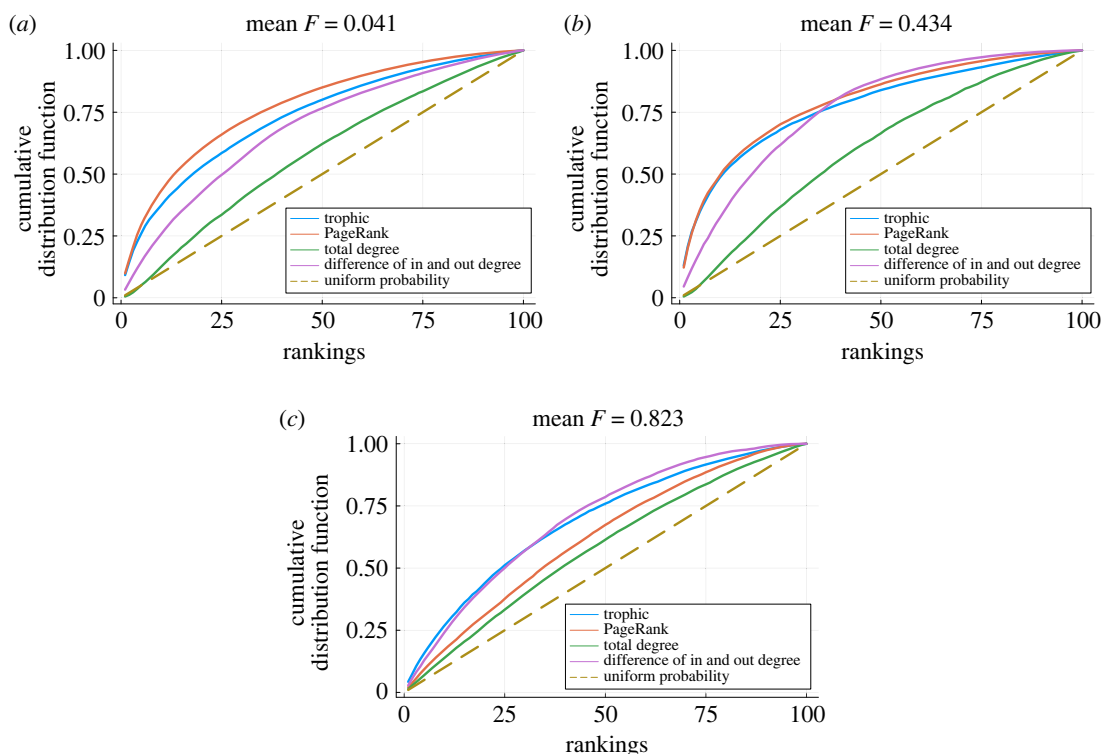
Figure 6. Cumulative distributions of probability of playing a strategy per player per game per network by different centrality rankings and different mean incoherence over 1000 network samples generated by the model in §2.2 at fixed temperatures corresponding to low, medium and high incoherence: $T_{GEN} = 0.02$, 1 and 100, respectively. Each network has $N = 100$, $\langle k \rangle = 5$ and 1000 players playing a generalized rock–paper–scissors game. (a) Networks of low trophic incoherence. (b) Networks of intermediate trophic incoherence. (c) Networks of high trophic incoherence.

understand network structure, since it helps to explain why the best centrality measure depends on the network. Similar results can be found for many standard network types like random graphs or scale-free networks, where again trophic level is competitive with the best other centrality metrics. It is noteworthy that trophic level is a good predictor of strategy success even in highly incoherent networks. However, if the network were maximally incoherent ($F = 1$), all the nodes would have the same trophic level and hence, according to this predictor, an equal probability of success. This is the case, for instance, in the standard rock–paper–scissors game.

4. Influence and influenceability of structure

In this section, we demonstrate how several topological notions of 'influence' and 'influenceability' based on the eigenvectors of the adjacency matrix can be understood and interpreted by looking at the trophic structure of directed networks. This section provides an explanation for the effect of trophic coherence on dynamics seen in the last section, and how networks can be better understood through trophic analysis. In particular, we show why the importance of low trophic level nodes depends on the network's trophic coherence. We provide results for numerically generated networks to control for variations in size, as well as results on a diverse dataset of real-world networks [15]. In real-world networks, such as the ones contained within our sample, there are a variety of network sizes, mean degrees and degree distributions. However, we show in this section that the general trends hold in both the numerically generated and real-world networks.

4.1. Localization of eigenvectors

One simple way to quantify influence in a network is eigenvector centrality, a widely used centrality metric [55]. The eigenvector centrality score of a node is determined by the centrality scores of its neighbours [55]. This makes sense as we assume that important nodes also connect to other nodes of high importance. This can be shown to be equivalent to the principal eigenvector of the adjacency
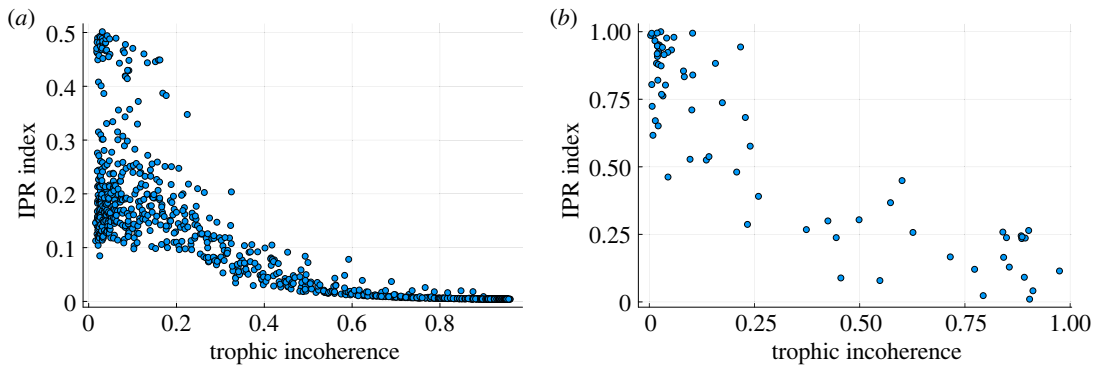
Figure 7. Average IPR of L^2 normalized eigenvectors of the adjacency matrix of real and generated networks for varying trophic incoherence. (*a*) 1000 data points each representing a network generated by the model in §2.2 of $N = 500$ nodes, $\langle k \rangle = 20$. Generation temperatures are logarithmically spaced between 10^{-2} and 10^2. (*b*) Real networks—taken from [15]. Each point represents a single network.

matrix and provides one measure of the importance of nodes in a network [55]. As we shall see, in trophically coherent networks the eigenvectors are localized. This has many implications for the dynamics and affects the sensitivity of the network to perturbation and being disrupted by a small subset of nodes. The localization of eigenvectors was first discovered in physics [56] where the phenomenon of electron localization known as Anderson localization was put forward. It has more recently been widely studied in random matrices [57]. It has wide implications for the dynamics of both physical systems and biological neural networks [58–60].

