Intervality and coherence in complex networks

Virginia Domínguez-García,1 Samuel Johnson,2 and Miguel A. Muñoz1
1) Departamento de Electromagnetismo y Física de la Materia and Instituto Carlos I de Física Teórica y Computacional. Universidad de Granada. E-18071, Granada, Spain.
2) Warwick Mathematics Institute, and Centre for Complexity Science, University of Warwick, Coventry, CV4 7AL, United Kingdom.

Food webs – networks of predators and prey – have long been known to exhibit “intervality”: species can generally be ordered along a single axis in such a way that the prey of any given predator tend to lie on unbroken compact intervals. Although the meaning of this axis – identified with a “niche” dimension – has remained a mystery, it is assumed to lie at the basis of the highly non-trivial structure of food webs. With this in mind, most trophic network modelling has for decades been based on assigning species a niche value by hand. However, we argue here that intervality should not be considered the cause but rather a consequence of food-web structure. First, analysing a set of 46 empirical food webs, we find that they also exhibit predator intervality: the predators of any given species are as likely to be contiguous as the prey are, but in a different ordering. Furthermore, this property is not exclusive of trophic networks: several networks of genes, neurons, metabolites, cellular machines, airports, and words are found to be approximately as interval as food webs. We go on to show that a simple model of food-web assembly which does not make use of a niche axis can nevertheless generate significant intervality. Therefore, the niche dimension (in the sense used for food-web modelling) could in fact be the consequence of other, more fundamental structural traits, such as trophic coherence. We conclude that a new approach to food-web modelling is required for a deeper understanding of ecosystem assembly, structure and function, and propose that certain topological features thought to be specific of food webs are in fact common to many complex networks.

PACS numbers: 5.65.+b, 87.10.Mn, 89.75.-k, 89.75.Fb, 89.75.Hc

For decades food-web modelling has been based on the idea of a “niche dimension”, according to which the species in an ecosystem are considered to be arranged in a specific order, which is tantamount to the existence of a one-dimensional hidden dimension. This assumption is justified by the empirical observation of a topological feature, exhibited by many food webs to a significant degree, called “intervality”. We show here that intervality is not necessarily the hallmark of a hidden niche dimension, but may ensue from other food-web structural properties, such as trophic coherence. In fact, we find instances of networks of genes, neurons, metabolites, cellular machines, airports, and words which exhibit intervality as significant as that of food webs. These results support a new approach to food-web modelling, and suggest that certain features of trophic networks are relevant for directed networks in general.

I. INTRODUCTION

Charles Darwin concluded On the Origin of Species reminiscing on his famous entangled bank, “clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, [...] so different from each other, and dependent on each other in so complex a manner”[1] Charles Elton later developed the concept of a food web – a network of predators and prey – as a description for a community of species[2] and in the eighties such systems were among the first to be explicitly modelled as random graphs with specific constraints.[3,4] With the advent of ever better ecological data and the explosion of research on complex networks, much work has gone into analysing and modelling the structure of food webs, and its relation to population dynamics and ecosystem function.[5,6] Not least among the motivations for such research has been an awareness that the sixth mass extinction is under way, and that we must strive to understand ecosystems if we are to protect them.[7]

Field ecologists apply a variety of techniques to infer the predation links which exist between (and sometimes within) the dozens, or hundreds, of species making up specific ecosystems. The results of such observations are sets of trophic networks, or food webs, which can now be analysed quantitatively, as we go on to do here. A food web with $S$ species can be encoded in an $S \times S$ adjacency matrix $A$, such that the element $A_{ij}$ is equal to one if species $i$ (the predator) consumes species $j$ (its prey), and zero if not. In other words, a food web can be regarded as an unweighted, directed network in which the nodes are species and the directed edges represent predation.

When Joel Cohen first examined a set of such food webs in the seventies, he discovered that they exhibited a topological property which he named intervality: the species could be ordered in a line in such a way that the prey of any given predator would form a compact
interval. In terms of the adjacency matrix, this meant that the columns could be ordered so that elements would form unbroken horizontal blocks (see Figure 1A). This observation re-invigorated the use of an important and old concept in ecology: the niche. The term was originally used simply to refer to a species’ habitat or ecological role, but was then defined by Hutchinson as a “position in a multi-dimensional hyperspace” – each dimension being some biologically relevant magnitude. The observed intervality of food webs suggested that predators consume every species within a particular compact hypervolume of niche space – and, moreover, that niche space was (at least to a good approximation) one-dimensional.

Motivated by the belief that complex systems could come about from simple rules, several models were put forward to explain the non-trivial structure of trophic networks. The cascade model was the first such attempt. The key idea behind this model is that food webs reflect some inherent hierarchy in which certain species are above others. Hence, species are organised on a hierarchical (niche) axis and are constrained to consume only prey which are below them. The approach is loosely based on the fact that predators tend to be larger than their prey, at least in certain kinds of ecosystem, and so the hierarchy could be regarded as one defined by body size. The cascade model was better at reproducing food-web structure than a fully random graph, but there were features – intervality, in particular – which it could not account for. Then Williams and Martinez put forward the well-known niche model, in which species are again ordered along a hierarchical or niche axis, but species are allowed to select prey now only on a contiguous interval below them (not a random selection, as in the cascade model). Thus, the niche model has intervality as a built-in property. It also proved quite successful in explaining certain other food-web properties, and, thanks to its simplicity, it is often taken as the reference model for generating synthetic food webs. The popularity enjoyed by the niche model has served to reinforce the belief that something akin to its niche axis does in fact form the backbone of real ecosystems.

As progressively better collections of food-web data were gathered, it became apparent that most trophic networks were not perfectly interval (see Figure 1B). Methods for a more precise definition). Several modifications of the niche model have been proposed to overcome the drawback of perfect intervality that the niche model introduces by construction. Both the “generalized niche” mode and the “minimum potential niche” model relax the assumption that the prey of a given predator must form a perfectly continuous interval.
nale is that the imperfect intervality of food webs might be a sign that there are, in fact, more than one niche dimension. Despite these developments, some kind of niche axis is still generally thought to underlie food-web structure.