The localization of the eigenvectors can be measured in a variety of ways and it is also important to consider which nodes the eigenvectors localize around. This is because the ability to influence a network depends on the distribution of the centrality and where the high centrality nodes are in the network.

4.1.1. Inverse participation ratio

The inverse participation ratio (IPR) measures the localization of an eigenvector. It is defined for the nth eigenvector as

$$\mathrm{IPR}_n = \frac{\sum_{i=1}^{N} |\Psi_i^n|^4}{\left(\sum_{i=1}^{N} |\Psi_i^n|^2\right)^2}. \tag{4.1}$$

Here $|\Psi_i^n|$ is the absolute value of each component of the nth eigenvector. This sum is maximized when all non-zero components are concentrated at one value and takes its minimum value when the non-zero components are exactly equally distributed. In figure 7 we show, for each network, the average IPR_n over eigenvectors, $\mathrm{IPR} = N^{-1} \sum_n \mathrm{IPR}_n$. The localization varies with trophic incoherence: it is large when the network is more coherent and the eigenvalues are localized; it decreases when the network is more incoherent and the eigenvectors delocalize. Moreover, at low incoherence there is a higher variability in IPR between different networks.

For numerically generated networks, figure 7*a*, the effect is very clear. The inverse participation also follows the same trend in a dataset of real-world networks, figure 7*b*, which includes food webs, trade networks, neural and social networks [15]. However, the real networks vary in size, density and degree heterogeneity, and thus present a noisier picture. The maximum localization is also larger than in the numerically generated case. This helps to explain why in the previous section we found that a small number of low trophic level nodes can influence the dynamics when the network is more coherent, as this is the regime where the eigenvectors localize and 'influence' is concentrated.

4.1.2. Entropy of eigenvectors

The entropy of an eigenvector is another way to measure localization based on the information about the system carried by the vector. It is defined for the nth eigenvector Φ^n (normalized so that the absolute values of the elements sum to 1) as

$$S(\Phi^n) = \frac{-1}{\ln N} \sum_{i=1}^{N} |\Phi_i^n| \ln |\Phi_i^n|. \tag{4.2}$$
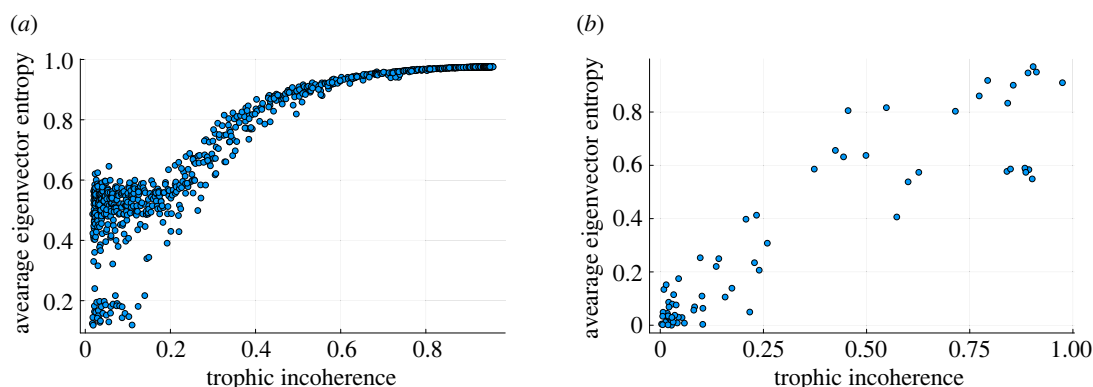
(a)

(b)

Figure 8. Average entropy of L^1 normalized eigenvectors of the adjacency matrix of real and generated networks for varying trophic incoherence. (a) 1000 data points each representing a network generated by the model in §2.2 of $N = 500$ nodes, $\langle k \rangle = 20$. Generation temperatures are logarithmically spaced between 10^{-2} and 10^2. (b) Real networks from [15]. Each point represents a single network.

This is minimized when all the weight of the eigenvector is concentrated in one node, and hence we know the most information about the system as all the importance is concentrated in that node. The entropy is maximized when the weight of all nodes is equal and all nodes are equivalent, so we have the least knowledge about the structure of the system. We normalize by the maximum value, $\ln N$, so that $0 \leq S \leq 1$ and we can compare networks of different size. A network can be characterized by the average, $\bar{S} = N^{-1} \sum_n S(\Phi^n)$. In figure 8, we show $\bar{S}$ for networks of varying trophic incoherence.

The trend in the numerically generated networks, figure 8a, is very clear, with the entropy of the very coherent networks being low and then increasing with incoherence. The real networks, figure 8b, produce a similar trend to the numerically generated networks, but again the data are noisier because of the more varied structure of real systems. This again is consistent with networks being sensitive to perturbations or attacks targeting the nodes of low trophic level, but the degree of sensitivity depending on the trophic coherence of the network.