Here, we challenge this view by appealing to three observations.

First, we highlight that predator intervality (the extent to which the predators of a given species can be arranged on unbroken intervals), i.e. the predator intervality of a matrix $A$ is the prey intervality of its transpose, $A^T$) seems to be as general and non-trivial a feature of food webs as prey-intervality. While it is conceivable that predators might “choose” their prey according to a niche axis, how can prey simultaneously select their predators according to a different ordering?

Second, we show that intervality is not a feature peculiar to food webs, as has usually been assumed: complex networks of various kinds – including those of genes, neurons, metabolites, cellular machines, airports, and words – exhibit levels of both prey (column) and predator (row) intervality similar to those of food webs. Yet it appears far-fetched to suggest that such diverse systems all owe their structure to some kind of hidden niche axis.

Third, we show that the recently proposed “preferential preying” model of food-web assembly, which does not involve a niche axis, generates significant intervality. This is the first model to correctly reproduce the trophic coherence of food webs – i.e. the fact that species can be assigned trophic levels and predators have a tendency to prey upon subsets of prey which are on similar such levels – and we find that a degree of intervality is a by-product of this topological feature. We go on to propose a version of the preferential preying model, amended to take account of phylogenetic constraints, which generates realistic values of intervality without fitting additional parameter values.

Taken together, we believe these observations call into question the concept of a niche dimension, whether as an operationally useful construct for food-web modelling, or as a reality to be uncovered in nature. We conclude by discussing what these findings might mean for our understanding of food-webs and other complex networks.

II. RESULTS

A. Intervality

Let us consider a directed network with $S$ nodes and $L$ edges, defined by the $S \times S$ adjacency matrix $A$. As we have said, in the case of a trophic network the nodes are species and the edges represent predation. The in- and out-degrees of node $i$ are $k_i^{in} = \sum_j A_{ji}$ and $k_i^{out} = \sum_j A_{ij}$, and correspond to the numbers of prey and predators of species $i$, respectively; and the mean degree is $\langle k \rangle = L/S$. Basal species, or autotrophs, have no in-coming edges, and are thus represented by nodes with $k_i^{in} = 0$. We shall denote with $B$ the number of basal nodes in a given network.

The top left panel of Fig.1B shows the adjacency matrix for the food web of Mondego Estuary, on the Atlantic coast of Portugal, with species ordered so as to maximize intervality $\xi$ (see Methods). Since there are $S!$ possible orderings of the columns it is not feasible in general to perform an exhaustive search for the most interval one. Therefore, we proceed as Stouffer et al. and use a simulated annealing (SA) algorithm – described in the Methods section. For comparison, in the top right panel of Fig.1B we show the best ordering for a randomisation of the network which preserves $k_i^{in}$ and $k_i^{out}$ for each node. These plots readily reveal that the empirical food web is remarkably more interval than its randomised counterpart. Observe, however, that the random network exhibits a non-zero level of intervality, since this magnitude is defined for the most interval ordering out of a great many possible choices. It is therefore always necessary to compare empirical values of intervality to the corresponding random expectations in a null model, in order to determine the significance of this measurement.

In the bottom left panel of Fig.1B we show the same Mondego Estuary adjacency matrix, but this time an ordering has been found which maximises predator-intervality, $\eta$: the extent to which the predators of a given prey species are contiguous (as above, the same procedure has been applied to an ensemble of network randomisations). The real food web is significantly more predator-interval, while the random graph has some degree of predator-intervality (there is a preponderance of vertical intervals). Interestingly, the ordering of maximum predator-intervality is different from the one yielding the highest prey-intervality, as is obvious to the naked-eye upon inspection of the two patterns. In fact, we shall see that predator-intervality seems to be as general and non-trivial a feature of food webs as the oft-cited intervality, to distinguish the two concepts.

We now extend this analysis to the entire database of directed networks we are going to use throughout this work (listed in Table I), which includes food webs, gene transcription networks, metabolic networks, networks of cellular signalling, the neural network of C. elegans, a word-adjacency network, and a US airport recommendation network. The two panels of Fig.2 show average prey-intervality $\xi$, and average predator-intervality $\eta$, for all networks in the database, compared to their randomisations. The first observation is that, as we can see in Fig.1 for Mondego Estuary, and more generally in Table I, prey-intervality is indeed a common feature of food webs. In all cases the empirical values lie several standard deviations from the random expectations (z-score values greater than 3 in all but 4 cases out of 46 food webs). Similarly, food webs also exhibit significant levels of predator-intervality $\eta$. In all cases, the orderings which maximise prey- and predator-intervality are different. What is perhaps more interesting is that intervality...
does not seem to be limited to food webs; indeed, almost all the other networks in our database show a similar trend, with levels of both interivalities similar to those observed in food webs.

Since the original observation that food webs tended to be interval, the existence of a hidden “niche dimension” has been assumed to be at the root of this deviation from randomness. However, this explanation fits ill with the fact that food webs and other kinds of network exhibit both prey- and predator- intervality to similar degrees of significance. It would seem, rather, that a more general reason must exist which can account for this topological feature.

B. Trophic coherence

One of the most striking characteristics of food webs – perhaps the first that springs to mind upon contemplating even a child’s drawing of an ecosystem – is the existence of a trophic structure. At the base there are plants, which are consumed by herbivores, which in turn might be preyed upon by omnivores or primary carnivores, and so the biomass flows from producers all the way up to top predators. Ecologists quantify the position of a species in a food web with its trophic level, such that basal species have \( l = 1 \), herbivores \( l = 2 \), etc. More generally – following Levine\cite{27} – one can define the (non-integer) trophic levels in a recursive, self-consistent way: the trophic level of a given species is equal to the average level of its prey plus one unit. That is,

\[
l_i = \frac{1}{k_{in}} \sum_j A_{ij} l_j + 1
\]

(1)

If \( k_{in} > 0 \), or \( l_i = 1 \) if \( k_{in} = 0 \). Observe that this definition is easily generalizable to other types of directed network, and each node can be assigned a trophic level by simply solving the set of linear equations given by Eq. (1).\cite{29} The only requirements for all nodes to have a well-defined level are that there should exist at least one basal node \( (B > 0) \), and that every non-basal node must be on at least one path including a basal node.\cite{29} In fact, this definition of trophic level is similar to other measures of centrality, such as PageRank, with the difference that the “+1” term establishes a natural hierarchy.\cite{29}

We have recently shown that the trophic structure of a directed network can be characterised by a degree of order we call trophic coherence.\cite{29} For this we define a variable \( x \) for each edge, as \( x_{ij} = l_i - l_j \), and consider the distribution of \( x \) over the \( E \) edges of a network. The mean is \( \langle x \rangle = 1 \) by definition, and the homogeneity of the distribution is the trophic coherence of the network. We can therefore quantify this feature simply with the standard deviation of \( p(x) \), which we refer to as an incoherence parameter: \( q = \sqrt{\langle x^2 \rangle} - 1 \). A highly coherent network \( (q \approx 0) \) is one in which the nodes fall into clear (almost integer) trophic levels, while a more random system is less coherent \( (q > 0) \). We have shown that trophic coherence is key to the linear stability of food webs and, since it can invert the usually positive relationship between diversity and stability, might be the solution to Robert May’s famous paradox.\cite{29,31,32} Trophic coherence has subsequently been found to play an important role in directed networks of many kinds, including those of neurons, genes, metabolites, cellular signalling, words, P2P, trade and transportation, and this feature is intimately related to cycles and feedback loops, graph eigenspectra and the ubiquity of ‘qualitatively stable’ systems.\cite{29,32}

Elsewhere in this issue, Klaise & Johnson show that trophic coherence also determines the extent and duration of spreading processes such as epidemics or neuronal cascades.\cite{33}

Networks with tunable trophic coherence can be generated with the recently proposed preferential preying model (PPM),\cite{29} which works as follows. We begin with \( B \) initial nodes (basal species) and no edges. New nodes (consumer species) are added sequentially to the system until a total of \( S \) nodes is reached. When a node enters the system its in-neighbours (prey) are awarded from among available nodes (those already in the network) in the following way: the first prey species is chosen randomly, and the rest are chosen with a probability that decays exponentially with the absolute trophic distance to their initial prey (i.e. with the absolute difference of trophic levels between its first prey and the subsequent ones). This probability is set by a parameter \( T \) that determines the degree of trophic specialization of consumers, and normalised so as to produce an expected number of edges \( L \). The lower the value of \( T \) (while \( T > 0 \)), the more the network will have a coherent trophic structure \( (q \approx 0) \). From the perspective of evolutionary ecology, this simply means that if a given predator is good

Figure 2. Average z-score for prey (ξ) and predator (η) interivalities for different type of networks (with respect to a null model consisting in randomisations of each network that preserve both \( k_{in} \) and \( k_{out} \) for each node). Empirical values of the z-score have been averaged over all networks in each category. It is evident that empirical values lie well outside the random expectation, with high values of z-scores in all cases. Intervality is therefore a non-trivial feature of many kinds of complex network.
at consuming species X, then its other prey are likely to have similar trophic levels to that of X, as is observed in nature.\footnote{\textsuperscript{21}}

Food webs tend to be very significantly coherent, and this is key to their stability and other properties. However, niche-based models are not able to reproduce this feature, and generate networks which are barely more coherent than random graphs.\footnote{\textsuperscript{20}} This shows that intervality is not a sufficient condition for trophic coherence. Is it possible, though, that trophic coherence induces intervality? Figure 3A displays an example of a maximally coherent network (\(q = 0\)), obtained with a low T, while figure 3B is an instance of a less coherent network (\(q > 0\)), obtained with a high T (see also Fig 4). Beside each of them appears its adjacency matrix, where species have been ordered in such a way as to maximise prey intervality, \(\xi\). The maximally coherent network is also perfectly interval, while the incoherent one has gaps which cannot be eliminated by changing the arrangement of nodes in any way. This trend is more clearly illustrated in Fig 4, which shows that both prey- and predator-intervalities decrease monotonically with T.

For an intuitive understanding of why intervality emerges in the presence of trophic coherence, imagine a simple food web consisting only of \(n_a\) predators that prey upon a different set of \(n_p\) species. It will be more likely that there is an ordering of the prey such that \(\xi\) is high – or one of predators such that \(\eta\) is – if \(n_a\) and \(n_p\) are small (i.e. if the adjacency matrix contains few rows and columns). If a food web is perfectly stratified (\(q = 0\)) so that species on level \(l\) are only consumed by those on level \(l + 1\), then the network can be seen as a superposition of many of these (independent) simple situations (A will have a “modular” or block structure). Given that at each pair of levels the number of species is significantly smaller than the total number \(S\), a global ordering can be found yielding as good a \(\xi\) as in the simple example – or a different ordering for as good an \(\eta\). Networks that are thus highly coherent will have higher values of both intervalities than those lacking this structure.

C. Topological features contributing to Intervality

The reasoning described above suggests that trophic coherence should not be the only feature related to intervality. Indeed, as illustrated below, trophic coherence can only account for a fraction of the total intervality observed in empirical networks.