4.2. Correlation between left and right eigenvectors

A node can be considered important if it is able to receive information from and input information into the rest of the network. In directed networks, the underlying hierarchy of trophic levels may mean that the nodes which are able to reach the most nodes in the network and the nodes which are reached by the most in the network are distinct groups. This is different from the undirected case where a hub node can effectively do both. This can be understood by looking at the correlation between the left and right principal eigenvectors. The left eigenvector can be thought of as a measure of the centrality of a node based on in-degree and its ability to receive information, and the right eigenvector can be thought of as the ability to emit information. This is just convention and transposing the adjacency matrix swaps these roles. One way to understand this is to look at the scalar product between the eigenvector centrality, given by the principal eigenvector of the adjacency matrix, and that of its transpose. Due to the Perron–Frobenius theorem both of these eigenvectors are real and non-negative so the multiplication of each element of the vector is real and non-negative. This is shown in figure 9a,b. We see that the correlation is small when the networks are coherent, as nodes of high centrality do not overlap; while when the network does not have such a well-defined hierarchical structure (i.e. it is more incoherent) there is a larger overlap, tending towards 1 as the eigenvectors are both normalized. This highlights the effects of hierarchy on the notion of node importance and that hierarchy can induce an asymmetry in node behaviour.

In the generated networks, the correlation between the eigenvectors follows a smooth trend while again in the case of real networks the overall correlation is reproduced but the results are much more noisy. These results can also be viewed in terms of matrix normality, which measures how well the adjacency matrix commutes with its transpose. This has previously been shown to be linked to trophic incoherence [5,16]. However, the correlation between right and left principal eigenvectors provides a more direct measure of whether the influential nodes are also those that are influenceable.
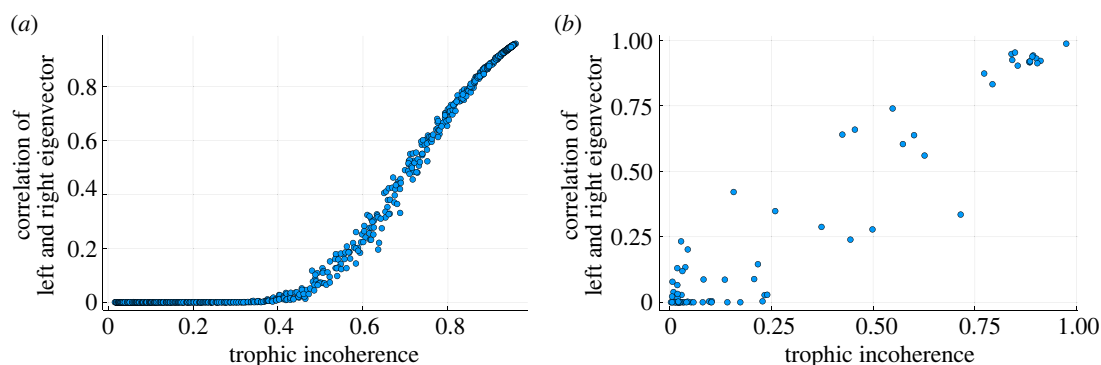
Figure 9. Correlation between left and right principal eigenvectors of the adjacency matrix of generated (*a*) and real (*b*) networks for varying trophic incoherence. Eigenvectors are L^2 normalized. (*a*) 1000 data points each representing a network generated by the model in §2.2 of $N = 500$ nodes, $\langle k \rangle = 20$. Generation temperatures are logarithmically spaced between 10^{-2} and 10^2. (*b*) Real networks from [15]. Each point represents a single network.
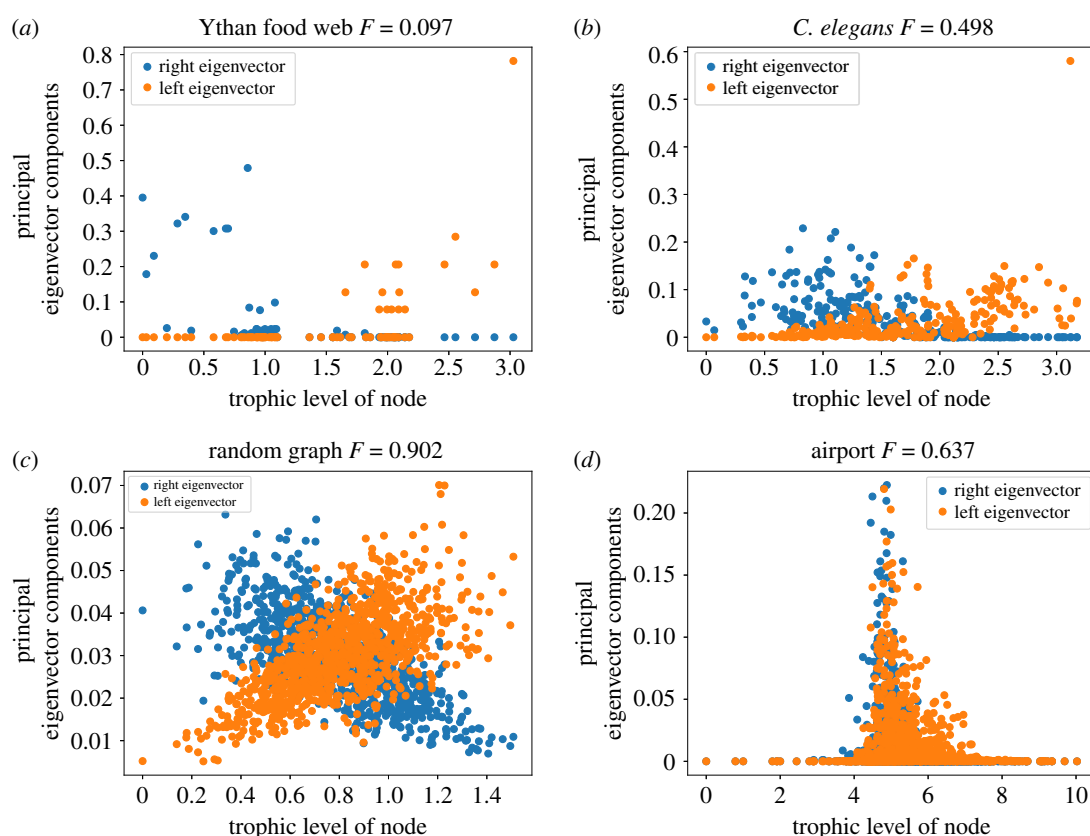

Figure 10. Principal left and right eigenvectors (L^2 normalized) for four example networks, with the trophic level of the nodes on the horizontal axis. (*a*) Ythan estuary food web with low incoherence, $F = 0.097$ [15]. Scalar product between left and right eigenvectors 0.003. (*b*) *C. elegans* connectome of intermediate incoherence, $F = 0.498$ [15]. Scalar product between left and right eigenvectors 0.278. (*c*) Random graph with $N = 1000$, $\langle k \rangle = 10$ and high incoherence with $F = 0.902$. Scalar product between left and right eigenvectors 0.904. (*d*) FAA preferred routes between airports, $F = 0.637$. Taken from [61]. Originally sourced from [62]. Scalar product between left and right eigenvectors 0.831.