In general, any property which had the effect of creating modularity – i.e. de-coupling certain non-zero matrix elements from others, in the sense that their rows or columns could be shuffled without affecting the ordering of other elements – would be conducive to higher levels of intervality. In other words, we should expect a high prey-intervality in networks in which nodes sharing any in-neighbours tended to share a high proportion of them; and the same goes for predator-intervality and shared out-neighbours. To capture this feature, we can define in-complementarity \(c_{in}\) as the mean number of shared in-neighbours over all pairs of nodes with any shared in-neighbours; and, similarly, out-complementarity \(c_{out}\) as the mean number of shared out-neighbours over all pairs of nodes with any shared out-neighbours.

We have measured the relationships between intervality and several topological properties (including \(S\), \(\langle k\rangle\), \(q\) and complementarities \(c_{in}\) and \(c_{out}\)). The correlation between \(\xi\) and \(c_{in}\) has a Pearson coefficient of \(r = 0.83\), precisely the same as that of \(\eta\) and \(c_{out}\). In other words, complementarity – i.e. the proportions of shared in-neighbours and of shared out-neighbours – accounts for approximately 70% of the variance in prey- and predator-intervalities, respectively, across a broad range of empirical networks. The correlations with trophic coherence itself are lower but still significant (\(r \approx 0.55\)). The mean degree, \(\langle k\rangle\), and the number of nodes, \(S\), are both negatively correlated with both kinds of intervality (\(r \approx -0.67\) and \(r \approx -0.45\), respectively), as we would expect from the reasoning above.

D. Modelling networks with coherence and intervality

Most past attempts to explain food-web intervality have focused on identifying the biological factors which might yield a niche dimension, such as body size.\footnote{\textsuperscript{18,19,37}} However, the work by Cattin \textit{et al}.\footnote{\textsuperscript{21}} is different in that it considers the effects of phylogeny, albeit within the framework of a niche-based model. Crucially, the ex-

![Figure 3](image_url)

Figure 3. A: Network generated with the preferential preying model (PPM) with \(S = 12\), \(B = 2\) and \(T = 0\) (the vertical position of each node reflects its trophic level) and its corresponding adjacency matrix, with rows and columns ordered to maximize prey-intervality (\(\xi = 1\)). B: As A but for a similar randomised network, leading to a smaller intervality \(\xi = 0.90\).
Figure 4. Left panel: Incoherence parameter $q$ as a function of the tunable parameter $T$ as obtained in the preferential preying model (PPM)\textsuperscript{29} ($N = 31, L = 68$ as in the Chesapeake bay food web\textsuperscript{35,36}). For each $T$, the reported values (crosses) are averages over 100 independent network realisations. As $T$ increases the generated networks are more incoherent. The Central (Right) panel shows prey ( predator) intervality $\xi$ ($\eta$), as measured in the same networks. Discontinuous lines are just interpolations.

existence of phylogenetic relationships between species – i.e. of common ancestry – was not modelled as a niche axis, but as a “diet overlap” for certain pairs of species. The thinking is that species which are closely related to each other (phylogenetically similar) will tend to share many of their prey. Cattin and colleagues found that phylogenetic constraints, modelled in this way, increase intervality\textsuperscript{21} – and subsequent empirical work has reinforced this connection\textsuperscript{85,89}. These results are in keeping with the view that any mechanism with the effect of making nodes topologically similar (such as close relatives having many shared prey or predators, in the case of food webs) will also lead to seemingly non-random levels of intervality. But since topologically similar nodes can exist for many different reasons, it is not just food webs which exhibit intervality, but directed networks of many different kinds.

In the case of food webs, we know that both trophic coherence and a phylogenetic signal are present. To explore whether these properties might be sufficient to explain the observed intervality, we extend the PPM to account for the effect of phylogenetic relations in a simple parameter-free way. For this we draw inspiration from the “nested hierarchy” model of Cattin \textit{et al.}\textsuperscript{21} and define a version of the PPM in which new nodes are assigned in-neighbours (prey) according not only to their trophic levels, as before, but also with a preference for nodes which already share out-neighbours (predators) whenever possible (see Methods for a detailed description of this model). Figure 5 illustrates the performance of this modified PPM model compared to its original version for the Chesapeake bay food web. In particular, we generate networks with the modified model fixing $N, L$ and $B$ as in the empirical network, and choose the optimal value of $T$ (vertical dashed line) that leads to a value of $q$ as close as possible to the empirical one (horizontal dotted line in Fig. 5). For this same value of $T$, the obtained intervality $\xi$ fits quite well with its empirical counterpart (z-score $-0.22$).

Observe, that while the dependence of the incoherence parameter, $q$, on the only parameter, $T$, remains as in the original version, the intervality $\xi$ obtained with the modified model is higher and much closer to the empirical value. Similar results are obtained, in general, for other trophic networks (e.g. for Narragansay the z-score is $-0.43$, for St. Marks $-1.06$, and for St. Martin $-0.89$), while a few exceptions exist (e.g. for the Everglades marshes we obtain $-6.52$, indicating that in this case our model does not account for the observed intervality). Worse agreement is obtained for predator intervality $\eta$; in the example of Figure 5 the corresponding z-score is $-2.04$. The reason for this is that intervality is strongly influenced by the degree sequences (see the definitions of $\xi$ and $\eta$ and the PPM model –as well as its modified variant– fit the empirical in-degree distribution, but not the out-degree one, which is the one affecting $\eta$ the most.

The justification for the proposed combination of trophic specialization and phylogenetic relations has a clear meaning only in the case of food webs. We leave as an open question to determine which mechanisms akin to these ones are at work in systems other than food webs, or whether there are different mechanisms which have a similar effect on intervality.