4.2.1. Example networks

The previous results on the localization of the eigenvectors and the correlation between the left and right eigenvectors, on both numerically generated and real-world networks, can be better understood by looking at some specific network examples. In coherent networks with a clear hierarchy of trophic levels, such as the Ythan estuary food web, figure 10*a*, the left and right eigenvectors localize to different parts of the network hierarchy. The nodes of large left eigenvector centrality are towards the top of the hierarchy while the nodes of large right eigenvector centrality centre on the nodes of lower

trophic level. This reflects the intuition around the localization of eigenvectors and the difference between the left and right eigenvectors, as well as what we know about food webs where energy flows from the bottom to the top of the system. This phenomenon can also be observed in real networks of intermediate coherence such as the connectome of the nematode *C. elegans*, figure 10*b*. In this network, the localization is less pronounced than in the food web due to the network being more incoherent and having more feedback. There is, however, still clearly visible localization with the non-zero elements not being evenly distributed with respect to trophic level, and the distribution of the non-zero elements of the left and right eigenvectors being centred at different places in the hierarchy.

A non-trivial relationship between trophic level and the distribution of eigenvectors can also be observed in some random graphs, figure 10*c*, where one would not expect there to be much useful hierarchical structure to observe. In a random graph the left and right eigenvectors are not as strongly localized and there are many non-zero elements shared by both the left and right eigenvectors. However, there is a positive correlation between the left eigenvector and trophic level, and a negative correlation between the right eigenvector and trophic level. This implies that nodes which are the best at emitting information and reaching the whole network are at the bottom of the hierarchy, nodes which receive information are at the top and nodes which do both equally are in the middle. This agrees with the well-known bow tie structure observed in directed networks [63].

There do, however, exist networks which break this trend. Figure 10*d* shows the Federal Aviation Administration (FAA) preferred routes between airports. This network is directed, but since it is an airport network it features hubs which are hubs of both in- and out-degree. Due to the fact that certain airports are very important hubs in both senses, the left and right eigenvector peaks overlap more than in previous real-world examples, as we can see in figure 10*d*. Thus, this kind of network is quite different from, say, a food web, which has an overall directionality. The airport network is more coherent than a random graph (in part because the degree imbalance is larger). However, the level distribution does stretch out beyond the central peak to span a number of levels. This shows how trophic analysis can be used to identify structural features related to network function. In particular, it distinguishes clearly between networks with a directional flow, like a food web or neural network, and those lacking this property, like the airport network.

4.3. Pseudospectra and pseudospectral radius

A network can be influenced if it is sensitive to perturbations; in particular, if a disruption to a small number of nodes leads to a large change in the behaviour of the system. This kind of phenomenon can be linked to the non-normality of the adjacency matrix [64], i.e. how far it is from commuting with its transpose. Networks which are highly non-normal (a feature that has been shown analytically to be related to low trophic incoherence [5,16]) have the property that their eigenvalues are extremely sensitive to perturbations [64]. This can be impactful in many fields, such as fluid mechanics, acoustics, condensed matter physics and construction of numerical methods [64–68].

The pseudospectrum of a matrix A with eigenspectrum $\sigma(A)$ is defined as

$$\sigma_\epsilon(A) = \{\sigma(A + E) : \|E\| \leq \epsilon\}, \tag{4.3}$$

for some $\epsilon > 0$ and E a matrix of norm less than ϵ. It measures how the spectrum changes subject to perturbations, and can be computed using the Julia package [69] based on EigTool [70]. The pseudospectra have recently been shown to be important in complex networks and the stability of ecosystems [71].

One way to measure the sensitivity to perturbation is to compute the pseudospectral radius, the spectral radius of the perturbed matrix minus the original spectral radius, and divide by the size of the perturbation. If this quantity is of order 10^0 then the system is stable to perturbations of that scale. The size of perturbation used can be very small. In the following figures the perturbation is of order 10^{-3}. The pseudospectral radius exhibits a transition as trophic incoherence is varied for generated networks, figure 11*a*,*b*, being large for networks of low incoherence and then decreasing as incoherence increases, as we would expect. A similar result can be found for real-world networks, figure 11*c*.

We again produce a similar trend for both the real and generated networks, with the real network data appearing more noisy, as expected. As with the various dynamics described above, the pseudospectral radius demonstrates an increased sensitivity to perturbations—in this case structural—as networks become more coherent.
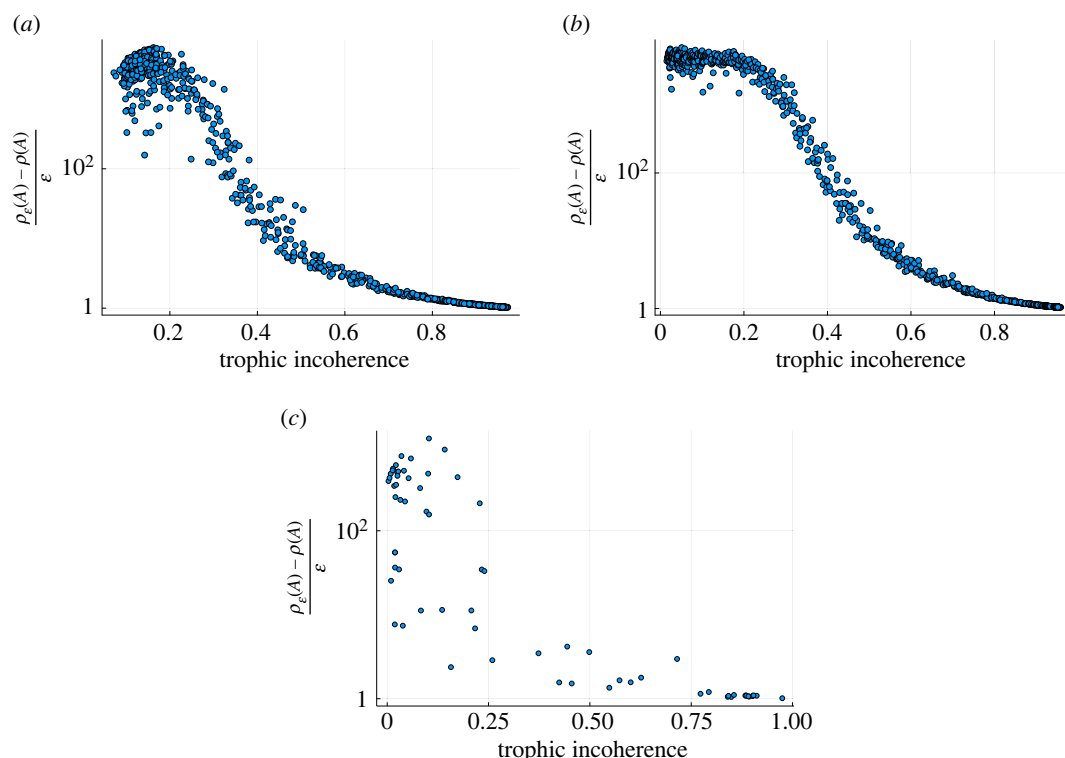
Figure 11. Pseudospectral radii scaled by perturbation size with trophic incoherence for numerically generated networks and real-world networks [15]. (a) 1000 data points each representing a network generated by the model in §2.2 of $N = 100$ nodes, $\langle k \rangle = 20$. Spectral perturbation size 10^{-3}. Generation temperatures are logarithmically spaced between 10^{-2} and 10^2. (b) 1000 data points each representing a network generated by the model in §2.2 of $N = 500$ nodes, $\langle k \rangle = 20$. Spectral perturbation size 10^{-3}. Generation temperatures are logarithmically spaced between 10^{-2} and 10^2. (c) Real networks—taken from [15] and excluding genetic networks for run time. Each point represents a single network. Spectral perturbation size 10^{-3}.

4.4. Out-components

One way to understand the transition in influence of nodes with hierarchical ordering is to look at the number of nodes which are reachable from a a given node—its 'out-component'—and how that changes with the trophic level of the node, and with the trophic coherence of the network. This is shown in figure 12.

When the network is very incoherent it is strongly connected and there is no heterogeneity in the ability of one node to influence other nodes. When the incoherence is intermediate there is still a large strongly connected component; however, the nodes at the top of the network are unable to influence the network as they lie outside this component. When the network is very coherent and F is close to zero, the strongly connected component is small and the size of the node's out-component is strongly linked to its trophic level. This leads to an asymmetry in the network due to the hierarchy of trophic levels, where some nodes can reach the whole network whereas others have a very small out-component. These results are a simple representation since the precise size of the out-component can vary widely depending on the type of network and how cycles are distributed throughout it. However, the general principle that networks of low incoherence lead to an asymmetry in component size is generally true due to the lack of feedback, and agrees with results from the literature [10]. This is also in agreement with other work on trophic analysis and the intuition behind it, which has analytically linked the emergence of a giant strongly connected component in a network to its trophic properties [7].

5. Discussion and conclusion

These results provide a different perspective on influence and influenceability in directed complex networks by using the fact that directed networks have an underlying hierarchical structure, unlike the undirected case. Thus, trophic analysis can give some insight into the function of a complex
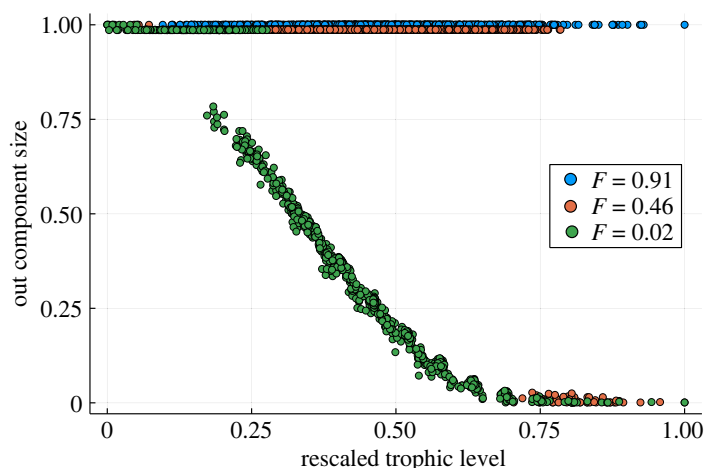
Figure 12. Size of the out-components of nodes with their trophic level rescaled by the maximum trophic level. Networks, generated by the model in §2.2 at a single temperature, of low, intermediate and high incoherence with $N = 1000$ and $\langle k \rangle = 10$. And $T_{GEN} = 0.02, 1, 100$ for the low, intermediate and high incoherence networks, respectively.

system. It reveals the links between many network properties, and how a node's trophic level can determine its ability to influence the rest of a network.

Trophic analysis is based on a simple calculation, a linear equation dependent only on the degrees and the adjacency matrix, and is straightforward to interpret. We hope that highlighting the diverse ways it can shed light on network function and structure might prompt its use in future work. In particular, when applied to real-world networks for analysis of structural features or understanding of dynamics, as we expect the phenomena we describe here to be quite general and applicable to many real-world systems. Since the equation is linear it can be solved for large real-world networks. Additionally, the results can be related to network properties such as the spectral radius, non-normality or strong connectivity [5,7]. In particular, networks of low trophic incoherence are highly non-normal. It has been shown recently that non-normal networks are very common [71], that non-normality of the adjacency matrix can have a large effect on the stability of a system [72], and it has been related to economic bubbles [73].