III. DISCUSSION

The niche concept has been central to ecology for many decades\textsuperscript{10,11} McInerny & Etienne, however, have recently called into question its prominent role, arguing that the ecological niche is often not well defined, can create needless confusion, and may not even be necessary for ecology\textsuperscript{12,13}. The results we describe here are relevant mainly for the niche idea as used in food-web modelling. This approach has its origins in the intervality of food webs first reported by Cohen\textsuperscript{12} and became firmly established thanks to the success of Williams & Martinez’s celebrated niche model\textsuperscript{15} All subsequent “structural” food-web models –i.e. those which attempt to generate realistic networks without modelling evolutionary or population dynamics explicitly– have been based on
the niche model. Going on this, attempts have been made to identify the niche axis with some biological magnitude, such as body size, but no such measure has ever been found to align completely with the orderings of maximum intervality. It is worth noting that Rossberg et al. were able to generate interval networks in an \( N \)-dimensional niche space provided phylogenetic relationships were present, thus proving that, contrary to general belief, a one-dimensional space was not a necessary condition for the emergence of intervality. In their model species evolution is represented as a random walker in the \( N \)-dimensional niche space, and over time species can go extinct or speciate into new ones. As long as new species are close to the parent ones in niche space, the resulting networks are highly interval.

We propose here that the intervality of food webs does not justify invoking hidden or niche dimensions. We have shown that food webs exhibit as much predator-intervality as prey-intervality, which seems incompatible with the interpretation of niche-based models. Moreover, we find that many different kinds of biological and artificial networks are as significantly interval as food webs. This means that either these various networks were all assembled according to the stringent constraints of a niche-axis analogue—which appears highly unlikely— or significant intervality need not be the hallmark of hidden dimensions. The latter explanation is further supported by the strong correlation between intervality and the average proportion of shared neighbours (complementarity) found across all networks. Finally, a one-parameter, entirely niche-free network model, based on the “preferential preying” and the “nested hierarchy” models, which emulates trophic specialization and phylogenetic relations, yields similar levels of intervality as those observed in food webs and other networks.

There are two messages to be drawn from this paper. For ecologists, we cannot say whether our results should change the interpretation of the niche concept more broadly, but as regards food-web modelling we believe the case is clear for a new approach. The niche axis—like Ptolemy’s epicycles, phlogiston, caloric, luminumiferous aether, and other constructs eventually found lacking in empirical support—should be abandoned. This does not mean that Williams & Martinez’s idea of food-web complexity arising from simple rules is wrong, only that we must update our concepts of these rules.

For the study of complex networks more generally, these results show that there are still open questions about the basic mechanisms behind their formation. Shared neighbours explain 70% of the variance in intervality, but whence the rest? Do other properties induce such apparent order, or are there, after all, hidden hierarchies underlying the architectures of certain complex systems? The high significance of topological properties such as intervality and trophic coherence found in many different kinds of system, from neural networks to trading relations, and word graphs to gene regulatory networks, suggests that we, like Darwin admiring the complexity of his entangled bank, have yet much to learn from ecosystems.

IV. METHODS

Measuring intervality. Given a food web and a specified ordering of species \( O \), we define the prey-intervality, \( \xi^O \), of each predator species \( i \) as the largest number of its prey to lie on an unbroken interval divided by the total number of prey it has. The prey-intervality of the food web itself is defined as that of the ordering with the highest prey-intervality: \( \xi = \max(\xi^O) \). Given that the search cannot be made exhaustively, we implement a standard simulated annealing algorithm. For this, one begins with a random ordering \( O_1 \) and measures its prey-intervality \( \xi^{O_1} \); then two species are randomly chosen and their positions in the ordering exchanged, yielding a new ordering \( O_2 \). The new prey-intervality \( \xi^{O_2} \) is calculated; and the permutation is accepted with probability

\[
P_{1 \rightarrow 2} = \min \left\{ \frac{\exp \left( \frac{\xi^{O_2} - \xi^{O_1}}{p} \right)}{1}, 1 \right\},
\]

where \( p \) is a temperature-like parameter, while with probability 1-\( P_{1 \rightarrow 2} \) the permutation is rejected and the algorithm is iterated. As is standard, we gradually lower \( p \) (starting with \( p = 0.01 \) and then decreasing by a factor 0.99 after every 1000 proposed permutations), and perform several runs from different initial conditions.

Figure 5. Figure analogous to Fig. 4 for the Chesapeake Bay food web, but employing the modified—rather than the original—version of the PPM, which includes a preference to choose prey which already share some predators (or prey). Results for the original PPM are also reported (light colors) for the sake of comparison. Dotted horizontal lines indicate the empirical values as measured in the real food web, while vertical ones stand for the value of \( T \) providing the best fit to \( q \). Observe that, while trophic coherence is not significantly changed by this modification to the model, the levels of intervality (\( \xi \) and \( \eta \)) are clearly augmented by the effect of a phylogenetic signal and become closer to the real values.
for each measurement. In all cases, the algorithm converges to a unique value of ξ for a given network. The predator-intervality, η, is obtained in an analogous way. Observe that this measure is slightly different from the one defined by Stouffer et al. [25], which is based on the “generalized niche model”; but both measures are strongly correlated (we have found a Pearson’s coefficient r = 0.82 between the two measurements over our network dataset).

Measuring complementarity. Given a food web we define its matrix of shared prey as M = AAᵀ, where A is the adjacency matrix. The element mᵢⱼ is the number of prey that are shared by species i and j. Similarly, the matrix of shared predators is M = AᵀA. The in-complementarity between i and j is defined as wiᵢⱼ = mᵢⱼ/ max(kᵢᵢ ∩ kᵢⱼ), and similarly, for the out-complementarity wiᵢⱼ = mᵢⱼ/ max(kᵢⱼ ∩ kᵢᵢ). By averaging over all possible pairs with at least one common shared in/out neighbour, we determine the overall complementarities cin and cout, respectively.