The general principle that the influence of a node is a function of its position in a hierarchy does not necessarily require the use of trophic levels to measure this and another metric such as PageRank could provide a similar result. However, trophic analysis allows a local measure of node importance to be related to trophic coherence, which gives a measure of how important the hierarchy is expected to be to the system. This is useful when trophic analysis is applied to real-world systems, such as when it has been used to investigate the hierarchy in a network of sustainable development goals [74,75]. This paper highlights how the importance of node-level measures of centrality (such as trophic level or PageRank) depends strongly on how these measures are distributed among the nodes, as well as on global network properties (trophic coherence) which may hinder or accentuate their effects. Trophic analysis can be a useful way to understand directed networks as it does not simply rely on either the in- or out-degree. In a directed network the ability to influence or be influenced may be of different relative importance depending on the dynamics, and it is possible that no correlation (nor more complex statistical relationship) exists between in- and out-degree. So nodes of importance in one measure may not be important in another. Trophic level simplifies this as the low level nodes are good at emitting and have many nodes downstream of them while nodes at the top of the network receive many paths from below. The interplay between trophic levels and degree distributions is something that could have a large impact on influence and will be the subject of future work. For example, a network could have a scale free distribution in either in- or out-degrees or both, which may or not be correlated, or bear some relationship to the trophic structure.

In conclusion, we have highlighted how there are many disparate ways to view influence and influenceability in real-world and generated networks, such as the ability to shape various discrete and continuous dynamics, the localization of eigenvectors, the sensitivity of the spectra to structural perturbations, or the distribution of out-components within networks. We have shown that all these phenomena can be thought of in terms of the local placement of nodes within the hierarchical structure of trophic levels, which is mediated by the global directionality given by trophic coherence. And we have described how this insight can be used to understand many structural and dynamical

processes in directed complex networks, which we hope prompts the use of trophic analysis in the study of specific real-world systems.

References

1. Lopes MA, Junges L, Woldman W, Goodfellow M, Terry JR. 2020 The role of excitability and network structure in the emergence of focal and generalized seizures. *Front. Neurol.* **11**, 74. (doi:10.3389/fneur.2020.00074)

2. Morone F, Makse H. 2015 Influence maximization in complex networks through optimal percolation. *Nature* **524**, 65–68. (doi:10.1038/nature14604)

3. Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, Makse HA. 2010 Identification of influential spreaders in complex networks. *Nat. Phys.* **6**, 888–893. (doi:10.1038/nphys1746)

4. Beguerisse-Díaz M, Garduño-Hernández G, Vangelov B, Yaliraki SN, Barahona M. 2014 Interest communities and flow roles in directed networks: the Twitter network of the UK riots. *J. R. Soc. Interface* **11**, 20140940. (doi:10.1098/rsif.2014.0940)

5. MacKay RS, Johnson S, Sansom B. 2020 How directed is a directed network? *R. Soc. Open Sci.* **7**, 201138. (doi:10.1098/rsos.201138)

6. Rodgers N, Tiňo P, Johnson S. 2022 Network hierarchy and pattern recovery in directed sparse Hopfield networks. *Phys. Rev. E* **105**, 064304. (doi:10.1103/PhysRevE.105.064304)

7. Rodgers N, Tiňo P, Johnson S. 2023 Strong connectivity in real directed networks. *Proc. Natl Acad. Sci. USA* **120**, e2215752120. (doi:10.1073/pnas.2215752120)

8. Eom YH, Shepelyansky DL. 2015 Opinion formation driven by PageRank node influence on directed networks. *Physica A* **436**, 707–715. (doi:10.1016/j.physa.2015.05.095)

9. Wright EAP, Yoon S, Ferreira AL, Mendes JFF, Goltsev AV. 2019 The central role of peripheral nodes in directed network dynamics. *Sci. Rep.* **9**, 13162. (doi:10.1038/s41598-019-49537-8)

10. la Cruz Cabrera OD, Matar M, Reichel L. 2019 Analysis of directed networks via the matrix exponential. *J. Comput. Appl. Math.* **355**, 182–192. (doi:10.1016/j.cam.2019.01.015)

11. Masuda N, Kawamura Y, Kori H. 2009 Analysis of relative influence of nodes in directed networks. *Phys. Rev. E* **80**, 046114. (doi:10.1103/PhysRevE.80.046114)

12. Masuda N, Kori H. 2010 Dynamics-based centrality for directed networks. *Phys. Rev. E* **82**, 056107. (doi:10.1103/PhysRevE.82.056107)

13. Chen DB, Gao H, Lü L, Zhou T. 2013 Identifying influential nodes in large-scale directed networks: the role of clustering. *PLoS ONE* **8**, e77455. (doi:10.1371/journal.pone.0077455)

14. Johnson S, Jones NS. 2017 Looplessness in networks is linked to trophic coherence. *Proc. Natl Acad. Sci. USA* **114**, 5618–5623. (doi:10.1073/pnas.1613786114)

15. Johnson S. 2020 Data repository. See https://www.samuel-johnson.org/data (accessed October 2020).

16. Johnson S. 2020 Digraphs are different: why directionality matters in complex systems. *J. Phys.: Complex.* **1**, 015003. (doi:10.1088/2632-072X/ab8e2f)

17. Johnson S, Domínguez-García V, Donetti L, Muñoz MA. 2014 Trophic coherence determines food-web stability. *Proc. Natl Acad. Sci. USA* **111**, 17923–17928. (doi:10.1073/pnas.1409077111)

18. Shuaib C, Syed M, Halawi D, Saquib N. 2022 Trophic analysis of a historical network reveals temporal information. *Appl. Netw. Sci.* **7**, 31. (doi:10.1007/s41109-022-00469-9)

19. Pilgrim C, Guo W, Johnson S. 2020 Organisational social influence on directed hierarchical graphs, from tyranny to anarchy. *Sci. Rep.* **10**, 4388. (doi:10.1038/s41598-020-61196-8)

20. Klaise J, Johnson S. 2017 The origin of motif families in food webs. *Sci. Rep.* **7**, 16197. (doi:10.1038/s41598-017-15496-1)

21. Klaise J, Johnson S. 2016 From neurons to epidemics: how trophic coherence affects spreading processes. *Chaos* **26**, 065310. (doi:10.1063/1.4953160)