Modified Preferential preying model with phylogenetic constraints. The original preferential preying model was designed to generate networks with tunable trophic coherence. Here we introduce an additional mechanism implementing the idea that a predator is likely to choose prey which are similar, i.e. that share other predators or other prey. In the standard PPM one starts with B basal nodes and adds progressively new species, up to a total of S. Each one selects its prey from existing network nodes, following these rules: i) The in-degree (kᵢᵢ) of species i is selected from a Beta distribution, so as to obtain on average the same number of links L as the empirical network to be modelled. ii) The first prey species (l) is randomly selected from the available ones. iii) Subsequent prey (j) are chosen with a probability Pᵢⱼ that decays with the trophic distance between j and l:

Pᵢⱼ ∼ exp(- |sᵢᵢ - sᵢⱼ|)

where T is a “temperature” parameter that sets the degree of trophic specialization. In this version of the PPM, in order to include an additional preference for phylogenetically related species as done by Cattin et al. [21] we modify rule iii) as follows. First we chose a prey j as above and consider all other possible prey with the same trophic level (plus/minus 1%) and from this group we randomly select one among the subgroup that already has some predator that shares prey with the existing prey of i. If there is no species obeying these constraints, then j is selected.

ACKNOWLEDGMENTS

Acknowledgments – We acknowledge the Spanish-MINECO grant FIS2013-43201-P (FEDER funds) for financial support.

REFERENCES

1. C. Darwin, On the Origin of Species (John Murray, London, UK, 1859).
2. C. S. Elton, Animal Ecology (Sidgwick and Jackson, London, 1927).
3. J. E. Cohen and C. M. Newman, “A stochastic theory of community food webs I. models and aggregated data,” Proc. R. Soc. London Ser. B. 224, 421–448 (1985).
4. J. E. Cohen, Food Webs and Niche Space (Princeton Univ. Press, Princeton, New Jersey, 1976).
5. S. L. Pimm, The Balance of Nature? Ecological Issues in the Conservation of Species and Communities (The University of Chicago Press, Chicago, 1991).
6. B. Drossel and A. J. McKane, Modelling Food Webs, in A Handbook of Graphs and Networks: From the Genome to the Internet (Wiley-VCH, Berlin, 2003).
7. J. A. Dunne, The network structure of food webs, in Ecological Networks: Linking Structure to Dynamics in Food Webs, edited by M. Pascual and e. J.A. Dunne (Oxford University Press, Oxford, UK, 2006).
8. R. V. Solé and J. Bascompte, Self-Organization in Complex Ecosystems (Princeton University Press, Princeton, USA, 2006).
9. J. Camacho, R. Guimerà, and L. A. N. Amaral, “Robust patterns in food web structure,” Phys. Rev. Lett. 88, 228102 (2002).
10. R. V. Solé and M. Montoya, “Complexity and fragility in ecological networks,” Proc. R. Soc. Lond. B 268, 20392041 (2001).
11. J. M. De Vos, L. N. Joppa, J. L. Gittleman, P. R. Stephens, and S. L. Pimm, “Estimating the normal background rate of species extinction,” Conservation Biology 29, 452–462 (2015).
12. J. E. Cohen, “Food webs and the dimensionality of trophic niche space,” Proc. Natl. Acad. Sci. USA 74, 4560–4563 (1977).
13. J. Grinnell, “The niche-relationships of the California Thrasher,” Auk 34, 427–433 (1917).
14. G. E. Hutchinson, “Concluding remarks,” Cold Springs Harbor Symp. Quant. Biol. 22, 415–427 (1957).
15. R. J. Williams and N. D. Martinez, “Simple rules yield complex food webs,” Nature 404, 180–183 (2000).
16. R. J. Williams and N. D. Martinez, “Success and its limits among structural models of complex food webs,” Journal of Animal Ecology 77, 512–519 (2008).
17. C. Guill and B. Drossel, “Emergence of complexity in evolving niche-model food webs,” Journal of Theoretical Biology 251, 108120 (2008).
18. D. B. Stouffer, J. Camacho, and L. A. N. Amaral, “A robust measure of food web interactivity,” Proc. Natl. Acad. Sci. USA 103, 19015–19020 (2006).
19. A. E. Zook, A. Eklof, U. Jacob, and S. Allesina, “Food webs: Ordering species according to body size yields high degree of interactivity,” Journal of Theoretical Biology 271, 106–113 (2011).
20. I. Patricio, “Network analysis of trophic dynamics in south florda ecosystems, fy 99: The graminoid ecosystem.” Master’s Thesis. University of Coimbra, Coimbra, Portugal (2000).
21. M. F. Cattin, L. F. Bersier, C. Banasek-Richter, R. Baltenesperger, and J. P. Gabriel, “Phylogenetic constraints and adaptation explain food-web structure,” Nature 427, 835–9 (2004).
22. M. Girvan and M. E. J. Newman, “Community structure in social and biological networks,” Proc. Natl. Acad. Sci. USA 99, 7821–7826 (2002).
23. S. Johnson, V. Domínguez-García, and M. A. Muñoz, “Factors determining nestedness in complex networks,” PloS ONE 8, e74025 (2013).
24. D. B. Stouffer, J. Camacho, R. Guimerà, C. A. Ng, and L. A. N. Amaral, “Quantitative patterns in the structure of model and empirical food webs,” Ecology 86, 13011311 (2005).
25. S. Allesina, D. Alonso, and M. Pascual, “A general model for food web structure,” Science 320, 658–661 (2008).
26. S. Johnson, V. Domínguez-García, L. Donetti, and M. A. Muñoz, “Trophic coherence determines food-web stability,” Proc. Natl. Acad. Sci. USA 111, 17923–17928 (2014).
27. S. Levine, “Several measures of trophic structure applicable to complex food webs,” J. Theor. Biol. 83, 195–207 (1980).
28. V. Domínguez-García, S. Pigolotti, and M. A. Muñoz, “Inherent directionality explains the lack of feedback loops in empirical networks,” Scientific Reports 4, 7497 (2014).
29. S. Johnson and N. S. Jones, “Spectra and cycle structure of trophically coherent graphs,” arXiv:1505.07332 (2015).
30. V. Domínguez-García and M. A. Muñoz, “Ranking species in mutualistic networks,” Scientific reports 5 (2015).
31. R. M. May, “Will a large complex system be stable?” Nature 238, 413 (1972).
32. R. S. McCann, “The diversity-stability debate,” Nature 405, 228–33 (2000).
33. J. Klaise and S. Johnson, “From neurons to epidemics: How trophic coherence affects spreading processes,” arXiv:1603.00670 (2016).
This is the model we use to generate coherent networks in this work; however, we note that Klaise & Johnson propose a slightly different version of the PPM, the main difference being that at high T their model limits in random graphs instead of acyclic cascade model networks.