22. Pagani A, Mosquera G, Alturki A, Johnson S, Jarvis S, Wilson A, Guo W, Varga L. 2019 Resilience or robustness: identifying topological vulnerabilities in rail networks. *R. Soc. Open Sci.* **6**, 181301. (doi:10.1098/rsos.181301)

23. Pagani A, Meng F, M Musolesi GF, Guo W. 2020 Quantifying resilience via multiscale feedback loops in water distribution networks. *J. Water Resour. Plan. Manag.* **146**, 04020039. (doi:10.1061/(ASCE)WR.1943-5452.0001231)

24. Bacco CD, Larremore DB, Moore C. 2018 A physical model for efficient ranking in networks. *Sci. Adv.* **4**, eaar8260. (doi:10.1126/sciadv.aar8260)

25. Kichikawa Y, Iyetomi H, Iino T, Inoue H. 2019 Community structure based on circular flow in a large-scale transaction network. *Appl. Netw. Sci.* **4**, 92. (doi:10.1007/s41109-019-0202-8)

26. Iyetomi H, Ikeda Y, Mizuno T, Ohnishi T, Watanabe T. 2017 International trade relationship from a multilateral point of view. In *6th Int. Conf. on Complex Networks & Their Applications, Lyon, France, 29 November–1 December 2017*, pp. 253–255. (doi:10.5281/zenodo.8019395)

27. Fairbanks J, Besançon M, Simon S, Hoffiman J, Eubank N, Karpinski S. 2021 Juliagraphs/graphs.jl: an optimized graphs package for the Julia programming language. See https://github.com/JuliaGraphs/Graphs.jl.

28. Volkening A, Linder DF, Porter MA, Rempala GA. 2020 Forecasting elections using compartmental models of infection. *SIAM Rev.* **62**, 837–865. (doi:10.1137/19M1306658)

29. Kim M, Yook SH. 2021 Majority-vote model with degree-weighted influence on complex networks. *Phys. Rev. E* **103**, 022302. (doi:10.1103/PhysRevE.103.022302)

30. Moinet A, Barrat A, Pastor-Satorras R. 2018 Generalized voterlike model on activity-driven networks with attractiveness. *Phys. Rev. E* **98**, 022303. (doi:10.1103/PhysRevE.98.022303)

31. Pena J, Pestelacci E, Tomassini M, Volken H. 2009 Conformity and network effects in the prisoner's dilemma. *IEEE Congress on Evolutionary Computation, Trondheim, Norway, 18–21 May 2009*. (doi:10.1109/CEC.2009.4982988)

32. Kuramoto Y. 1984 *Chemical oscillations, waves, and turbulence*. Berlin, Germany: Springer-Verlag.

33. Acebrón JA, Bonilla LL, Vicente CJP, Ritort F, Spigler R. 2005 The Kuramoto model: a simple paradigm for synchronization phenomena. *Rev. Mod. Phys.* **77**, 137–185. (doi:10.1103/RevModPhys.77.137)

34. Bick C, Goodfellow M, Laing CR, Martens EA. 2020 Understanding the dynamics of biological and neural oscillator networks through exact mean-field reductions: a review. *J. Math. Neurosci.* **10**, 9. (doi:10.1186/s13408-020-00086-9)

35. Cumin D, Unsworth CP. 2007 Generalising the Kuramoto model for the study of neuronal synchronisation in the brain. *Physica D* **226**, 181–196. (doi:10.1016/j.physd.2006.12.004)

36. Guo Y, Zhang D, Li Z, Wang Q, Yu D. 2021 Overviews on the applications of the Kuramoto model in modern power system analysis. *Int. J. Electr. Power Energy Syst.* **129**, 106804. (doi:10.1016/j.ijepes.2021.106804)

37. Lindner M, Lincoln L, Drauschke F, Koulen JM, Würfel H, Plietzsch A, Hellmann F. 2021 NetworkDynamics.jl: composing and simulating complex networks in Julia. *Chaos* **31**, 063133. (https://pubmed.ncbi.nlm.nih.gov/34241293/)

38. Muolo R, O'Brien JD, Carletti T, Asllani M. 2023 Persistence of chimera states and the challenge for synchronization in real-world networks. (http://arxiv.org/abs/2306.00237)

39. Fernández-Gracia J, Suchecki K, Ramasco JJ, San Miguel M, Eguíluz VM. 2014 Is the voter model a model for voters? *Phys. Rev. Lett.* **112**, 158701. (doi:10.1103/PhysRevLett.112.158701)

40. Kirman A. 1993 Ants, rationality, and recruitment. *Q. J. Econ.* **108**, 137–156. (doi:10.2307/2118498)

41. Fichthorn K, Gulari E, Ziff R. 1989 Noise-induced bistability in a Monte Carlo surface-reaction model. *Phys. Rev. Lett.* **63**, 1527–1530. (doi:10.1103/PhysRevLett.63.1527)

42. Considine D, Redner S, Takayasu H. 1989 Comment on 'Noise-induced bistability in a Monte Carlo surface-reaction model'. *Phys. Rev. Lett.* **63**, 2857. (doi:10.1103/PhysRevLett.63.2857)

43. Szabó P, Czárán T, Szabó G. 2007 Competing associations in bacterial warfare with two toxins. *J. Theor. Biol.* **248**, 736–744. (doi:10.1016/j.jtbi.2007.06.022)

44. Garcia Lorenzana G, Altieriiulia A. 2022 Well-mixed Lotka-Volterra model with random strongly competitive interactions. *Phys. Rev. E* **105**, 024307. (doi:10.1103/PhysRevE.105.024307)

45. Mambuca AM, Cammarota C, Neriiulia I. 2022 Dynamical systems on large networks with predator-prey interactions are stable and exhibit oscillations. *Phys. Rev. E* **105**, 014305. (doi:10.1103/PhysRevE.105.014305)

46. Szolnoki A, Mobilia M, Jiang LL, Szczesny B, Rucklidge AM, Perciulia M. 2014 Cyclic dominance in evolutionary games: a review. *J. R. Soc. Interface* **11**, 20140735. (doi:10.1098/rsif.2014.0735)