H. E. Ulanowicz and D. Baird, “Nutrient controls on ecosystem dynamics: the Chesapeake mesohaline community,” Journal of Marine Systems 19, 159 – 172 (1999).

L. G. Abarca-Arenas and R. E. Ulanowicz, “The effects of taxonomic aggregation on network analysis,” Ecological Modelling 149, 285 – 296 (2002).

J. Capitán, A. Arenas, and R. Guimerá, “Degree of intervality of food webs: From body-size data to models,” Journal of Theoretical Biology 334, 35–44 (2013).

E. Eklöf and D. B. Stouffer, “The phylogenetic component of food web structure and intervality,” Theor Ecol (2015).

D. Mouillot, B. Krasnov, and R. Poulin, “High intervality explained by phylogenetic constraints in host-parasite webs,” Ecology 89(7), 20432051 (2008).

G. E. Hutchinson, An Introduction to Population Biology (Yale Univ Press, New Haven, CT, USA., 1978).

R. K. Colwell and T. F. Rangel, “Hutchinson’s duality: The once and future niche,” Proc. Natl. Acad. Sci. USA 106, 19651 (2009).

R. E. Ulanowicz and R. Etienne, “Pitch the niche: Is the niche a useful concept in ecology or species distribution modelling?” J. Biogeogr. 39, 20962102 (2012).

G. McOerny and R. Etienne, “Pitch the niche: Taking responsibility for the concepts we use in ecology and species distribution modelling,” J. Biogeogr. 39, 21122118 (2012).

B. B. Stouffer, E. L. Rezende, and L. A. N. Amaral, “Characterising foodwebs in two new zealand streams,” New Zealand Journal of Marine and Freshwater Research 32, 271–286 (1998).

R. M. Thompson and C. R. Townsend, “Energy availability, spatial heterogeneity and ecosystem size predict food-web structure in stream,” Oikos 108, 137148 (2012).

R. M. Thompson, E. D. Edwards, A. R. McIntosh, and C. R. Townsend, “Allocation of effort in stream food-web studies: the best compromise?” Marine and Freshwater Research 52, 339345 (2001).

R. E. Ulanowicz, C. Bondavalli, and M. Egnotovich, “Spatial and temporal variation in the structure of a freshwater food web,” Network Analysis of Trophic Dynamics in South Florida Ecosystems, FY 97: The Florida Bay Ecosystem. (1998).

D. Mason, “Quantifying the impact of exotic invertebrate invaders on food web structure and function in the great lakes: A network analysis approach,” Interim Progress Report to the Great Lakes Fisheries Commission- yr 1 (2003).

M. E. Monaco and R. E. Ulanowicz, “Comparative ecosystem trophic structure of three u.s mid-Atlantic estuaries,” Marine Ecology Progress Series 161, 239–254 (1997).

S. Opitz, “Trophic interactions in Caribbean coral reefs,” ICLARM Tech. Rep. 43, 341 (1996).

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

R. Warren, “Spatial and temporal variation in the structure of a freshwater food web,” Oikos 55, 299–311 (1989).

R. R. Christian and J. J. Luczkovich, “Organizing and understanding a winter’s Seagrass foodweb network through effective trophic levels,” Ecol. Model. 117, 99–124 (1999).

L. Goldwasser and J. A. Roughgarden, “Construction of a large Caribbean food web,” Ecology 74, 1216–1233 (1993).

U. Jacob, A. Thierry, U. Brose, W. Arntz, S. Berg, T. Brey, I. Fieberg, T. Jonsson, K. M. Mänttänen, C. Milham, O. Petchey, J. Riede, and J. Dunne, “The role of body size in complex food webs,” Advances in Ecological Research 45, 181–233 (2011).

J. D. Duch and A. Arenas, “Community identification using extremal optimization,” Physical Review E 72, 027104 (2005).

J. Sanz, J. Navarro, A. Arbós, C. Martín, P. C. Marijuan, and Y. Moreno, “The transcriptional regulatory network of mycobacterium tuberculosis,” PLoS ONE 6, e2178 (2011).

H. Yu and M. Gerstein, Proc. Natl. Acad. Sci. USA.

A. Ma’ayan, G. A. Cecchi, J. Wagner, A. R. Raob, R. Iyengar, and G. Stolovitzky, “Ordered cyclic motifs contribute to
dynamic stability in biological and engineered networks,” PNAS 105, 19235–19240 (2008).

79 C. Rodríguez-Caso, B. Corominas-Murtra, and R. V. Sol, “On the basic computational structure of gene regulatory networks,” Mol. BioSyst. 5, 1617–1629 (2009).