47. Avelino PP, De Oliveira BF, Trintin RS. 2020 Performance of weak species in the simplest generalization of the rock-paper-scissors model to four species. *Phys. Rev. E* **101**, 062312. (doi:10.1103/PhysRevE.101.062312)

48. Avelino PP, de Oliveira BF, Trintin RS. 2022 Lotka-Volterra versus May-Leonard formulations of the spatial stochastic rock-paper-scissors model: the missing link. *Phys. Rev. E* **105**, 024309. (doi:10.1103/PhysRevE.105.024309)

49. Avelino PP, de Oliveira BF, Trintin RS. 2022 Parity effects in rock-paper-scissors type models with a number of species $NS \leq 12$. *Chaos Solit. Fractals* **155**, 111738. (doi:10.1016/j.chaos.2021.111738)

50. Milne PA, Brown PJ, Roming PWA, Avelino PP, Oliveira BFD, Trintin RS. 2021 Weak species in rock-paper-scissors models. *Europhys. Lett.* **134**, 48001. (doi:10.1209/0295-5075/134/48001)

51. Menezes J, Barbalho R. 2023 How multiple weak species jeopardise biodiversity in spatial rock–paper–scissors models. *Chaos Solit. Fractals* **169**, 113290. (doi:10.1016/j.chaos.2023.113290)

52. Postlethwaite CM, Rucklidge AM. 2022 Stability of cycling behaviour near a heteroclinic network model of Rock–Paper–Scissors–Lizard–Spock. *Nonlinearity* **35**, 1702–1733. (doi:10.1088/1361-6544/ac3560)

53. Brin S, Page L. 1998 The anatomy of a large-scale hypertextual web search engine. *Comput. Netw. ISDN Syst.* **30**, 107–117. (doi:10.1016/S0169-7552(98)00110-X)

54. Gleich DF. 2015 PageRank beyond the web. *SIAM Rev.* **57**, 321–363. (doi:10.1137/140976649)

55. Newman M. 2010 *Networks: an introduction*. Oxford, UK: Oxford University Press.

56. Anderson PW. 1958 Absence of diffusion in certain random lattices. *Phys. Rev.* **109**, 1492–1505. (doi:10.1103/PhysRev.109.1492)

57. Lucas Metz F, Neri I, Rogers T. 2019 Spectral theory of sparse non-Hermitian random matrices. *J. Phys. A: Math. Theor.* **52**, 434003. (doi:10.1088/1751-8121/ab1ce0)

58. Zhang GH, Nelson DR. 2019 Eigenvalue repulsion and eigenvector localization in sparse non-Hermitian random matrices. *Phys. Rev. E* **100**, 052315. (doi:10.1103/PhysRevE.100.052315)

59. Amir A, Hatano N, Nelson DR. 2016 Non-Hermitian localization in biological networks. *Phys. Rev. E* **93**, 042310. (doi:10.1103/PhysRevE.93.042310)

60. Tanaka H, Nelson DR. 2019 Non-Hermitian quasilocalization and ring attractor neural networks. *Phys. Rev. E* **99**, 062406. (doi:10.1103/PhysRevE.99.062406)

61. Peixoto TP. 2022 The Netzschleuder network catalogue and repository: FAA Routes Network. See https://networks.skewed.de/net/faa_routes (accessed December 2022).

62. FAA. 2010 Air traffic control system command center. See http://www.fly.faa.gov/.

63. Timár G, Goltsev AV, Dorogovtsev SN, Mendes JFF. 2017 Mapping the structure of directed networks: beyond the bow-tie diagram. *Phys. Rev. Lett.* **118**, 078301. (doi:10.1103/PhysRevLett.118.078301)

64. Trefethen LN, Embree M. 2020 *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*. Princeton, NJ: Princeton University Press.

65. Okuma N, Sato M. 2020 Hermitian zero modes protected by nonnormality: application of pseudospectra. *Phys. Rev. B* **102**, 014203. (doi:10.1103/PhysRevB.102.014203)

66. Gebhardt T, Grossmann S. 1994 Chaos transition despite linear stability. *Phys. Rev. E* **50**, 3705–3711. (doi:10.1103/PhysRevE.50.3705)

67. Symon S, Rosenberg K, Dawson STM, McKeon BJ. 2018 Non-normality and classification of amplification mechanisms in stability and resolvent analysis. *Phys. Rev. Fluids* **3**, 053902. (doi:10.1103/PhysRevFluids.3.053902)

68. Sujith RI, Juniper MP, Schmid PJ. 2016 Non-normality and nonlinearity in thermoacoustic instabilities. *Int. J. Spray Combust. Dyn.* **8**, 119–146. (doi:10.1177/1756827716651571)

69. Github Repository. Pseudospectra.jl. See https://github.com/RalphAS/Pseudospectra.jl (accessed August 2022).

70. Wright TG. 2022 EigTool. See http://www.comlab.ox.ac.uk/pseudospectra/eigtool/ (accessed August 2022).

71. Asllani M, Lambiotte R, Carletti T. 2018 Structure and dynamical behavior of non-normal networks. *Sci. Adv.* **4**, eaau9403. (doi:10.1126/sciadv.aau9403)

72. Duan C, Nishikawa T, Eroglu D, Motter AE. 2022 Network structural origin of instabilities in large complex systems. *Sci. Adv.* **8**, 8310. (doi:10.1126/sciadv.abm8310)

73. Sornette D, Lera SC, Lin J, Wu K. 2022 Non-normal interactions create socio-economic bubbles. (http://arxiv.org/abs/2205.08661)

74. Dawes JHP. 2022 SDG interlinkage networks: analysis, robustness, sensitivities, and hierarchies. *World Dev.* **149**, 105693. (doi:10.1016/j.worlddev.2021.105693)

75. Dawes JHP, Zhou X, Moinuddin M. 2022 System-level consequences of synergies and trade-offs between SDGs: quantitative analysis of interlinkage networks at country level. *Sustain. Sci.* **17**, 1435–1457. (doi:10.1007/s11625-022-01109-y)

76. Rodgers N, Tiňo P, Johnson S. 2023 Influence and influenceability: global directionality in directed complex networks. *Zenodo*. (doi:10.5281/zenodo.8063465)