80 D. J. Watts and S. H. Strogatz, “Collective dynamics of ‘small-world’ networks,” Nature 393, 440–442 (1998).
### Table I: List of empirical networks considered in this work. The first column shows their names, the fourth their types, and the last one references to the literature. The z-scores of intervalities $\xi$ and $\eta$ are computed with respect to a null model consisting in many randomisations of each network which preserve both in- and out-degree sequences. The number of “stars” after each z-score indicates whether the z-score is between 1 and 2 (**), 2 and 3 (***), or larger than 3 (***)

| Name                               | S    | Type   | $\xi$ z-score | $\eta$ Z-score | ref       |
|------------------------------------|------|--------|---------------|----------------|-----------|
| Akatore Stream                     | 85   | Stream | 2.67          | 2.650          | (***)     |
| Florida Bay (dry season)           | 122  | Marine | 14.75         | 68.536         | (**)      |
| Florida Bay (wet season)           | 122  | Marine | 14.75         | 55.779         | (**)      |
| Benguela Current                   | 29   | Marine | 7.00          | 11.492         | (**)      |
| Berwick Stream                     | 79   | Stream | 3.04          | 9.206          | (**)      |
| Blackrock Stream                   | 87   | Stream | 4.31          | 6.350          | (**)      |
| Bridge Brook Lake                  | 25   | Lake   | 4.28          | 14.676         | (**)      |
| Broad Stream                       | 95   | Stream | 5.95          | 15.423         | (**)      |
| Scotch Broom                       | 85   | Terrestrial | 2.62  | 9.206         | (**)      |
| Florida Bay (dry season)           | 102  | Stream | 6.83          | 13.678         | (**)      |
| Florida Bay (wet season)           | 249  | Marine | 13.31         | 32.603         | (**)      |
| Carpenteria Salt Marsh Reserve     | 128  | Marine | 4.23          | 24.123         | (**)      |
| Cayman Islands                     | 261  | Marine | 4.43          | 52.739         | (**)      |
| Chesapeake Bay                     | 31   | Marine | 2.19          | 7.997          | (**)      |
| Crystal Lake (control)             | 20   | Lake   | 2.55          | 5.454          | (**)      |
| Crystal Lake (delta)               | 20   | Lake   | 1.65          | 2.236          | (**)      |
| Cuba                               | 261  | Marine | 6.89          | 18.791         | (**)      |
| Cypress (dry season)               | 65   | Terrestrial | 6.89  | 18.791        | (**)      |
| Cypress (wet season)               | 65   | Terrestrial | 6.75  | 18.791        | (**)      |
| Dempsters Stream (autum)           | 86   | Stream | 4.83          | 10.116         | (**)      |
| El Verde Rainforest                | 155  | Terrestrial | 9.74  | 56.909         | (**)      |
| Everglades (G. Marshes)            | 63   | Terrestrial | 9.79  | 51.266         | (**)      |
| Florida Bay                        | 122  | Marine | 14.48         | 55.670         | (**)      |
| G. Marshes (dry)                   | 63   | Terrestrial | 9.79  | 51.266         | (**)      |
| G. Marshes (wet)                   | 63   | Terrestrial | 9.79  | 51.266         | (**)      |
| Grassland                          | 61   | Terrestrial | 1.59  | 2.917          | (**)      |
| Healy Stream                       | 96   | Stream | 6.60          | 11.242         | (**)      |
| Jamaica                            | 263  | Marine | 15.61         | 82.760         | (**)      |
| Kuyburg Stream                     | 98   | Stream | 6.42          | 8.142          | (**)      |
| Little Rock Lake                   | 92   | Lake   | 10.84         | 54.022         | (**)      |
| Lough Hyne                         | 349  | Marine | 14.66         | 67.464         | (**)      |
| Mangrove Estuary (dry season)      | 91   | Marine | 12.65         | 36.163         | (**)      |
| Mangrove Estuary (wet season)      | 91   | Marine | 12.65         | 36.163         | (**)      |
| Michigan Lake                      | 35   | Lake   | 3.69          | 24.402         | (**)      |
| Mondego Estuary                    | 42   | Marine | 6.64          | 13.237         | (**)      |
| Narragansett Bay                   | 31   | Marine | 3.65          | 6.841          | (**)      |
| Caribbean Reef                     | 50   | Marine | 11.12         | 11.179         | (**)      |
| N.E. Shelf                         | 79   | Marine | 17.76         | 24.212         | (**)      |
| Skipwith Pond                      | 25   | Lake   | 7.88          | 2.504          | (**)      |
| St. Marks Estuary                  | 48   | Marine | 4.60          | 21.388         | (**)      |
| St. Martin Island                  | 42   | Terrestrial | 4.88  | 12.730         | (**)      |
| Stony Stream                       | 109  | Stream | 7.61          | 13.525         | (**)      |
| Troy Stream                        | 78   | Stream | 2.32          | 5.686          | (**)      |
| Weddell Sea                        | 483  | Marine | 31.81         | 120.062        | (**)      |
| Ythan Estuary                      | 82   | Marine | 4.82          | 7.713          | (**)      |
| Biological                         |      |        |               |               |           |
| C. Elegans metabolic               | 453  | TRN    | 4.50          | 7.726          | (**)      |
| E. Coli transcription              | 1033 | TRN    | 2.50          | 51.538         | (**)      |
| M. musculus transcription          | 73   | TRN    | 1.62          | 3.908          | (**)      |
| Mammalian signalling               | 599  | Cell signalling | 2.34  | 21.686        | (**)      |
| B. Subtilis transcription          | 814  | TRN    | 1.69          | 2.305          | (**)      |
| M. Tuberculosis transcription      | 1624 | TRN    | 1.98          | -9.637         | (**)      |
| Other Networks                     |      |        |               |               |           |
| C. Elegans neural                  | 297  | neural | 7.90          | 22.429         | (**)      |
| Word adjacency                     | 50   | words  | 2.02          | 6.479          | (**)      |
| U.S.A airports                     | 1226 | airports | 2.13  | 49.869        | (**)      |

Table I: List of empirical networks considered in this work. The first column shows their names, the fourth their types, and the last one references to the literature. The z-scores of intervalities $\xi$ and $\eta$ are computed with respect to a null model consisting in many randomisations of each network which preserve both in- and out-degree sequences. The number of “stars” after each z-score indicates whether the z-score is between 1 and 2 (**), 2 and 3 (***), or larger than 3 (***).